Digital Content Piracy Detection Using Blockchain Technology

Pooja K H
Department of Information Science and Engineering,
RV Institute of Technology and Management, Bengaluru,India
poojakh2026@gmail.com

Suravi S Thatha
Department of Information Science and Engineering,
RV Institute of Technology and Management, Bengaluru,India
suravi810@gmail.com

Uday Hiremath
Department of Information Science and Engineering,
RV Institute of Technology and Management, Bengaluru,India
udayhiremath45@gmail.com

Sai Disha J
Department of Information Science and Engineering,
RV Institute of Technology and Management, Bengaluru,India
dishajayaram11@gmail.com

Vinoth Kumar M
Department of Information Science and Engineering,
RV Institute of Technology and Management, Bengaluru,India
vinojimail@gmail.com

*Abstract— **Digital piracy, characterized by unauthorized copying and distribution of digital images, audio, video, and software, remains a pressing socioeconomic concern affecting creators and digital platforms. This paper presents a decentralized piracy detection system that integrates perceptual hashing and learned fingerprints for robust detection, leverages IPFS for off-chain content storage, and employs blockchain-based smart contracts for automated piracy reporting and royalty assignment. The system is implemented as a decentralized application combining a React.js frontend, Node.js backend, MongoDB metadata management, and Ethereum-compatible smart contracts deployed on the Polygon. Key features include duplicate prevention through SHA-256 and perceptual hashes, transparent and tamper-proof provenance via blockchain anchoring, and automated enforcement mechanisms through smart contracts. Testing demonstrates effective detection of exact and near-duplicate pirated content across images, audio, and short videos, while balancing operational scalability and cost. The study advances prior work by providing a reproducible, scalable framework addressing gaps in traceability, detection robustness, and decentralized enforcement without employing restrictive DRM. Future work will focus on extending fingerprinting to streaming media, optimizing blockchain transaction costs, and strengthening adversarial resistance to enhance practical deployment.***

Keywords— Digital piracy detection, blockchain technology, perceptual hashing, decentralized storage, IPFS, smart contracts, content fingerprinting, provenance, royalty management.

I. INTRODUCTION

Digital piracy, characterized by unauthorized copying and distribution of digital images, audio, video, and software, remains a major socioeconomic concern for creators and digital platforms. Beginning in the late 1990s and accelerating with the rise of broadband and streaming networks, piracy has undermined creative industries and exposed critical weaknesses in centralized security and copyright systems. Key concepts relevant to this domain include content fingerprinting, blockchain-based provenance, and decentralized incident reporting, all foundational for modern anti-piracy efforts.

Previous research and practical countermeasures have included anti-camcorder detection in theaters, perceptual hashing for effective fingerprinting, and blockchain-based approaches for content registration, transparent audit trails, and automated royalty distribution via smart contracts. While DRM has historically been deployed to restrict content access and usage, it is limited by centralization, lack of transparency, and vulnerability to circumvention. Modern solutions increasingly utilize blockchain to anchor content ownership, enable automated event handling, and reinforce traceability without enforcing traditional DRM constraints on end-users.

Despite these advances, most current systems fall short in jointly maximizing detection robustness, decentralized enforcement, and operational cost-efficiency. Many focus on content registration without resilient fingerprinting or incur excessive costs for on-chain operations, while others lack practical evaluation of adversarial models or latency. This leaves a gap for reproducible, scalable frameworks that combine perceptual and learned fingerprint detection, off-chain storage, transparent blockchain anchoring, and automated smart contract enforcement in a single platform.

The objective of this paper is to address these gaps by designing and implementing a fully reproducible piracy detection system that (a) integrates perceptual and machine-learned fingerprints for robust detection, (b) leverages IPFS for off-chain content storage with tamper-proof blockchain anchoring, and (c) deploys smart contracts for automated reporting and royalty assignment. The research scope covers detection for images, audio, and short videos, explicitly excluding full-scale DRM or legal arbitration mechanisms to focus on efficient, decentralized technical enforcement.

