Journal of Information Systems Engineering and Management
2025, 10(55s)
e-ISSN: 2468-4376
https://jisem-journal.com/

Research Article

Secure Document Automation Using Blockchain Anchors and AI-Validated Semantic Hashing for Invoice Integrity

Ranadheer Reddy Charabuddi [1]
[1] *Aventis Inc, USA. Email: ranadheer30@gmail.com*

ARTICLE INFO

Received: 14 July 2025

Revised: 18 July 2025

Accepted: 21 July 2025

ABSTRACT

During the age of digital transformation, invoice automation is central to operation efficiency but is vulnerable to fraud, semantic manipulation, and structural tampering. Legacy document processing systems based on Optical Character Recognition (OCR) and rule-based field extraction tend not to incorporate semantic comprehension as well as cryptographic security needed to provide document integrity. In response to these key shortcomings, this research introduces a new framework referred to as Double-Layered Integrity Validation using Blockchain-Coupled Structural and Semantic Hashing (DLIV-BCSH). The system integrates cryptographic and artificial intelligence-based methods to implement a tamper-evident verification process. To that end, it employs the BLAKE3 algorithm for the fast, collision-resistant structural hashing and Sentence-BERT with SimHash for semantic similarity capture. This two-stage hashing is then rooted in a blockchain, developing immutable, timestamped audit trails without on-chain exposure of sensitive invoice content. The scheme is tested with the Customer Invoices Dataset from Kaggle, which contains realistic but synthetic structured invoices. A preprocessing pipeline consisting of field extraction, text normalisation, data imputation, and tokenisation gets the data ready for secure hashing and embedding. Implemented in Python, the DLIV-BCSH system shows excellent performance statistics with 98.5% accuracy of tamper detection, 96.7% semantic sensitivity, and an average latency of blockchain anchoring at 0.85 seconds. It precisely identifies byte-level as well as meaning-level tampering and performs better than current models like ARCHANGEL and VBlock. This study highlights the strong synergy between Blockchain Technology and Artificial Intelligence, providing a secure, low-latency, and privacy-preserving way for document integrity validation.

Keywords: Semantic Hashing, Invoice Integrity, Blockchain Anchoring, BLAKE3, Sentence-BERT.

INTRODUCTION

In the evolving environment of digital transformation, document automation has become a very important element of business processes, especially in the sectors of manufacturing such as in terms of finance, supply chain management, healthcare, and government[1]. Invoice processing is highly prone to fraud, manipulation, and functional inefficiencies since it is repetitive, sensitive, and business-critical among all the listed ones. The e-invoicing systems are largely structured through optical character recognition (OCR), rule-based extraction of information and predetermined checks of validations[2]. Unlike the original version of such tool, it is not useful in data integrity, fine tampering or traceability which is very essential in financial transparency and regulatory compliance[3]. The rising rates of electronic invoice fraud like manipulation of inappropriate contents, value manipulation, and semantic misrepresentation, as well as the unauthorized use of contents observed confirm to us the importance of needing secure, verifiable, and intelligent document handling mechanisms[4]. Most modern-day security systems are based on file level or structure hashing functions such MD5, SHA-256, or SHA-3 and are designed to detect differences in the raw byte representation, but do not have the capability to contextually or semantically read what the invoice is saying. Due to this, its vulnerability is that attackers can exploit by injecting certain linguistic or formatting changes that alter the meaning of an invoice without hash mismatches[5].

Journal of Information Systems Engineering and Management
2025, 10(55s)
e-ISSN: 2468-4376
https://jisem-journal.com/

Research Article

To address the existing limitations, this study proposes a novel architecture with two layers, which is DLIV-BCSH (Double-Layered Integrity Validation using Blockchain-Coupled Structural and Semantic Hashing). It unites cryptographic and semantic verification to defend the integrity of invoices. BLAKE3 is used to store the structure of the documents at the byte level, Sentence-BERT and SimHash generate semantic hashes of documents about invoices. These hashes are weaved to a blockchain, producing recordings that cannot be modified, and have a timestamp. It is semantically and structurally tampered in real-time, and hence the resistance is high even when the phrasing or structure of the document is altered. The method enhances sets of security, scalability, and reliability in the automated accounts payable, as well as in general digitalized document landscapes. The key contribution of this study follows as:

- Combines BLAKE3 structural hashing with Sentence-BERT and SimHash semantic hashing to detect both low-level and meaning-level tampering in invoices.
- Anchors dual-layered hash records to blockchain for permanent, tamper-evident audit trails and integrity verification.
- Achieves high performance with 98.5% tamper detection accuracy, 96.7% semantic sensitivity, and 0.85-second blockchain anchoring latency.
- Implements a scalable, modular pipeline adaptable for real-world invoice automation and extensible to other document types.

The subsequent sections of this study are arranged as follows. Section II reviews prior research, Section III discusses the problem statement for the method, while Section IV discusses the effectiveness of the methodology and summary of findings. Section V concludes and summarizes future work.

