Blockchain-Based Certificate Authenticity Validation Using OCR and AI-Assisted Tampering Detection

Prof. P. R. Patil[1], Om Balgude[2], Aryan Patil[3], Neeraj Yamaji
[1,2,3,4]*Department of Computer Engineering, TSSM's Bhivarabai Sawant College of Engineering and Research, Narhe, Pune, India*

Abstract—**The rapid growth of digital education systems and online recruitment platforms has significantly increased the risk of forged and manipulated academic certificates. Traditional certificate verification methods are mostly manual, time-consuming, and dependent on centralized authorities, making them inefficient and vulnerable to tampering. As a result, educational institutions, employers, and verification agencies face major challenges in ensuring the authenticity and integrity of academic credentials. To address these issues, this research proposes a blockchain-based certificate authenticity validation system integrated with Optical Character Recognition (OCR) and Artificial Intelligence (AI) techniques for secure and automated document verification. The proposed system combines blockchain technology, OCR, smart contracts, and intelligent document analysis to detect forged or altered academic certificates in a reliable and transparent manner. Trusted certificate records are maintained using decentralized blockchain-based verification mechanisms, ensuring immutability and protection against unauthorized modifications. When a certificate is uploaded to the system, OCR technology is used to extract important information such as the student's name, institution name, certificate ID, and issue date. The extracted information is then validated against trusted institutional records using blockchain-based verification and intelligent validation techniques. In addition, the system performs document integrity analysis to identify any signs of tampering or unauthorized modifications in certificate files. The proposed platform consists of multiple integrated components, including a user interface developed using React.js, a backend server built with Node.js and Express.js, a MongoDB database, and an AI-powered verification service implemented using FastAPI. The blockchain layer is implemented using Ethereum smart contracts to securely store certificate hashes and maintain immutable verification records. The system also incorporates role-based access control, activity logging, and real-time verification notifications to improve transparency, accountability, and security. By storing document hashes on the blockchain, the system ensures that certificate records cannot be modified without authorization. The proposed solution aims to provide a decentralized, tamper-resistant, scalable, and cost-effective framework for certificate verification that can be used by educational institutions, recruiters, and organizations. Experimental analysis demonstrates that the integration of OCR with blockchain technology improves the reliability and efficiency of the verification process, reduces verification time, and enhances the accuracy of document authenticity validation when compared to traditional verification systems. Overall, the proposed blockchain-based certificate authenticity validation system offers a secure and transparent solution for issuing, storing, and verifying academic certificates. The integration of blockchain technology and OCR significantly reduces the possibility of certificate fraud and provides a trustworthy verification mechanism suitable for modern digital education and recruitment ecosystems.**

I. INTRODUCTION

In recent years, digital technologies and online verification systems have significantly transformed the way academic credentials are created, managed, and shared. Educational institutions, employers, and government organizations increasingly rely on academic certificates for admissions, recruitment, identity verification, and professional qualification validation. However, the rapid advancement of digital editing tools and document manipulation techniques has also led to a substantial increase in fake and manipulated academic certificates. These fraudulent credentials can result in the recruitment of unqualified individuals, financial losses, reputational damage, and a decline in trust within educational and professional ecosystems [1], [2]. Traditional certificate verification methods are primarily manual, time-consuming, and