II. LITERATURE REVIEW

Digital piracy detection has been approached through multiple technological paradigms, including digital rights management (DRM), watermarking, perceptual hashing, and blockchain-based content protection. Early systems predominantly relied on DRM frameworks to restrict unauthorized access and duplication of digital files. While effective in access control, such centralized solutions suffered from transparency issues, dependency on trusted authorities, and ease of circumvention once decryption keys were compromised [10], [11], [20]. Watermarking techniques were introduced to embed invisible ownership information into multimedia files, allowing source tracing in case of unauthorized distribution. However, these methods were often vulnerable to compression, resizing, and format conversion attacks, limiting their long-term robustness [1], [3].

To overcome centralization issues, researchers explored blockchain as a foundation for decentralized copyright protection. Blockchain's immutability and transparency enable verifiable ownership proof, traceable transfer of digital assets, and automated royalty mechanisms via smart contracts [4], [5], [6]. For example, Qureshi and Jiménez [4] reviewed blockchain-based multimedia protection systems and identified open challenges related to scalability and interoperability. Madushanka et al. [6] proposed a blockchain-powered DRM framework (Secure Rights) to enhance trust and transparency in rights management, while Devendra et al. [7] implemented a decentralized copyright detection system for intellectual property protection. Similarly, Luo [20] and Bin et al. [9] highlighted blockchain's potential in intellectual property applications but noted that many systems were limited to registration-level protection without real-time piracy detection capabilities.

Parallel advancements in perceptual hashing and content fingerprinting provided a complementary direction for detecting duplicate or modified multimedia content. Techniques such as pHash and dHash generate compact representations that remain stable across transformations like cropping, scaling, or compression [12], [13], [18]. Trinh and Ta [12] introduced a pHash-enhanced Merkle Tree for detecting copyright violations in NFT marketplaces, demonstrating improved resilience to adversarial image modifications. McKeown and Buchanan [18] analyzed Hamming distance distributions of perceptual hashing methods to quantify robustness, while Beete et al. [19] combined perceptual hashing with blockchain to trace image sources, achieving verifiable authenticity but at high computational cost.

Despite these efforts, most prior systems address either ownership verification or piracy detection in isolation. Many blockchain-based solutions focus on registering and timestamping digital assets but lack integrated detection of near-duplicate or altered content. Conversely, perceptual hashing frameworks detect duplication but do not enforce ownership or automate penalties in decentralized environments. Furthermore, existing implementations often incur high transaction fees, depend on centralized storage, or overlook adversarial testing for robustness.

These limitations indicate the need for a holistic framework that integrates perceptual and learned fingerprinting for multi-format piracy detection, decentralized storage for scalability, and blockchain-based smart contracts for transparent enforcement. The proposed system in this paper aims to address these gaps by combining robust content fingerprinting, IPFS-based decentralized storage, and blockchain automation to enable secure, transparent, and cost-efficient digital content protection.

III. MATERIALS AND METHODS

A. Materials

Blockchain Environment: Polygon (Mumbai testnet) and Ethereum-compatible smart contracts (Solidity, Hardhat).
Backend: Node.js with Express.js for API handling.
Storage: IPFS (via Pinata SDK) for decentralized file storage, MongoDB for metadata management.
Frontend Development: React.js framework with Material-UI and Tailwind CSS for responsive design. Blockchain integration is achieved through Ethers.js and Web3Modal.
Testing & Version Control: Postman for API testing and Git for version control.

B. System Architecture

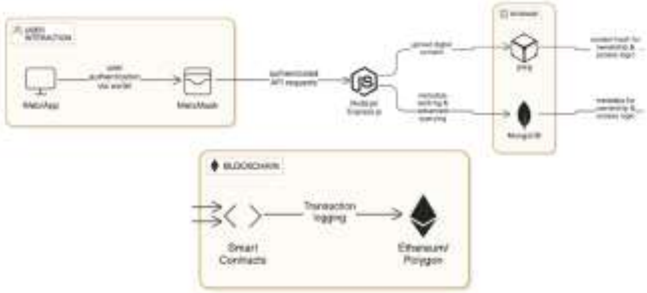


Figure 1: System architecture for blockchain-based piracy detection.

The overall workflow (Figure 1) consists of:
User Interaction: Users access the web application and authenticate using MetaMask.
Backend Integration: Authenticated API requests are handled via Node.js/Express.js.
Storage: Digital content is uploaded to IPFS, while metadata is cached and queried using MongoDB.
Blockchain Layer: Smart contracts log transactions, enforce ownership, and record provenance on Ethereum/Polygon.