LITERATURE REVIEW

“Azzam et al.[6] aims to increase the security and integrity of document processing in e-government environments through the combination of Optical Character Recognition (OCR) and blockchain technology. The primary goal of the system is to ensure that official documents, for example, incoming invoices, cannot be tampered with once they have been accepted by government departments. The system leverages OCR to read textual data from paper or electronic documents and produce a cryptographic hash, which is stored safely on a blockchain. SECHash has several key benefits: less manual labor, lower fraud, guaranteed document longevity, and improved auditability.

Radha, Kuehlkamp, and Nabrzyski [7] solves the inefficiencies in traditional document attestation procedures—delays, fraud, and non-transparency—through the design of a blockchain and self-sovereign identity (SSI)-driven framework. The approach combines decentralized ledger technology for safe, tamper-proof storage and SSI for user-managed identity verification. Major benefits are prevention from fraud, privacy boost, real-time monitoring, and lower reliance on physical presence. Outcomes indicate better security and process efficiency, especially in remote situations such as amidst the COVID-19 pandemic.”

Oluwaferanmi [8] proposes an SSI-based online document verification solution to mitigate the increasing threat of digital document forgery and the intricacy of existing verification processes. The solution increases security, privacy, and interoperability, especially in online lending scenarios. Findings identify decreased verification time, increased user control of data, and higher trust levels, showing the efficiency of the SSI-based method in simplifying digital credential verification. Meanwhile, Liu et al.,[9] targets the improvement of electronic voting and audit log system security and reliability through the application of blockchain technologies. No particular dataset was utilized but rather prototypes for the simulation of actual environments. The DBE-voting system proposed employs a double blockchain (private and public) with linkable ring signatures for privacy, authentication, and data integrity. Experiments indicate that the DBE-voting system performs 29 transactions per second and the audit log system saves 50% storage and withstands one-third compromised nodes. Challenges are system adoption and integration complexity.

Journal of Information Systems Engineering and Management
2025, 10(55s)
e-ISSN: 2468-4376
https://jisem-journal.com/

Research Article

Romero and Hernandez [10]offers a Blockchain-based solution for multipolicy management for insurance firms, presenting a standardized policy model that ensures efficient workings and improved interoperability among organizations. The model provides consistent policy management, ensuring scalability and flexibility to accommodate changing market needs. The solution utilizes Merkle trees to ensure secure data management, with every policy tied to an isolated Merkle tree, allowing updates and additions without compromising existing policies. The architecture, which runs on a private Ethereum network via Hyperledger Besu and Tessera, provides secure and transparent transactions, effective dispute resolution, and fraud prevention systems. The validation stage proved the model effective in minimizing data redundancy and maintaining policy data consistency and integrity. Furthermore, technical management of the system has been eased, redundancies in operations have been avoided, and privacy is increased.

Dragomirescu et al.,[11] create a scalable and sustainable invoice processing model using the combination of DevOps practices and machine learning (ML). The strategy applies dynamic model testing, training, deployment, and monitoring in a Proof of Concept (PoC) to automate the financial processes and minimize manual handling. There is no explicit dataset referred to, but the analysis comes from survey-based analysis. Major benefits comprise improved operational insight, decreased errors, and maximized resource utilization. Yet, issues like integration difficulty and resistance to new systems persist as limiting factors.

Pandey et al.,[12]develop an Intelligent Document Management System (IDMS) capable of accurately extracting and processing data from medical bills, Aadhar cards, and PAN cards. The study employs EasyOCR and a hybrid method combining Natural Language Processing (Regular Expressions) and Computer Vision (OCR) for data extraction. Although the specific dataset is not named, it involves real-world document types related to healthcare and identity. The hybrid NLPCV approach showed superior accuracy—97% for hospital invoices, 71% for Aadhar cards, and 78% for PAN cards. However, limitations include variable accuracy across document types and challenges in handling poor-quality or unstructured inputs.

Bisetty et al.,[13]implement invoice verification automation using ERP systems to promote financial accuracy, efficiency, and transparency. The study utilizes a comparative study of ERP systems and suggests an architectural framework that combines automated workflows, data verification, and machine learning. Although there is no specific dataset, real-world case studies underpin the research. The benefits are less processing time, fewer errors, and savings. Yet, setbacks include complexity of initial implementation and integration issues with existing systems.

DUAL HASH BLOCKCHAIN FRAMEWORK FOR INVOICE INTEGRITY

The proposed methodology combines structured preprocessing, semantic and cryptographic hashing, and blockchain anchoring to provide secure invoice integrity. BLAKE3 is employed for structural hash and Sentence-BERT with SimHash for semantic checks, creating a two-layered validation mechanism. Blockchain anchoring provides tamper-evident and auditable records for extended document authenticity. The visual representation is shown in Figure 1.