dependent on centralized databases. Such systems are often inefficient and vulnerable to tampering, unauthorized access, and cyber-attacks. In most conventional systems, certificate records are stored and managed by a single authority, creating a centralized architecture with several limitations, including single points of failure, lack of transparency, risks of unauthorized modifications, and inefficient verification workflows. Furthermore, manual verification processes usually require direct communication between institutions and employers, which increases operational complexity and causes delays in authentication. Since scanned certificates and digital documents can easily be edited or forged, visual inspection alone is no longer sufficient for determining the authenticity of academic credentials [1], [2], [6]. Blockchain technology has emerged as a promising solution for secure, transparent, and tamper-resistant data management due to its decentralized and immutable architecture [9], [10]. Blockchain-based certificate verification systems ensure that once certificate information is recorded on the blockchain, it cannot be modified without consensus from network participants. Smart contracts further automate certificate issuance, validation, and revocation processes, thereby improving efficiency, security, and trust among stakeholders [1], [3]. Recent research studies have proposed various blockchain-based educational certificate verification frameworks using technologies such as Ethereum, Hyperledger Fabric, InterPlanetary File System (IPFS), and other decentralized architectures to establish secure and trustworthy credential management systems [2], [4], [12]. In addition to blockchain, Optical Character Recognition (OCR) and image-processing techniques play a crucial role in automated document analysis and authenticity verification. OCR enables the extraction of textual information from scanned certificates and physical documents, while image-processing and computer vision algorithms assist in detecting anomalies, inconsistencies, and possible tampering attempts [5], [13], [14]. Previous studies have shown that integrating OCR with intelligent image analysis can significantly improve the accuracy and reliability of automated certificate verification systems [5].

This research proposes a blockchain-based certificate authenticity validation system integrated with OCR and AI-assisted document analysis techniques. The proposed system aims to provide a decentralized, efficient, and secure certificate verification framework by combining trusted institutional record management, deterministic hashing, smart contracts, OCR-based field extraction, confidence scoring, tampering detection, and role-based access control mechanisms [1], [2], [3]. When certificates are uploaded, they are processed using OCR and image analysis techniques to extract necessary information such as student details, institution name, certificate ID, and issue date. The extracted information is then compared with trusted institutional records stored in the database and optionally anchored on the blockchain to ensure immutable verification [1], [5].

The proposed system architecture consists of a React.js-based frontend interface, a backend server developed using Node.js and Express.js, a MongoDB database, a Python FastAPI-based AI microservice, and Ethereum smart contract integration. The AI module supports OCR-based extraction, contextual validation, template consistency analysis, and heuristic tampering detection for certificate verification. The blockchain layer provides secure storage and transparent verification of certificate hash values [1], [2], [5]. Additionally, role-based access control, audit logging, and real-time verification notifications are implemented to improve system security and accountability.

By integrating blockchain technology with OCR and AI-assisted verification mechanisms, the proposed system aims to reduce certificate fraud, improve trust in digital credential validation, and provide a cost-effective framework for certificate verification. The system can assist educational institutions, employers, and organizations in secure and efficient certificate verification while preventing the use of forged or manipulated credentials. Overall, the proposed solution has the potential to significantly improve the reliability, transparency, and efficiency of modern academic credential verification systems [1], [2], [3], [5].

II. LITERATURE REVIEW

Academic credential verification has become a major challenge in the modern education and employment ecosystem. Traditional methods of certificate verification are often manual, time-consuming, costly, and highly vulnerable to forgery and manipulation. With the increasing number of graduates and the rapid

digitization of educational systems, organizations and institutions face significant difficulties in validating the authenticity of academic documents. Fraudulent certificates not only damage the credibility of educational institutions but also create unfair advantages in recruitment and higher education opportunities [1], [2], [6].

Blockchain technology has emerged as a promising solution to address these issues due to its decentralized, immutable, transparent, and tamper-resistant nature [9], [10]. By storing academic records on a distributed ledger, blockchain ensures that certificates cannot be altered once issued. The technology eliminates dependence on centralized authorities and enables secure verification of credentials through consensus mechanisms and cryptographic validation. Furthermore, blockchain-based systems enhance trust, traceability, accountability, and transparency in academic credential management [1], [3], [10].

To improve storage efficiency and scalability, several researchers have integrated blockchain with the InterPlanetary File System (IPFS). IPFS provides decentralized off-chain storage where certificate files are securely stored, while their unique cryptographic hash values are maintained on the blockchain. This approach reduces the storage burden on blockchain networks while maintaining data integrity and security. Companies, institutions, and employers can instantly verify certificates using blockchain records and IPFS-based document retrieval systems [4], [8].

Recent studies have proposed various blockchain-powered academic verification systems. The Verifi-Chain model introduced a decentralized certificate verification framework using Ethereum blockchain and IPFS, where certificates are stored securely and verified using unique hash codes [1], [4]. Similarly, ShikkhaChain proposed a blockchain-powered credential verification platform for Bangladesh that supports certificate issuance, revocation, and verification through Ethereum smart contracts, QR-code validation, and role-based access control [2], [3].