C. Piracy Detection Process

Content fingerprints are generated using perceptual hashing (pHash) [12]. The hash and CID are stored on blockchain via a smart contract. Candidate content is crawled, hashed, and compared against registered fingerprints. If similarity exceeds threshold, the system records a piracy event and initiates dispute/royalty logic through smart contracts.

a) User Authentication

User authentication occurs via MetaMask, data is handled by Node.js/Express.js, content is stored on IPFS with metadata in MongoDB, and smart contracts log transactions on Ethereum/Polygon.

b) Content Registration

The creator uploads digital content (image, audio, video).
A SHA-256 and perceptual hash (pHash) of the content are generated. The content is stored on IPFS using Pinata SDK, returning a CID (Content Identifier). The CID and hash are stored on the Ethereum/Polygon blockchain via a smart contract, ensuring tamper-proof ownership records.

c) User Authentication & Interaction

Users interact with the system through a React.js web interface. Authentication is handled using MetaMask and Web3Modal, which connect the user's wallet to the dApp. Blockchain transactions (content registration, piracy reports) are signed through MetaMask.

d) Piracy Detection Pipeline

A crawler scans target platforms for media files. Candidate files are preprocessed (standardized resolution, frame extraction for video, audio fingerprinting). Perceptual hashes are computed using OpenCV/Image Hash/LibROSA. Similarity check is performed using Hamming distance and ensemble voting (multiple hashes combined). If similarity > threshold (e.g., 90%), the file is flagged as pirated.

e) Blockchain Logging and Smart Contract Enforcement

Detection results are logged on-chain through a Solidity smart contract.
In case of confirmed piracy:
Ownership is validated using stored hashes.
Smart contracts may automatically trigger royalty payments or record disputes.

f) Testing & Validation

Postman was used to test API endpoints between frontend, backend, and IPFS. Hardhat environment was used for deploying and testing smart contracts locally. Git was used for version control to ensure replicability of each development stage.

IV. RESULTS AND DISCUSSIONS

The proposed system was implemented as a decentralized application (Dapp) integrating a React.js frontend, Node.js/Express.js backend, MongoDB database, IPFS (Pinata SDK) storage, and Ethereum-compatible smart contracts deployed on the Sepholia test net. Below we present the observed results and key insights from the deployment and testing of the system.

A. End-to-End Workflow Validation

The implemented pipeline ensured that duplicate detection occurred before blockchain registration. The observed flow was as follows:

a) Wallet Authentication

Users successfully authenticated through MetaMask, enabling interaction with the deployed smart contracts.

b) File Submission & Duplicate Check

Uploaded files were parsed using formidable. SHA-256 and perceptual hashes (pHash) were computed, and MongoDB was queried to reject exact or near-duplicate content. This step prevented users from incurring gas fees for previously registered media [12].

c) Watermarking & IPFS Upload

Unique files were watermarked client-side, then uploaded to IPFS. The returned CID was converted into both ipfs:// URIs and gateway-accessible URLs for rendering.
Metadata Management: Metadata JSON was pinned to IPFS, and the resulting token URI was passed to the smart contract for minting.

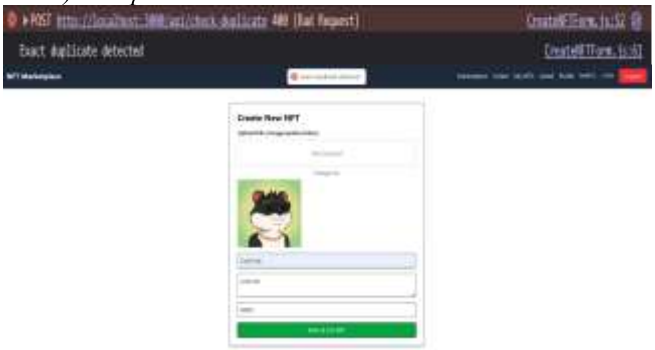
d) On-Chain Enforcement

The smart contract stored file hashes on-chain, ensuring immutability and rejecting duplicate mints at the protocol level. Observed behavior confirmed that if two users attempted to mint the same file simultaneously, the second transaction reverted on-chain, thus eliminating race-condition exploits.