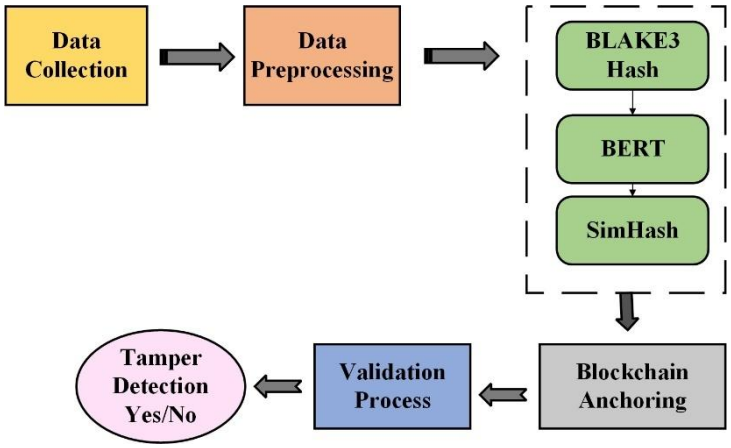

Figure 1: DLIV-BCSH Workflow for Secure Invoice Validation Using AI and Blockchain

Journal of Information Systems Engineering and Management
2025, 10(55s)
e-ISSN: 2468-4376
https://jisem-journal.com/

Research Article

3.1. Data Collection

The data used for the study is the Customer Invoices Dataset [14], which is available on Kaggle The data includes realistic synthetic invoices with fields such as customer names, invoice numbers, dates, and totals. It facilitates document automation, semantic embedding, blockchain anchoring, and hashing. It also allows testing of tampered forms to ensure the strength of a dual-layered integrity validation framework for secure invoice processing.

3.2. Data Preprocessing

Data preprocessing prepares raw invoice data for safe and correct processing by extracting the principal fields, cleaning and normalizing text, processing missing values, and formatting documents.

3.2.1. "Field Extraction and Structuring

The initial preprocessing operation is to extract important fields like invoice number, customer name, date, item description, total amount, and payment terms from every invoice. All these fields are then formatted into a standard format like JSON or CSV, allowing downstream processing to be done consistently and efficiently both for hashing and semantic processing.

3.2.2. Text Cleaning and Normalization

Text normalization includes the stripping of punctuation, special characters, stopwords, and redundant whitespace along with standardization of formats such as dates and currencies. This enables uniform invoice text for semantic embedding accurately and trusted cryptographic hashing during integrity checks. The cleaned text T_c is derived from the raw text T_r as shown in (1)

$$T_c = Normalize(Clean(T_r)) \quad (1)$$

Where, Tr represents raw invoice text, Clean is the function to remove noise and Normalize is the function to standardize format.

3.2.3. Missing Value Handling

To maintain consistency and accuracy in document processing, missing value handling is a core preprocessing operation. Numerical fields like invoice values or tax percentages containing missing entries in this research are filled using the mean imputation method. The missing value $x_{missing}$ is estimated as shown in (2)

$$x_{imputed} = \frac{1}{n}\sum_{i=1}^{n} x_i \quad (2)$$

Where, x_i represents the known data values, n represents count of known values and $x_{imputed}$ represents the estimated value."

3.2.4. Tokenization and Embedding Preparation

In order to pre-process invoice text for semantic embedding, the cleaned text is tokenized into semantically significant units of sentences and encoded in UTF-8. This ensures compatibility with natural language processing models such as Sentence-BERT.

3.2.5. Document-Level Formatting for Hashing

For document-level formatting, each invoice is reformatted into a normalized text structure in a uniform field ordering and layout. The formatted output is subsequently employed in structural hashing with BLAKE3 and semantic hashing with SimHash, allowing for correct and tamper-evident document integrity verification in downstream processing.

3.3. Structural Hash Generation (Cryptographic Layer)

To provide byte-level security and identify unauthorized changes in invoice content, this research uses the BLAKE3 cryptographic hash function, which is a high-speed cryptographic hash function, against normalized invoice data[15]. Preprocess and arrange each invoice document of the normalized Customer Invoices Dataset first in a uniform field

Journal of Information Systems Engineering and Management
2025, 10(55s)
e-ISSN: 2468-4376
https://jisem-journal.com/

Research Article

order, such as invoice number, date, customer name, item description, total amount, and payment terms. The organized content D is then sent to the BLAKE3 function to produce a structural hash that is unique as shown in (3)

$$H_s = BLAKE3(D) \quad (3)$$

Where, D represents the pre-processed and normalized invoice data, H_s is the generated structural hash value. BLAKE3 facilitates detection of even slight content alteration through creation of unique hash values, making it even more secure against forgery. It surpasses SHA-2 and SHA-3 in terms of speed and security and is best suited for invoice-level automation. When paired with semantic hashes, it provides a two-stage document anchoring process on the blockchain.

