International Journal of Science and Research (IJSR)
ISSN: 2319-7064
Impact Factor 2025: 7.089

Integrating Digital Forensics and Machine Learning for Retail Return Fraud Detection: A Comprehensive Review

Girish Kurkure[1], Dipita Dhande[2], Manisha Shirsath[3]

[1]Consultant, Fractal Analytics
Email: *girishkurkure111[at]gmail.com*

[2]Assistant Professor, Ashoka Centre of Business and Computer Studies Nashik, India
Email: *dipitadhande10[at]gmail.com*

[3]Associate Professor, Ashoka Centre of Business and Computer Studies Nashik, India
Email: *manishas.acbcs[at]aef.edu.in*

Abstract: ***Return fraud in retailing is the misuse of the return policy offered by stores for personal gain through the return of goods, which causes substantial losses to the retailers every year. Others include "wardrobing" which is returning heavily priced clothing after only wearing it once or twice, and returning stolen components, empty packaging, and even fake receipts. Indeed, return fraud causes retailers to hike even higher prices. The field of digital forensics provides an effective remedy by playing the role of a virtual detective that tracks down digital trails left by cases of fraudulent returns. Methods such as device fingerprinting, analysing metadata, and confirming shipping information are some of the ways that detect a pattern of fraud, including that of a fictitious identity and that of a device. Artificial Intelligence and Machine Learning technologies further amplify these functions by analysing massive data, recognizing anomalies in consumer patterns, and leveraging computer vision to verify the image of a product. The study emphasizes a shift from traditional physical checks to digital forensic strategies integrated with AI, addressing emerging fraud trends and reducing reverse logistics costs. By bridging gaps in current practices, this research provides actionable insights for academia and industry, supporting profit protection while maintaining customer satisfaction.***

Keywords: Return Fraud, Digital Forensics, Artificial Intelligence, Retail, Wardrobing

1. Introduction

E-commerce and omnichannel retailing, growing at a rapidly increasing rate, have drastically changed the purchasing patterns of people by providing increased convenience, variety, and easy return policies. Although these return policies have increased customer satisfaction and differentiation, they also incite many return products, hence becoming a material challenge in the retailing sector in terms of operational and financial implications, including sustainability. It has been reported that the total loss of return products every year worldwide is in the range of billions of dollars, with a large component of it attributed neither to consumer-related activities nor to return losses, but related to return-related fraudulent activities [1, 2].

Retail return fraud has been noted as a major emerging threat within contemporary supply chains. Retail return fraud can be defined as a broad category of fraudulent activities whereby a person uses the return policy method to gain financial or physical gain illegally. Types of return fraud include wardrobing, receipt fraud, return of stolen merchandise, item switching fraud, and friendly fraud whereby a customer fraudulently claims a defective or undelivered item of merchandise [3]. Unlike other kinds of retail fraud crimes, return fraud crimes can occur post-transaction. As a result of enhanced fraudster sophistication and the high volume of transactions and increased return time limits, more traditional methods used to control fraud have lost effectiveness.

Conventional methods of fraud analysis in the retail returns process are predominantly dependent on rules-based models, manual auditing, and threshold-based notifications. Despite ensuring some baseline control, these models have several in-built flaws. For instance, rules-based models do not tend to vary according to fluctuations in fraud patterns, manual audits entail significant manual effort and are susceptible to manual errors, and models dependent on threshold criteria tend to produce material numbers of false alarms, having negative ramifications for genuine clients themselves [4]. Further, all such models tend to be predominantly reactive in their functionality, flagging occurrences of fraud only after establishing material losses in the financial accounts of concerned entities. With the rising numbers of returns, there exists an imperative need for models that tend to provide intelligent and proactive support for analysing complex and nonlinear patterns of fraud.

Digital forensics in this case has assumed prominence as an investigative strategy in retail return-related fraud. Digital forensics is defined as the examination, analysis, interpretation, and assessment of digital evidence produced by information systems with the aim of determining accountability [5]. Digital evidence in retailing may involve transaction records, point-of-sale records, online transaction records on e-commerce platforms, customer behaviour records, device identity records, IP records, CCTV records, and supply-chain records. Digital forensics in return-related fraud assists in taking investigations beyond the superficial level where anomalies are usually sought and allows retailers to investigate in depth with the aim of determining root causes

of fraud and ascertaining compliance with fraud policies in retailing. However, retail data is huge and presents considerable difficulties against which conventional forensic procedures are put to the test.