Several systematic review studies have also highlighted the growing adoption of blockchain in higher education. Researchers emphasize that blockchain can significantly reduce diploma fraud, improve administrative efficiency, and establish globally trusted educational ecosystems [6], [11]. However, these studies also identify major challenges such as governance issues, scalability limitations, lack of interoperability, privacy concerns, and the absence of standardized regulatory frameworks [2], [8], [12].

In addition, blockchain technology is increasingly being explored for educational governance, accreditation systems, and micro-credentialing. Modern educational ecosystems require flexible, transparent, and interoperable systems capable of supporting lifelong learning, digital credentials, and international academic mobility. Blockchain-based governance frameworks can improve trust, decentralization, and accountability among educational institutions, employers, accreditation agencies, and governments [3], [6], [11].

The proposed blockchain-based certificate verification systems generally include multiple layers such as user interfaces, smart contracts, digital wallets, verification portals, and decentralized storage systems. Smart contracts automate certificate issuance, validation, and revocation, while decentralized applications (DApps) provide secure interaction among stakeholders including students, institutions, employers, and regulatory bodies [1], [3], [10].

Despite the advantages of blockchain technology, researchers have come across several limitations and challenges. These include scalability issues, high transaction costs, interoperability challenges between blockchain platforms, governance complexities, privacy concerns, and lack of global standards [2], [8], [12]. Moreover, many existing systems focus primarily on technical implementation while giving limited attention to policy frameworks, governance mechanisms, and socio-economic factors influencing adoption.

Overall, blockchain combined with decentralized storage technologies such as IPFS presents a secure, transparent, and efficient solution for academic certificate verification. The integration of smart contracts, decentralized governance, and tamper-proof digital records can significantly improve trust and reliability in educational credential management systems while reducing fraud and administrative overhead [1], [4], [10]. However, future research is still required to address challenges related to governance, interoperability, privacy, and large-scale institutional adoption [2], [8], [12].

III. SYSTEM ARCHITECTURE AND DESIGN

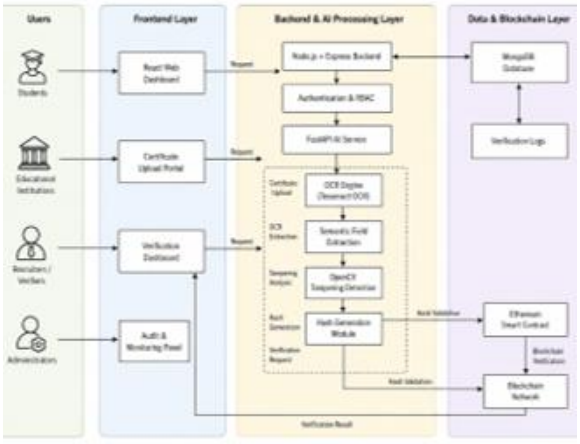

Fig. 1. System Architecture

The proposed certificate authenticity validation system is designed as a decentralized and intelligent framework that combines blockchain technology, Optical Character Recognition (OCR), image-processing techniques, and artificial intelligence for secure and reliable certificate verification. The architecture is developed to ensure transparency, security, scalability, and efficient validation of academic credentials. Each component of the system works together to automate the verification process and detect forged or manipulated certificates with high accuracy [1], [2], [5].

The overall architecture of the system consists of multiple layers, including the user interface layer, communication layer, application layer, AI-based verification layer, database layer, and blockchain layer. These layers communicate securely through APIs and real-time communication mechanisms to ensure seamless interaction among users, verification services, and blockchain networks. The layered architecture improves modularity, maintainability, and scalability of the system [1], [3].

A. User Interface Layer

The user interface layer provides an interactive web-based platform through which users can upload certificates, request verification, and view verification results. The frontend is developed using React.js to create a responsive and user-friendly interface. Real-time communication and status updates are implemented using Socket.IO, enabling users to receive instant notifications regarding certificate verification progress and results. The interface is designed to support different user roles such as administrators, institutions and recruiters through role-based access control mechanisms [1], [2].