3.4. Semantic Hash Generation (AI Layer)

Although structural hashing can efficiently detect byte level changes, it proves to be inadequate in detecting semantic content changes in text data. For addressing this weakness, the methodology suggested includes a semantic hashing layer fueled by artificial intelligence. Here, the preprocessed invoice text is passed through a pretrained transformer-based language model, namely Sentence BERT. This model represents the text in dense vector representations that preserve the semantic content of important invoice details, including payment terms, product descriptions, and customer instructions[16]. The high dimensional embeddings thus obtained are then passed to $SimHash$, which compresses the semantic information into a fixed-length binary hash, which is known as the semantic hash. This two-layered methodology ensures that tampering on the semantic level, for instance, invoice meaning manipulation, is successfully identified, enhancing document integrity verification robustness.

3.5. Blockchain Anchoring

To provide long term integrity and non-repudiation in invoice verification, both structural hash H_s and semantic hash H_{sem} are rooted onto a blockchain. Structural hash is computed with BLAKE3, while semantic hash is computed using Sentence-BERT embeddings and $SimHash$. They collectively create a two-layer integrity tuple with embedded metadata like document ID, issuer ID, and timestamp[17]. This metadata is logged in a blockchain transaction, producing a tamper-evident, time-stamped ledger record. Public blockchains such as Ethereum provide decentralized trust, whereas permissioned blockchains such as Hyperledger Fabric ensure controlled privacy and access. Importantly, no invoice content is kept on-chain—merely hashes and metadata—securing data privacy. This anchoring provides future confirmation by checking re-generated hashes against the originals. Any difference indicates possible tampering. This technique forms an unalterable, audit-tracked record, increasing trust, security, and regulatory compliance in automated accounts payable processing systems without revealing sensitive information.

3.6. Integrity Verification Process

The integrity validation process ensures the accessed invoice is not altered structurally or semantically. The invoice, in validation, is subject to the same preprocessing as initial registration. BLAKE3 produces a fresh structural hash, while Sentence-BERT and SimHash result in a fresh semantic hash. These are compared to the original blockchain-anchored hashes. If the hashes match, document authenticity is verified. A structural mismatch predicts low-level modifications, whereas a semantic mismatch indicates changed content meaning. In both, the mismatches predict full tampering. This two-layer authentication successfully identifies illegal changes, providing high trust, security, and integrity in automated invoice processing of secure document automation systems.

Algorithm: Invoice Integrity Verification

Input: Raw Invoice

Output: Integrity Status (Valid / Tampered)

1. Preprocess: Extract → Clean → Normalize → Impute → Format

2. Hash Generation: $H_s = BLAKE3(D)$

Journal of Information Systems Engineering and Management
2025, 10(55s)
e-ISSN: 2468-4376
https://jisem-journal.com/

Research Article

3. Blockchain Anchoring: Store H_s , H_{sem}, and metadata on blockchain

4. Verification: Recompute H_s', H_sem'

Compare with anchored hashes

Match: Valid

Mismatch: Tampered (Structural / Semantic / Both)

RESULT AND DISCUSSION

The section analyses the performance of the suggested framework in identifying tampering of invoices. Through visual inspection and accuracy measures, the results are compared to current approaches. The discussion points out the robustness, reliability, and real-world usability of the system in providing safe, tamper-evident document automation through double-layered integrity checking and blockchain anchoring. The simulation parameter is shown in Table 1.

Table 1: Simulation Parameter

Parameter	Value
Dataset	Customer Invoices Dataset (Kaggle)
Hashing Algorithm (Structural)	BLAKE3
Hashing Algorithm (Semantic)	Sentence-BERT + SimHash
Block Size (Tested)	512 KB
Software	Python