Machine Learning (ML) has been identified as a significant facilitator for dealing with these issues, as it helps reproduce certain behaviours automatically from big data. Machine Learning algorithms, starting from classification algorithms, based on machine learning, aiming at recognizing patterns and behaviours, have been identified with a strong potential for distinguishing fraudulent transactions, hard to be detected by rules [6], [7]. In a returns process, machine learning algorithms help identify, based on past transactions, a difference between a legit and a fraudulent return, while new developments within Explainable Artificial Intelligence have encountered solutions regarding transparency, an issue identified for tools of such type, used for decision-making, concerning customers.

Despite increasing attention from the academic and industrial communities to fraud analytics, current literature on retail return fraud serves only to be piecemeal. Other researches might independently be concerned with financial fraud detection, e-commerce security, supply chain risk management, or general machine learning methods. On the other hand, the current literature on digital forensics mostly concentrates on such genres as cybercrime, incident response, or legal investigations, to the seeming neglect of retail-related processes for returns. There seems to be a clear need for more extensive literature reviews undertaken to combine digital forensic practices with machine learning algorithms for the detection of retail-related returns. More importantly, critical issues such as forensic preparedness, evidence integrity, privacy, or the ethics of automatic fraudulent detection systems are still uncharted territories or standalone [8].

This literature gap highlights the importance of a comprehensive review that approaches both digital forensics and machine learning from an interdisciplinary angle and seeks to integrate these two areas to provide a comprehensive approach towards retail return fraud. A comprehensive view is necessary to appreciate how the tenets of forensics can benefit fraud detection techniques that rely on machine learning algorithms and make them more reliable and robust from a legal and interpretability point of view.

As such, the aim of the paper is to provide an integrative literature review on retail return fraud. This literature review is targeted towards the involvement of digital forensics and machine learning. This literature aims to achieve the following: (i) It seeks to systematically investigate types or elements of retail return fraud. It aims to investigate how various approaches under digital forensics are applicable in retail returns. It seeks to investigate models and methods that have been deployed using machine learning for fraud analysis. It seeks to investigate the gaps and potential developments within the area. By integrating various streams of knowledge within different disciplines, this paper seeks to provide meaningful contributions towards developing effective fraud analysis frameworks.

The remaining portions of this paper are structured as follows. Section II includes a thorough literature survey conducted in retail return fraud, digital forensics, and machine learning-based fraud detection. Section III explains the methodological approach followed in the process of literature review and classification. Section IV includes discussions related to noteworthy findings and comparative insights gathered through the literature review process. Section V moves into open challenges and future research sections, followed by concluding remarks in Section VI.

2. Literature Review

This section undertakes a review of existing literature relevant to retail return fraud, digital forensics, and machine learning-based approaches to fraud detection. The literature is segregated into four thematic areas: A) retail returns and fraud issues in reverse logistics, B) traditional approaches to fraud detection and their limitations, C) digital forensics in fraud investigation, and D) machine learning techniques for fraud detection. This structured review highlights the evolution of research in each domain and identifies deficiencies that motivate the present study.

a) Retail Returns and Return Fraud in Reverse Logistics

Extensive research has considered reverse logistics as a crucial part of contemporary supply chains, especially under conditions of growing product return volumes. For instance, early work by Rogers et al. focused on the strategic and operational complexity brought about by reverse logistics to bear on inventory control, transportation, and cost management [1]. Liberal return policies, while being highly effective in improving customer satisfaction, have been shown in subsequent studies to drive up substantially both operational risk and financial exposure for retailers.

Return fraud has emerged as a particular subset of retail shrinkage and loss prevention research. Beck and Hopkins examined retail fraud trends and described return fraud as one of the fastest-growing loss categories-a trend driven by fraud via the exploitation of policies rather than traditional theft mechanisms [5]. Academic studies categorize return fraud further into activities such as wardrobing, receipt reuse, and fraudulent non-receipt claims, with all often challenging to distinguish from legitimate returns through standard transaction-level checks [4].