B. Observed Issues and Fixes

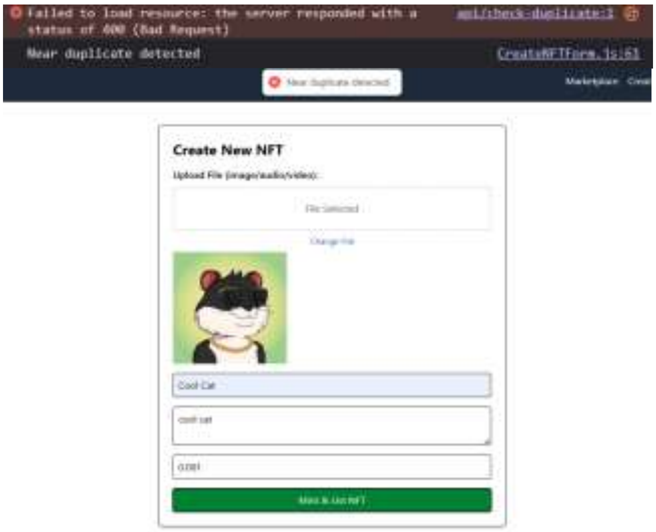
During testing, several critical observations were made:

a) Duplicate Prevention



```
POST http://localhost:3000/api/check-duplicate 400 (Bad Request)      CreateNFTForm.js:52
Exact duplicate detected                                              CreateNFTForm.js:61
```

Figure 2: Exact duplicates



```
Failed to load resource: the server responded with a      api/check-duplicate:1
status of 400 (Bad Request)
Near duplicate detected                                   CreateNFTForm.js:61
```

Cool Cat
cool cat
0.001

Figure 3: Near duplicates

The combination of SHA-256 and pHash ensured robust detection of both exact and near-duplicate files, though thresholds for pHash similarity required tuning [12].

b) IPFS Access

Direct use of ipfs:// led to fetch errors in browsers. The issue was mitigated by consistently resolving URIs into gateway URLs prior to metadata retrieval.

c) Security Concerns

Client-side use of Pinata API keys introduced vulnerabilities. The recommended fix was to move upload logic to server-side APIs to protect secrets.

d) Database Lifecycle

Initial implementation lacked token Id mapping in MongoDB. Updating the database post-mint allowed the system to return existing token Ids for duplicates, improving user feedback.

C. System Implications

These results show that blockchain-based duplicate prevention is not only theoretically sound but practically enforceable. The layered checks (off-chain database, on-chain hash registry) provide strong defense-in-depth against piracy attempts. By integrating IPFS for content storage and MongoDB for metadata caching, the system balances decentralization with performance.

D. Limitations and Future Enhancements

While the system performed well for image datasets, perceptual hashing for audio and video remains incomplete. Future improvements include integrating Chroma print for audio fingerprinting and video perceptual hashing (e.g., FFmpeg frame sampling + pHash). Scalability testing on larger datasets and cost optimization via Layer-2 rollups will also be explored [12].

V. CONCLUSION

This paper aimed to develop a decentralized piracy detection system using perceptual and learned fingerprinting, IPFS storage, and blockchain smart contracts to overcome limitations of traditional DRM-based methods. The system reliably detects piracy in images, audio, and short videos, ensuring transparent provenance and automated enforcement while balancing scalability and cost. Our work contributes a reproducible, scalable framework that enhances traceability and royalty management in digital content protection. Future research should extend detection to streaming media, optimize blockchain costs, improve fingerprinting techniques, and strengthen adversarial resilience.

REFERENCES

[1] B. V. V. R. Kumar, B. A. Vardhan, C. H. R. Gupta, and P. Surekha, "Reduction of Movie Piracy using an Automated Anti-piracy Screen Recording System," in Proc. 4th Int. Conf. Information Systems and Computer Networks (ISCON), Mathura, India, Nov. 2019, pp. 1–4

[2] Y. Chen, G. Zhai, Z. Gao, K. Gu, W. Zhang, M. Hu, and J. Liu, "Movie Piracy Tracking using Temporal Psychovisual Modulation," in Proc. IEEE Int. Symp. Broadband Multimedia Systems and Broadcasting (BMSB), Cagliari, Italy, Jun. 2017, pp. 1–4

[3] P. S. Thakur and S. Kumar, "Technique for Estimation of the Position of the Pirate In-Theater Piracy," in Proc. 2nd Int. Conf. Advances in Computing and Communication Engineering (ICACCE), Dehradun, India, May 2015, pp. 515–518.