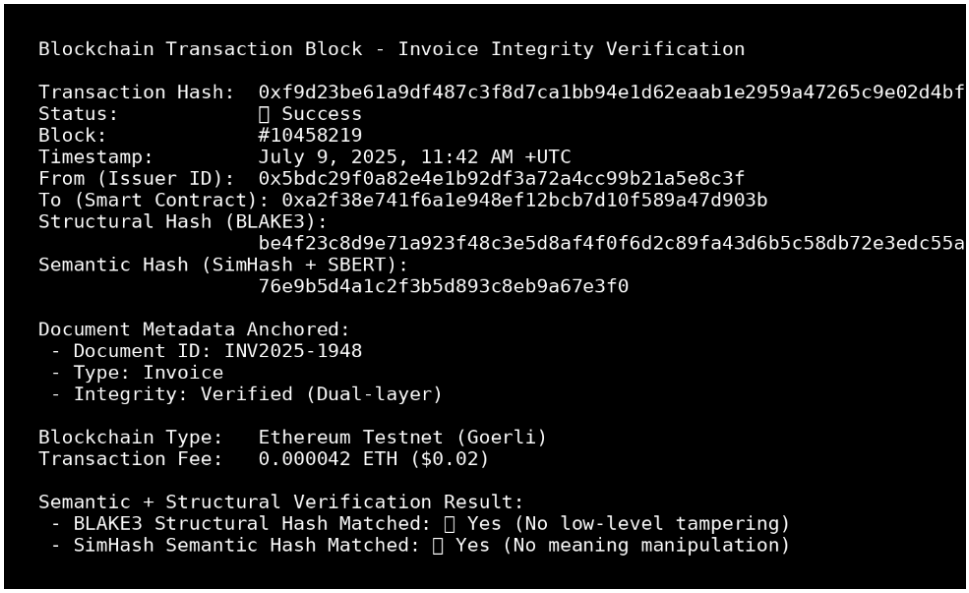

Figure 2: Blockchain Invoice Verification Record on Ethereum Testnet

Figure 2 shows This chart displays a blockchain-based integrity verification record of an invoice, safely anchored on the Ethereum Goerli testnet. It contains transaction metadata like the hash, timestamp, block number, issuer, and smart contract addresses. Structural and semantic hashes (BLAKE3 and SimHash+SBERT) validate the invoice's authenticity and semantic preservation. The verification status indicates success, with no tampering detected. Document metadata indicates invoice ID and two-layer integrity verification. The low transaction fee underscores cost-effectiveness of blockchain-based document authentication systems.

Journal of Information Systems Engineering and Management
2025, 10(55s)
e-ISSN: 2468-4376
https://jisem-journal.com/

Research Article

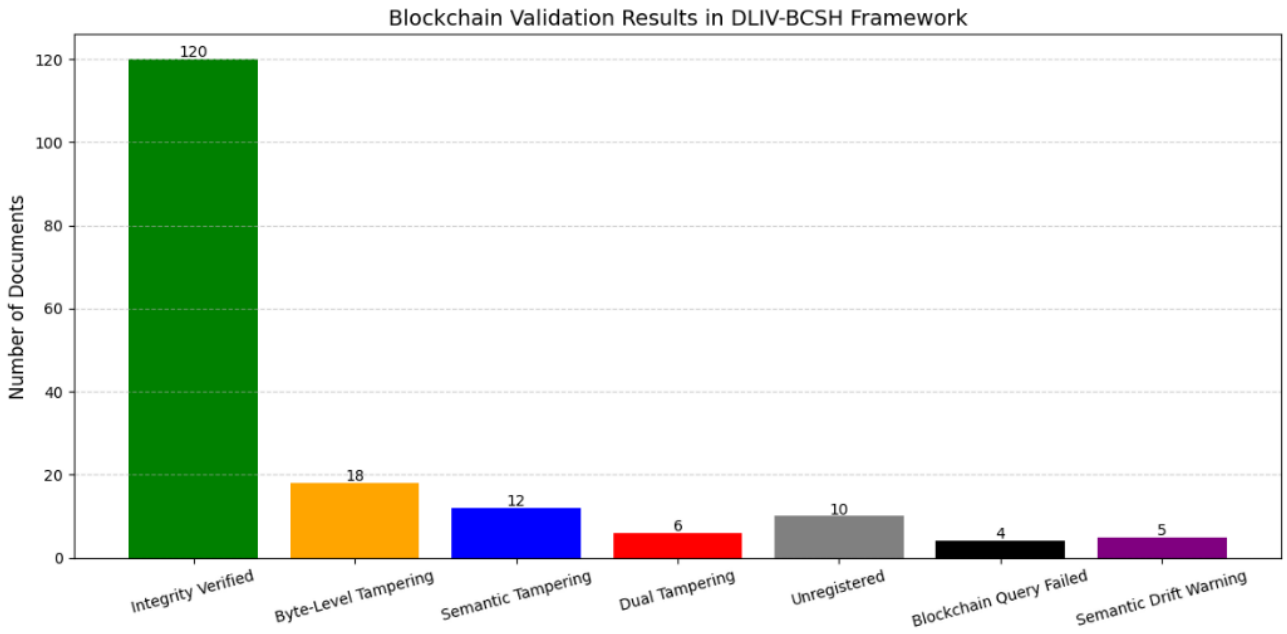

Figure 3: Blockchain validation results

Figure 3 shows the document validation results with the DLIV-BCSH framework. Of all the cases, 120 documents were confirmed as intact. There were 18 cases of byte-level tampering, 12 cases of semantic tampering, and 6 cases of dual tampering. 10 documents were also unregistered, 4 incurred blockchain query failures, and 5 raised semantic drift warnings. The framework is able to differentiate various integrity problems with invoice documents effectively.