While most of the industry-focused research-such as reports prepared by the National Retail Federation and Deloitte-point out the magnitude of losses due to fraudulent returns and operational burdens on retailers [2], [3], these studies tend to be descriptive in nature, focusing on financial impact and policy implications without giving significant insights into the technical mechanisms of detection or investigation.

b) Traditional Approaches to Fraud Detection and Their Limitations

Traditional fraud detection in retail environments relies on rule-based logic, statistical thresholds, and laborious processes of manual review. These methods typically use predefined heuristics, including return frequency limits or transaction value thresholds, to flag suspicious activity. While

such methods are relatively easy to implement, their effectiveness rapidly degrades as fraud patterns evolve.

Fawcett and Provost showed in their research that "a fraud behaviour which is complex and adaptive is poorly modelled by a static rule-based systems when fraudsters deliberately adapt their actions to get around the system". Recent studies still present this observation: adversarial behaviour and concept drift seriously lower fixed-rule system performance over time.

Further, traditional techniques typically return a high volume of false positives, which in turn translate into unnecessary friction on customers and added operational cost. In terms of scaling, an omnichannel retail environment with increasing volume and velocity of return transactions cannot practically continue with manual audits or post-transaction reviews. This has been a strong motivation for adaptive data-driven methods that are able to learn from historical patterns and change over time.

c) Fraud investigation and digital forensics

Digital forensics provides an ordered process for the investigation of incidents where digital evidence is to be considered. Foundational work by Casey identified evidence identification, preservation, and analysis followed by presentation as a core investigative process, focusing on the protection of evidentiary integrity throughout the whole lifecycle of the investigation itself. In turn, Beebe and Clark proposed similar frameworks that provided a structured approach to forensic investigation, mirroring the technical analysis to the objectives of the investigation to enhance consistency and accountability.

Traditionally, within retail and commercial environments, the application of digital forensics has been restricted to cyber incidents, insider threats, and financial misconduct. Standards such as NIST SP 800-86 outline best practices for integrating forensic techniques into incident response processes that further reinforce the need for forensic readiness in organizational systems. These are increasingly applicable principles within the domain of retail return fraud, where every fraud leaves behind a footprint at almost all layers of multi-interconnected systems: from point-of-sale platforms and e-commerce portals to payment gateways and logistics systems.

Despite its relevance, the application of digital forensics to retail return fraud is still scant in current literature. Most of the studies treat forensic analysis as a post-incident activity, not as an operational capability embedded in the system. Moreover, issues like volume, heterogeneity, and potential privacy constraints often make large-scale manual forensic analysis challenging or impossible. This represents an important gap that may be filled by automated and intelligent approaches while supporting forensic investigations without compromising evidentiary standards.

d) Fraud Detection with Machine Learning Techniques

The domain of applications of machine learning for fraud detection encompasses banking, insurance, and e-commerce. In this line, cost-sensitive learning methods, such as those proposed by Bahnsen et al., explicitly consider the imbalance between fraudulent and legitimate transactions in model training, considering misclassification costs [11]. Such techniques are particularly relevant in retail returns, where cases are typically rare but financially significant.

Unsupervised and semi-supervised learning methods also have gained momentum given the scarcity of labelled fraud data. Chandola et al. presented an extensive survey of anomaly detection techniques and outlined how they were suited for detecting fraud patterns that were not previously seen [13]. Streaming and real-time detection frameworks such as those by Carcillo et al. further present how ML-based systems can scale in high-volume transactional settings [12].

More recent studies explore explainability and transparency of fraud detection models. Explainable AI techniques seek to render model decisions interpretable to investigators, auditors, and compliance teams to address concerns over trust and regulatory accountability [14], [15]. However, many ML-based fraud detection studies primarily concentrate on predictive performance and rarely consider how model outputs can be utilized in forensic investigations, or provided as admissible evidence.

e) Overall Summary and Research Gaps Identified

Literature reviewed shows that although retail return fraud is acknowledged as a major problem, research now remains fragmented across disciplines. Research on reverse logistics and retail fraud mainly focuses on operational and financial impacts, while digital forensics research focuses on investigative rigor without scalability. Similarly, machine learning research demonstrates strong detection capability but often does not align with forensic principles such as traceability, evidence preservation, and explainability.