B. Application and Backend Layer

The application layer acts as the core processing unit of the system and is responsible for handling user requests, authentication, certificate management, and communication between different system components. The backend is implemented using Node.js and Express.js, which provide scalable server-side processing and API management. This layer validates user permissions, manages verification workflows, coordinates communication with the AI verification service, and interacts with the blockchain layer for certificate validation and storage operations [1], [3].

The backend also performs data synchronization, request handling, activity logging, and real-time notification management. Secure authentication and authorization mechanisms are implemented to ensure that users can only access functionalities permitted according to their roles and permissions.

C. AI-Based Verification Layer

The AI-based verification layer is responsible for document analysis, OCR processing, and tampering detection. This module is implemented using Python and FastAPI to provide efficient and scalable AI services. OCR techniques are used to extract important information from uploaded certificates, including student name, institution name, certificate ID, issue date, and course details [5], [14].

After data extraction, the AI module performs contextual validation, template consistency analysis, anomaly detection, and heuristic tampering analysis. Image-processing algorithms and computer vision techniques are used to detect possible modifications, inconsistencies, suspicious patterns, or alterations in certificate documents. The AI layer significantly improves the accuracy and reliability of automated certificate verification [5], [13].

D. Database Layer

The database layer stores all certificate records, user information, verification logs, and system metadata. MongoDB is used as the primary database due to its flexibility, scalability, and efficient handling of

structured and semi-structured data. The database maintains trusted institutional certificate records that are used for comparison during the verification process. Verification history, extracted OCR data, tampering analysis results, user activities, and blockchain transaction references are also stored in the database for auditing and traceability purposes. The database layer supports fast data retrieval and efficient management of large volumes of certificate records [1], [4].

E. Blockchain Layer

The blockchain layer provides decentralized and tamper-resistant certificate validation using Ethereum smart contracts. Certificate hashes generated through deterministic hashing algorithms are securely stored on the blockchain to ensure immutability and transparent verification. Once certificate information is anchored on the blockchain, it cannot be modified without network consensus, thereby preventing unauthorized changes or forgery [1], [9], [10].

Smart contracts automate certificate registration, validation, and verification operations while maintaining transparency and trust among stakeholders. The blockchain layer also maintains audit trails of verification activities, making the system more secure and accountable [3], [10].

F. Certificate Verification Workflow

When a user uploads a certificate, the system initiates a multi-stage verification workflow. First, the uploaded document is processed by the AI verification module, where OCR techniques extract relevant textual information and image-processing algorithms analyze the certificate for anomalies or signs of tampering [5], [13], [14].

Next, the backend layer compares the extracted information with trusted institutional records stored in the MongoDB database. Simultaneously, the certificate hash is validated against blockchain records through Ethereum smart contracts to ensure authenticity and integrity. If inconsistencies or unauthorized modifications are detected, the system flags the certificate as suspicious [1], [2].

After completing all verification stages, the system generates a verification result indicating whether the certificate is authentic, manipulated, or unverifiable. The verification details, analysis reports, and transaction records are then securely stored in the database and blockchain layers for future reference and auditing.

G. System Scalability and Security

The proposed system is designed to be scalable, secure, and efficient. The modular layered architecture enables the platform to support multiple concurrent users and process large numbers of certificate verification requests simultaneously. Real-time communication, distributed verification mechanisms, and decentralized blockchain storage improve overall system performance and reliability [1], [3], [4].

Security features such as role-based access control, secure APIs, immutable blockchain records, logging mechanisms, and AI-assisted tampering detection ensure strong protection against fraud and unauthorized access. The combination of blockchain technology, OCR, and AI-based verification provides a robust and transparent framework for academic certificate validation [1], [2], [5].

Overall, the proposed system offers a scalable, secure, and intelligent solution for certificate authenticity verification. By integrating decentralized blockchain technology with automated OCR and AI-assisted document analysis, the system significantly improves the efficiency, transparency, and trustworthiness of academic credential verification processes [1], [2], [3], [5].