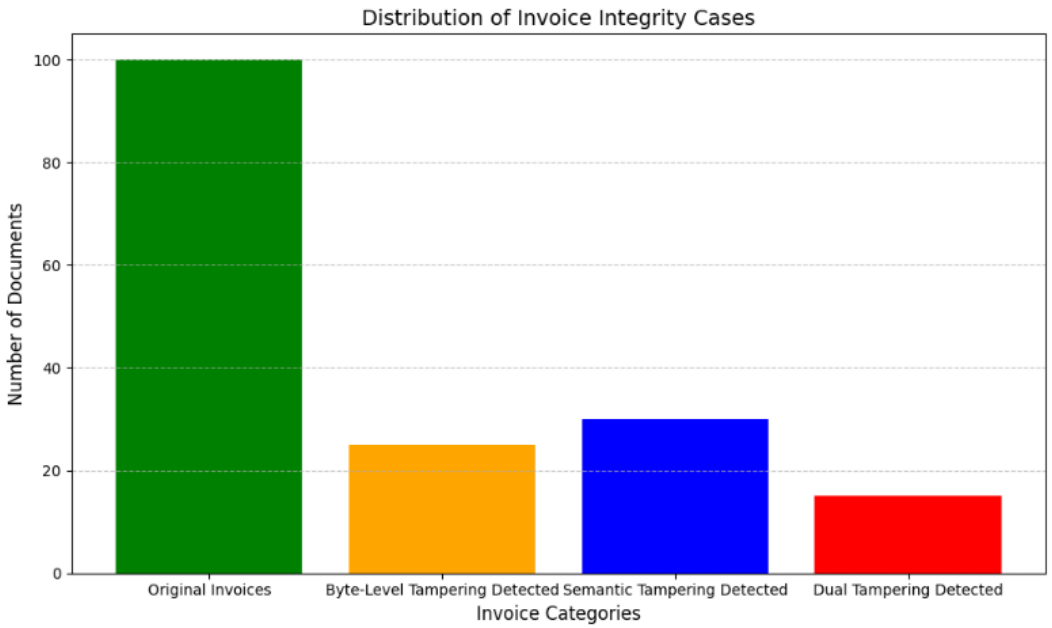

Figure 4: Distribution of Invoice Integrity Cases Across Detection Categories

Figure 4 displays the distribution of cases of invoice integrity examined through the DLIV-BCSH paradigm. It indicates that among all documents processed, 100 invoices were original and intact. In comparison, 25 invoices underwent byte-level tampering, 30 semantically tampered, and 15 exhibited dual tampering in structure and meaning. The display emphasizes how structural and semantic validation on top of each other strengthens subtle manipulation detection.

Journal of Information Systems Engineering and Management
2025, 10(55s)
e-ISSN: 2468-4376
https://jisem-journal.com/

Research Article

4.1 Performance Metrics

Performance measures test the capabilities of a system to identify the tampering with documents and manipulations of semantics.

4.1.1 Tamper Detection Accuracy (TDA):

Measures the accuracy of the detection of tampered and untampered invoices by the system. The mathematical expression is shown in (4)

$$TDA = \frac{TP+TN}{TP+TN+FP+FN} \times 100 \quad (4)$$

Where, TP is the true positive, TN is the true negative, FN is the false negative, and FP is the false positive.

4.1.2 False Acceptance Rate (FAR):

It evaluates the rate of tampered invoices accepted in error as untampered by the system. It indicates the system's failure to identify some modifications. The mathematical expression is shown in (5)

$$FAR = \frac{FN}{FN+TP} \times 100 \quad (5)$$

4.1.3 Blockchain Anchoring Latency (BAL)

It is the Average time taken anchoring the dual-layer integrity tuple (structural and semantic hash) of an invoice to the blockchain. The mathematical expression is shown in (6)

$$BAL = \frac{1}{n}\sum_{i=1}^{n}(t_{end,i} - t_{start,i}) \quad (6)$$

Where, $t_{start,i}$ is the time when anchoring starts for invoice i, $t_{end,i}$ represents the time when anchoring completes for invoice i, n represents the total number of invoices processed

4.1.4 Semantic Sensitivity Rate (SSR)

It measures how well the system can identify small but significant changes in document content. The mathematical expression is shown in (7)

$$SSR = \frac{Detected\ Sematic\ Tampering}{Total\ Semantic\ Changes} \times 100 \quad (7)$$

Where, SSR represents the semantic sensitivity rate.

Table 2: Performance Metrics

Metrics	Value
Tamper Detection Accuracy	98.5%
False Acceptance Rate	1.2%
Blockchain Anchoring Latency	0.85 seconds
Semantic Sensitivity Rate	96.7%

Figure 5 and Table 2 shows four most important evaluation criteria applied to measure an invoice integrity verification system's effectiveness. According to the chart, the best performers are Tamper Detection Accuracy (98.5%) and Semantic Sensitivity Rate (96.7%), proving the model's high capacity for detecting structural as well as semantic tampering. False Acceptance Rate remains low at 1.2%, which means that there are few undetected manipulations. The Blockchain Anchoring Latency, which appears as 0.85 seconds, indicates effective anchoring of integrity records to the blockchain.