The integration of digital forensics and machine learning provides a comprehensive and robust approach to detecting retail return fraud by combining evidentiary rigor with intelligent, scalable analysis. Digital forensics establishes a structured framework for identifying, preserving, and analysing digital evidence generated across retail systems such as point-of-sale terminals, e-commerce platforms, payment gateways, and logistics networks. By ensuring data integrity, maintaining chain of custody, and aligning investigations with established standards such as NIST SP 800-86, digital forensics enables reliable reconstruction of return events and supports legal and regulatory accountability.

Machine learning strengthens this forensic base by adding the capability to analyse high volume, diverse retail data. Methods such as cost-sensitive learning help to balance the class distribution, making it possible to learn from imbalanced data. Unsupervised, semi-supervised learning algorithms, such as anomaly detection, help to discover new, unknown patterns of fraudulent returns. Real-time machine learning algorithms help to detect returns by monitoring transactions and behavioural indicators. By working in tandem with other AI technologies, such as explainable AI, machine learning results can be traced to indicate why the transaction raised an alert.

This synergy makes possible the detection of accurate cases of retail return fraud, misuse of inside accounts, and financial

impropriety, all while maintaining acceptable standards of evidence, through the combined effort of machine learning, which serves as the triage and pattern recognition layer, to prioritize cases of suspected misuse, and digital forensic validation, examination, and documentation for audit and litigation purposes.

Noticeably, few works adopt an integrated viewpoint that merges digital forensics with machine learning for retail return fraud alone. Matters of forensic readiness, ethics of customer data utilization, and legal defensibility of automated fraud detection systems are barely addressed. It is essential to address these gaps through a holistic review that can bridge the technical, operational, and investigative dimensions, which is what this paper purport to achieve.

3. Review Methodology

This research employs a Structured Narrative Review methodology, in a bid to comprehensively evaluate the body of existing research that relates to the aspect of retail return fraud, digital forensics, and the use of machine learning approaches for the purpose of detecting fraud. This methodology will not involve a meta-analysis as the basis of this research.

a) Literature Search Strategy

The relevant literatures are collected through specific searches from the reputable digital libraries such as IEEE Xplore Digital Library, Science Direct, SpringerLink Digital Library, ACM Digital Library, and Google Scholar Digital Library. The search terms used are Retail Return Fraud, Reverse Logistics Fraud, Digital Forensics Studies, Machine Learning Fraud Detection Techniques, and Anomaly Detection Methods for Retailing. Only peer-reviewed journals and relevant industry reports from 2005 and 2024 are referenced.

b) Study Selection and Inclusion Criteria

Studies had to deal with, at least, one of the following: (i) retail returns or return fraud, (ii) use of digital forensics in fraud or business investigations, or (iii) machine learning solutions for fraud detection in retail or transactional settings. Studies solely on financial fraud domains that are irrelevant, or studies whose methods are not clearly stated, were excluded. High-citation studies with robust methods have been favoured.

c) Review and Classification Approach

The literature that met the search criteria for selection and analysis has been evaluated and categorized systematically on the basis of its prime theme: retail-return fraud characteristics, conventional and cyber forensic techniques for investigation, and ML-driven techniques for its identification. A comparative scrutiny of these studies also helped in highlighting similarities and gaps within these reviewed literature pieces.

d) Synthesis and Gap Identification

Findings from the classified literature were employed using qualitative synthesis to emphasize the challenges, best practices, and open issues within the research. Notable focus has been given to the recognition of the knowledge gaps within the fusion of the forensic processes with the machine learning paradigms, such as scalability, explainability, or the reliability of evidence. Knowledge gaps are identified to provide the foundation for discussion within the following sections.

4. Results and Discussion

It is evident from the reviewed literature that retail return-related fraud has become a common phenomenon with the rise of e-commerce and relaxed return policies. It has also been identified in existing research that it is quite challenging to detect return-related fraudulent activities by using the rule-based system, as it is a static solution and does not incorporate the dynamic nature of customer behaviour.

Machine learning algorithms show great potential in detecting complicated and clandestine patterns of fraud based on behavioural as well as historical information. Nevertheless, most research work on fraud detection emphasizes the issue of accuracy rather than explainability and accountability. Digital forensics offer organized tools for evidence collection and investigation but are generally employed post-fraud and are not scalable in the case of retail fraud.

In general, it seems that the integration of machine learning for early detection and the principles of digital forensics for investigation and validation can offer a more effective approach for the management of retail return fraud.