IV. SYSTEM METHODOLOGY

The proposed certificate authenticity validation system uses a combination of blockchain technology, Optical Character Recognition (OCR), artificial intelligence, image-processing techniques, smart contracts, and database management to ensure secure and reliable certificate verification. The methodology is designed to automate the verification process while improving transparency, scalability, and resistance against certificate forgery and tampering. The complete workflow integrates multiple technologies that work together to validate the authenticity of academic certificates efficiently [1], [2], [3].

The overall methodology of the system consists of several stages, including certificate registration, document preprocessing, OCR-based information extraction, tampering analysis, blockchain validation, and final authenticity verification. Each stage contributes to improving the accuracy and security of

the verification process while reducing the dependency on manual validation methods [1], [5].

A. Certificate Registration Process

The verification process begins when an authorized educational institution registers certificate information within the system. The uploaded information includes important details such as the student's name, institution name, certificate ID, course information, and issue date. Before storage, the system standardizes the certificate information into a structured format to ensure consistency and efficient comparison during future verification operations.

After preprocessing, the system generates a unique cryptographic hash value from the certificate data using deterministic hashing algorithms. This hash functions as a digital fingerprint of the certificate and uniquely represents the document content. The generated hash is stored in both the MongoDB database and the blockchain network through Ethereum smart contracts. This process ensures data immutability and prevents unauthorized modifications to certificate records [1], [9], [10].

B. Certificate Upload and Document Preprocessing

When a user wants to verify a certificate, the certificate document is uploaded through the web-based interface. The uploaded document first undergoes preprocessing operations to improve the quality and readability of the image before OCR analysis.

Document preprocessing includes image enhancement, noise removal, brightness correction, resizing, edge sharpening, and distortion reduction. These techniques improve OCR accuracy by making the text clearer and easier to extract. Preprocessing also helps handle certificates uploaded in different formats, resolutions, and lighting conditions [5], [13], [14].

C. OCR-Based Information Extraction

After preprocessing, the system applies Optical Character Recognition (OCR) techniques to extract textual information from the certificate image. The OCR module identifies and extracts important fields such as student name, institution name, certificate ID, issue date, and course details.

The extracted information is converted into machine-readable structured data for further analysis. OCR automation significantly reduces manual effort and enables rapid processing of large numbers of certificate documents [5], [14].

D. AI-Assisted Tampering Detection

Once the certificate information is extracted, the system performs AI-assisted document analysis to identify possible signs of forgery or tampering. Image-processing and computer vision techniques are used to detect anomalies such as inconsistent fonts, mismatched alignments, image alterations, irregular patterns, and suspicious modifications within the document [5], [13].

The AI module performs contextual validation and template consistency analysis to compare uploaded certificates against trusted certificate structures. If unusual inconsistencies or tampering indicators are detected, the certificate is flagged as suspicious for further verification.

E. Hash Generation and Blockchain Validation

After OCR extraction and tampering analysis, the system generates a new cryptographic hash from the extracted certificate data. This newly generated hash is then compared with the original hash values stored in both the MongoDB database and the blockchain network. The blockchain layer uses Ethereum smart contracts to validate certificate authenticity. Since blockchain records are immutable and decentralized, any modification in certificate data results in a mismatch between the generated hash and the stored hash. If both hash values match successfully, the certificate is considered authentic. Otherwise, the certificate is identified as manipulated, forged, or suspicious [1], [2], [10].

Smart contracts automate the validation process and maintain transparent verification records. The blockchain network ensures that certificate information cannot be altered without consensus from network participants, thereby enhancing trust and security within the verification ecosystem [3], [10].

F. Final Verification Decision

The final authenticity decision is based on multiple verification parameters, including OCR extraction confidence, AI-based tampering analysis, contextual validation results, and blockchain hash verification. The system combines these factors to generate a final verification outcome indicating whether the certificate is authentic, suspicious, manipulated, or unverifiable.

Verification results are displayed to users through the frontend interface and are simultaneously stored in the database for future reference, auditing, and reporting purposes. The system also maintains activity logs and transaction histories to improve transparency and accountability.