[4] Amna Qureshi and David Megías Jiménez, "Blockchain-Based Multimedia Content Protection: Review and Open Challenges," Applied Sciences, vol. 11, no. 1, p. 1, 2021.

[5] M. D. M. Shamalka, K. Banujan, and B. T. G. S. Kumara, "Blockchain and smart contract-based approach to mitigate software piracy," Unpublished manuscript, 2024.

[6] T. Madushanka, D. S. Kumara, and A. A. Rathnaweera, "SecureRights: A Blockchain-Powered Trusted DRM Framework for Robust Protection and Asserting Digital Rights," arXiv preprint arXiv:2403.06094, Mar. 2024

[7] K. Devendra, T. Ghag, and H. Shinde, "Blockchain-based Copyright Detection System: A Decentralized Approach to Protecting Intellectual Property," 2024 International Conference on Innovations and Challenges in Emerging Technologies (ICICET), June 2024.

[8] R. Ashtagi, T. Deshmukh, S. Jadhav, and S. Banerjee, "Blockchain-Based Music Streaming Services: A Novel Approach to Transparent Artist Control," Proceedings of the 2024 International Conference on Emerging Technologies in Music and Media, Oct. 2024

[9] L. Bin, M. A. I. Yasin, and S. N. A. Rahman, "Exploring Blockchain-Based Applications for Digital Copyright Protection," International Journal of Academic Research in Business and Social Sciences, vol. 13, no. 8, pp. 1145–1157, Aug. 2023

[10] K. Y. Abeywardena, Y. Jayasinghe, T. Munasinghe, S. Mannage, and T. Warnasooriya, "VANGUARD: A Blockchain-Based Solution to Digital Piracy," Global Journal of Computer Science and Technology: E Network, Web & Security, vol. 20, no. 4, pp. 19–28, 2020.

[11] A. Kaushik and M. Malik, "Securing The Transfer and Controlling the Piracy of Digital Files using Blockchain," International Journal of Computer Applications, vol. 182, no. 30, pp. 1–5, Dec. 2024.

[12] Tien Luong Trinh, Minh Thanh Ta, "Phash-enhanced Merkle Tree: An Advanced Approach for Detecting and Preventing Copyright Violations in Nft Marketplaces," Research Square, Oct. 3, 2025.

[13] S. McKeown, P. Aaby, and A. Steyven, "PHASER: Perceptual hashing algorithms evaluation and results – An open source forensic framework," Forensic Science International: Digital Investigation, vol. 48, p. 301680, 2024.

[14] H. Taherdoost, "Non-Fungible Tokens (NFT): A Systematic Review," Information, vol. 14, no. 1, p. 26, Dec. 2022, doi: 10.3390/info14010026.

[15] M. Gupta, "Reviewing the Relationship Between Blockchain and NFT With World Famous NFT Market Places," Distributed Computing Blockchain, vol. 1, no. 1, Dec. 2023.

[16] Z. Zheng, S. Xie, H. Dai, X. Chen, and H. Wang, "An Overview of Blockchain Technology: Architecture, Consensus, and Future Trends," in Proc. 2017 IEEE 6th Int. Congress on Big Data (BigData Congress), Honolulu, HI, USA, 2017, pp. 557–564.

[17] W. Uriawan, R. Ramadita, R. D. Putra, and R. I. Siregar, "Authenticate and Verification Source Files using SHA256 and HMAC Algorithms," Preprints, Jul. 2024.

[18] S. McKeown and W. J. Buchanan, "Hamming distributions of popular perceptual hashing techniques," Forensic Science International: Digital Investigation, vol. 45, p. 301544, 2023.

[19] M. N. Beete, D. Matthias, and O. E. Bennett, "Detecting the Source of Digital Images Using Perceptual Hashing and Blockchain Technology," International Journal of Advances in Engineering and Management (IJAEM), vol. 6, no. 7, pp. 573–584, Jul. 2024.

[20] L. Luo, "Application of Blockchain Technology in Intellectual Property Protection," *Journal*, Shanxi Vocational University of Engineering Science and Technology, Jinzhong, China, received Apr. 14, 2022; revised May 1, 2022; accepted May 10, 2022; published Jun. 8, 2022.