5. Conclusion

The review of study that was carried out in this study entails a comprehensive examination of the literature that exists in the field of retail return fraud, digital forensics, and machine learning. From the study, it can be seen that the conventional means of fraud discovery are not adequate in the contemporary retail setup. Machine learning in fraud discovery, on the other hand, provides the benefit of scalability and adaptability, in addition to the attribute of digital forensics that

Taken independently, these methods are handling partly the issue. There exists an enormous necessity to integrate machine learning and forensic techniques to produce an effective, scalable, as well as defensible method to discover cases of retail return-related fraud.

6. Future Scope

Future research endeavours could be directed towards creating a converged approach for fraud detection that integrates machine learning and digital forensic preparedness. There could be a emphasis on developing models for transparency, privacy-respecting analytics, and real-time detection capabilities for omnichannel retail environments. Simultaneously, a convergence of common benchmarks for research could go a long way in making research more consistent and comparable.

References

[1] D. S. Rogers, B. Melamed, and R. S. Lembke, "Modeling and analysis of reverse logistics," *Journal of Business Logistics*, vol. 33, no. 2, pp. 107–117, 2012.
[2] National Retail Federation and Appriss Retail, *2023 National Retail Security Survey*, Washington, DC, USA, 2023.
[3] Deloitte, *Managing the Returns Challenge in Retail*, Deloitte Insights, New York, NY, USA, 2021.
[4] J. Xu and Z. Chen, "A comprehensive study of retail fraud and loss prevention," *Decision Support Systems*, vol. 57, pp. 294–307, 2014.
[5] A. Beck and M. Hopkins, "Retail shrinkage and fraud: Trends and prevention strategies," *Security Journal*, vol. 32, no. 1, pp. 1–19, 2019.
[6] A. Dal Pozzolo, G. Bontempi, O. Snoeck, and D. Bontempi, "Adversarial drift detection in fraud detection," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 29, no. 10, pp. 4753–4767, Oct. 2018.
[7] T. Fawcett and F. Provost, "Combining data mining and machine learning for effective fraud detection," in *Proc. 2nd Int. Conf. Knowledge Discovery and Data Mining (KDD)*, Portland, OR, USA, 1997, pp. 9–15.
[8] E. Casey, *Digital Evidence and Computer Crime*, 3rd ed. Burlington, MA, USA: Academic Press, 2011.
[9] N. L. Beebe and J. G. Clark, "A hierarchical, objectives-based framework for the digital investigations process," *Digital Investigation*, vol. 2, no. 2, pp. 147–167, 2005.
[10] National Institute of Standards and Technology (NIST), *Guide to Integrating Forensic Techniques into Incident Response*, NIST Special Publication 800-86, Gaithersburg, MD, USA, 2006.
[11] A. C. Bahnsen, D. Aouada, and B. Ottersten, "Cost-sensitive decision trees for fraud detection," *Expert Systems with Applications*, vol. 39, no. 10, pp. 9546–9553, 2012.
[12] F. Carcillo, Y.-A. Bontempi, O. Snoeck, and G. Bontempi, "Scarff: A scalable framework for streaming credit card fraud detection," *Information Fusion*, vol. 41, pp. 182–194, 2018.
[13] V. Chandola, A. Banerjee, and V. Kumar, "Anomaly detection: A survey," *ACM Computing Surveys*, vol. 41, no. 3, Article 15, 2009.
[14] M. T. Ribeiro, S. Singh, and C. Guestrin, "Why should I trust you? Explaining the predictions of any classifier," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining*, San Francisco, CA, USA, 2016, pp. 1135–1144.
[15] A. B. Arrieta *et al.*, "Explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges," *Information Fusion*, vol. 58, pp. 82–115, 2020.
[16] E. W. T. Ngai, Y. Hu, Y. H. Wong, Y. Chen, and X. Sun, "The application of data mining techniques in financial fraud detection: A classification framework and an academic review of literature," *Decision Support Systems*, vol. 50, no. 3, pp. 559–569, 2011.
[17] S. Raghavan, V. Madani, and R. Jones, "Fraud analytics using descriptive, predictive, and social network techniques," *IBM Journal of Research and Development*, vol. 58, no. 5, pp. 1–13, 2014.