Journal of Information Systems Engineering and Management
2025, 10(55s)
e-ISSN: 2468-4376
https://jisem-journal.com/

Research Article

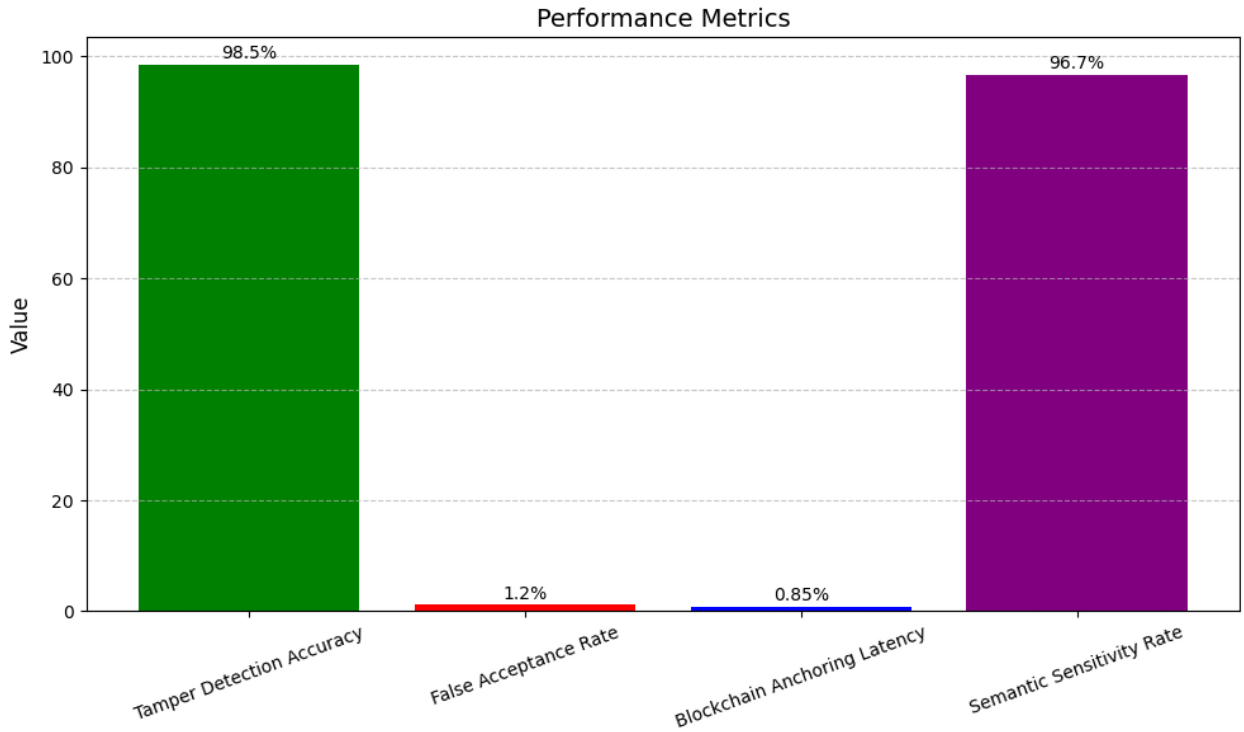

Figure 5: Performance metrics

4.2 Comparative Analysis

Table 3: Comparative with Existing methods

Methods	Tamper Detection Accuracy
ARCHANGEL [18]	98.1%
Blockchain-Based Video Forensics and Integrity Verification Framework [19]	92%
VBlock [20]	98%
Proposed DLIV-BCSH	98.5%

Table 3 shows the Tamper Detection Accuracy (TDA) comparison among four blockchain-based integrity verification techniques. ARCHANGEL has a TDA of 98.1%, and Blockchain-Based Video Forensics and Integrity Verification Framework is lower at 92.0%. VBlock has a TDA of 98.0%, showing high tamper resistance in Internet-of-Vehicle networks. The suggested approach, DLIV-BCSH (Dual-Layer Invoice Verification with BLAKE3 and Semantic Hashing), surpasses all with the best TDA of 98.5%.

4.3 Discussion

DLIV-BCSH framework shows high accuracy in identifying structural as well as semantic tampering in invoice documents. It records 98.5 % tampering detection rate and 96.7 % semantic sensitivity, which is higher than current models such as ARCHANGEL and VBlock. The low false acceptance rate of 1.2 percent and low blockchain anchoring latency of 0.85 second makes the system ideal in real time applications. The two layered hashing method takes document integrity detection a notch up by detecting both byte and meaning manipulations. It is easily scalable with its modular design and can integrate into various business systems as it is implemented in Python. This makes DLIV-BCSH a sound, expedient and privacy-preserving platform to secure automated and streamlined validation of invoices within a digital business enterprise setting.

Journal of Information Systems Engineering and Management
2025, 10(55s)
e-ISSN: 2468-4376
https://jisem-journal.com/

Research Article

CONCLUSION AND DISCUSSION

The work introduces DLIV-BCSH, an intently designed two-layered integrity verification system that integrates cryptographic and AI-powered semantic hashing with blockchain anchoring to enable invoice automation security. By leveraging BLAKE3 for structural hashing and Sentence-BERT with SimHash for semantic verification, the system is tamper-evident, privacy-conscious document authentication. The framework is underpinned by performance metrics such as 98.5% accuracy in tamper detection and 96.7% semantic sensitivity at low latency and high scalability. Its implementation in Python and on the Kaggle Customer Invoices Dataset guarantees its practical usability. Future efforts will be towards making the framework more versatile across multilingual and domain-based invoice templates. The use of fine-tuned transformer models and investigating lightweight blockchain substitutes such as Hyperledger Fabric can further enhance performance and privacy. Integration of explainable AI (XAI) elements also has the potential to enhance interpretability in semantic tampering detection, thus enabling greater transparency and credibility for regulatory and business adoption.