G. Security and Scalability Considerations

The proposed methodology is designed to support secure and scalable certificate verification operations. Blockchain technology provides tamper-resistant storage and decentralized validation, while role-based access control mechanisms ensure secure access management for institutions, administrators and recruiters [1], [2], [10].

The system architecture supports concurrent certificate verification requests and can process large volumes of certificates simultaneously. Automated OCR extraction, AI-assisted analysis, and smart contract-based validation significantly reduce verification time and operational overhead compared to traditional manual verification systems [2], [5].

Overall, the proposed methodology combines blockchain technology, OCR, artificial intelligence, and smart contracts to create a secure, scalable, and transparent certificate authenticity validation framework. The integration of these technologies improves verification accuracy, reduces certificate fraud, and provides an efficient solution for modern academic credential validation systems [1], [2], [3], [5].

V. IMPLEMENTATION

The implementation of the proposed certificate authenticity validation system was carried out using a hybrid technology stack that integrates blockchain technology, OCR-based document analysis, AI-assisted tampering detection, decentralized verification mechanisms, and real-time communication services. The system was developed as a modular web-based platform consisting of frontend, backend, AI verification, database, and blockchain subsystems. The implementation primarily focuses on secure certificate verification, automated document analysis, immutable validation, and scalable deployment suitable for real-world academic credential verification environments.

A. Frontend Implementation

The frontend layer of the system was implemented using React.js, Vite, Tailwind CSS, and Socket.IO to provide responsive user interfaces and real-time verification monitoring capabilities. The frontend architecture supports role-based dashboards for institutions, administrators, recruiters, and verification authorities. Users can upload certificates, request verification, monitor audit logs, and visualize verification results through an intuitive web-based interface. Real-time communication between the frontend and backend is handled using Socket.IO, which enables live updates related to certificate verification progress, tampering alerts, blockchain validation status, and system notifications. REST APIs and secure HTTP requests are used for communication between frontend services and backend modules.

B. Backend Implementation

The backend subsystem was implemented using Node.js and Express.js and functions as the core orchestration and control layer of the system. The backend handles authentication, authorization, certificate processing, blockchain interaction, audit logging, and integration with AI verification services. JWT-based authentication and role-based access control (RBAC) mechanisms were implemented to ensure secure access to system functionalities and administrative operations.

Multiple REST API endpoints were developed for certificate uploads, verification requests, institutional management, audit retrieval, and blockchain validation operations. Express controllers and middleware modules manage certificate workflows and facilitate communication between the frontend, AI verification microservice, database layer, and blockchain subsystem. Additionally, Socket.IO services were integrated within the backend to stream real-time notifications and verification updates to connected frontend clients.

C. AI Verification and OCR Implementation

The AI verification subsystem was implemented as a FastAPI-based Python microservice responsible for OCR extraction, image preprocessing, tampering detection, confidence scoring, and forensic analysis. The AI module integrates multiple libraries and frameworks, including Tesseract OCR, OpenCV, Pillow, NumPy, and PyMuPDF, for efficient

document analysis and semantic information extraction. Uploaded certificate images and PDF documents undergo several preprocessing operations such as grayscale conversion, Gaussian filtering, thresholding, segmentation, and morphological transformations to improve OCR accuracy. The OCR engine extracts important certificate details including student name, institution name, certificate ID, issue date, academic program, and grade information from uploaded documents.

In addition to OCR extraction, the AI verification service performs heuristic tampering analysis and anomaly detection using OpenCV-based forensic techniques. The implementation analyzes uploaded certificates for font inconsistencies, image mismatches, suspicious overlays, editing artifacts, template anomalies, and structural irregularities. Confidence values are generated based on OCR reliability metrics and detected inconsistencies. Certificates with abnormal confidence scores are automatically flagged as suspicious and may require manual review. This implementation improves fake certificate detection by combining semantic verification with visual forensic analysis rather than relying solely on blockchain hash validation.