REFERENCE

[1] K. Rauniyar, X. Wu, S. Gupta, S. Modgil, and A. B. Lopes de Sousa Jabbour, "Risk management of supply chains in the digital transformation era: contribution and challenges of blockchain technology," *Ind. Manag. Data Syst.*, vol. 123, no. 1, pp. 253–277, 2023.

[2] B. A. Masa, "Development of an Artificial Intelligence-based Solution for Document Processing Automation Using Machine Learning and NLP Techniques," 2022.

[3] M. I. Hossain, T. Steigner, M. I. Hussain, and A. Akther, "Enhancing data integrity and traceability in industry cyber physical systems (ICPS) through Blockchain technology: A comprehensive approach," *ArXiv Prepr. ArXiv240504837*, 2024.

[4] A. Rohilla, "Strengthening Financial Resilience: A Holistic Approach to Combatting Fraud," *Indian J. Econ. Finance IJEF*, vol. 4, no. 1, pp. 20–31, 2024.

[5] D. S. Soataliyev and N. Khandelwal, "Leveraging Blockchain-Backed Cloud Solutions for Secure, Compliant, and Optimized Remote Workforce Management".

[6] F. Azzam *et al.*, "The use of blockchain technology and OCR in e-government for document management: Inbound invoice management as an example," *Appl. Sci.*, vol. 13, no. 14, p. 8463, 2023.

[7] S. K. Radha, A. Kuehlkamp, and J. Nabrzyski, "The Future of Document Verification: Leveraging Blockchain and Self-Sovereign Identity for Enhanced Security and Transparency," *ArXiv Prepr. ArXiv241201531*, 2024.

[8] A. Oluwaferanmi, "Leveraging Blockchain and Smart Contract Technologies to Revolutionize US Supply Chain Finance: Implications for Trade Transparency, Real-Time Settlements, and Working Capital Optimization," 2025.

[9] Z. Liu and others, "Improving the security and reliability of application systems with blockchain technology," 2023.

[10] A. Romero and R. Hernandez, "Blockchain-Driven Generalization of Policy Management for Multiproduct Insurance Companies," *Future Internet*, vol. 16, no. 10, p. 356, 2024.

[11] O.-A. Dragomirescu, P.-C. Crăciun, and A. R. Bologa, "Enhancing Invoice Processing Automation Through the Integration of DevOps Methodologies and Machine Learning," *Systems*, vol. 13, no. 2, p. 87, 2025.

[12] M. Pandey, M. Arora, S. Arora, C. Goyal, V. K. Gera, and H. Yadav, "AI-based Integrated Approach for the Development of Intelligent Document Management System (IDMS)," *Procedia Comput. Sci.*, vol. 230, pp. 725–736, 2023.

[13] S. Bisetty, A. Ayyagari, A. Joshi, O. Goel, L. Kumar, and A. Jain, "Automating Invoice Verification through ERP Solutions," *Int. J. Res. Mod. Eng. Emerg. Technol.*, vol. 12, no. 5, p. 131, 2024.

[14] "Customer Invoices Dataset." Accessed: Jul. 09, 2025. [Online]. Available: https://www.kaggle.com/datasets/pradumn203/payment-date-prediction-for-invoices-dataset

[15] Z. E. Ahmed *et al.*, "A Complementary Approach for Securing and Anti-Counterfeiting of Valuable Documents Based on Encryption of Computer-Generated Hologram," *Sensors*, vol. 25, no. 8, p. 2410, 2025.

[16] D. Xu, N. Ren, and C. Zhu, "Integrity authentication based on blockchain and perceptual hash for remote-sensing imagery," *Remote Sens.*, vol. 15, no. 19, p. 4860, 2023.

Journal of Information Systems Engineering and Management
2025, 10(55s)
e-ISSN: 2468-4376
https://jisem-journal.com/

Research Article

[17] C. Zakaret *et al.*, "Blockchain and secure element, a hybrid approach for secure energy smart meter gateways," *Sensors*, vol. 22, no. 24, p. 9664, 2022.
[18] J. Ceron, C. Tinipuclla, and P. Shiguihara, "A Survey of Blockchain for Video Integrity," *Eng. Proc.*, vol. 42, no. 1, p. 4, 2023.
[19] H. F. Atlam, N. Ekuri, M. A. Azad, and H. S. Lallie, "Blockchain Forensics: A Systematic Literature Review of Techniques, Applications, Challenges, and Future Directions," *Electronics*, vol. 13, no. 17, p. 3568, 2024.
[20] T. H. Austin and F. Di Troia, "A blockchain-based tamper-resistant logging framework," in *Silicon Valley Cybersecurity Conference*, Springer Nature Switzerland Cham, 2022, pp. 90–104.