D. Database Implementation

MongoDB was implemented as the primary database management system for storing certificate records, institutional profiles, user data, verification logs, audit trails, and role-based access control information. The database schema was designed to support indexed querying and optimized retrieval of trusted certificate records during verification operations. Uploaded certificate documents are temporarily stored within the upload's module before OCR processing and tampering analysis. The database layer also stores OCR extraction results, blockchain transaction references, verification reports, confidence scores, and audit histories for transparency and traceability. Redis caching mechanisms can additionally be integrated for session management, real-time queue handling, and performance optimization.

E. Blockchain Implementation

The blockchain subsystem was implemented using Ethereum smart contracts integrated through ethers.js APIs. Smart contracts were developed to manage certificate hash registration, immutable verification, issuer validation, and certificate revocation workflows. Deterministic semantic hashes generated from normalized certificate data are securely stored on the blockchain to ensure tamper-resistant certificate validation. During certificate verification, hashes generated from uploaded certificate data are compared against trusted blockchain records to determine authenticity. Blockchain-based validation eliminates dependency on centralized trust systems and significantly improves transparency, immutability, and integrity within academic credential verification environments. The implementation also supports integration with decentralized storage technologies such as IPFS and distributed document management architectures for secure certificate sharing and immutable file storage. These decentralized storage mechanisms further enhance data security and prevent unauthorized modifications to uploaded documents.

F. Verification Workflow Implementation

The complete implementation workflow begins when an authorized institution uploads trusted certificate records into the system database. Certificate information is normalized and converted into deterministic semantic hashes before blockchain registration. When a user uploads a certificate for verification, the backend forwards the document to the AI verification microservice for OCR extraction and tampering analysis. The extracted information is semantically compared with trusted institutional records stored in MongoDB, while generated certificate hashes are validated against blockchain smart contracts. The system then generates a detailed verification report containing OCR confidence scores, confidence scores, tampering analysis results, blockchain validation status, and final authenticity decisions. Verification results are displayed on the frontend dashboard in real time and are simultaneously stored within audit logs for future reference, transparency, and traceability.

G. Overall System Implementation

The implemented system successfully demonstrates the integration of blockchain technology, OCR extraction, AI-assisted tampering detection, decentralized trust management, and real-time certificate verification within a unified architecture. The modular implementation design improves scalability, maintainability, interoperability, and

deployment flexibility while providing a secure, intelligent, and efficient solution for academic certificate authenticity validation.

VI. RESULTS AND ANALYSIS

The proposed blockchain-based certificate authenticity validation system was evaluated to analyze its effectiveness in certificate verification, OCR-based information extraction, tampering detection, blockchain validation, and overall system performance. The experimental analysis was conducted using a dataset consisting of genuine academic certificates, forged certificates, low-quality scanned documents, and synthetically manipulated certificates created for testing purposes. The evaluation focused on OCR extraction accuracy, tampering detection capability, blockchain verification reliability, response time, and overall certificate verification performance [1], [2], [5].

A. OCR Extraction Performance

The first stage of the evaluation analyzed the performance of the OCR extraction module. The system successfully extracted important certificate information such as student name, institution name, certificate ID, issue date, academic program, and grade information with high accuracy. Before OCR extraction, uploaded certificate documents underwent preprocessing operations including image enhancement, noise removal, thresholding, and segmentation to improve text readability and extraction quality [5], [13], [14].

Experimental results showed that the OCR module achieved an average extraction accuracy of approximately 95.75% across different certificate formats and document qualities. Among all extracted fields, certificate IDs achieved the highest extraction accuracy due to their structured formatting and consistent visual patterns. The integration of OCR preprocessing and image enhancement techniques significantly improved extraction reliability, especially for scanned and partially degraded certificate documents [5], [14].

B. Tampering Detection Analysis

The second phase of testing focused on the system's ability to detect forged and manipulated certificates. Various types of tampered certificates were tested, including certificates with modified names, altered grades, replaced institutional logos, edited image regions, and inconsistent formatting structures.

The AI-assisted forensic analysis module successfully identified most tampering attempts using image-processing techniques, confidence scoring, and heuristic validation mechanisms. The system generated confidence scores based on detected inconsistencies and classified certificates into categories such as authentic, suspicious, tampered, or unverified [5], [13].

Experimental evaluation demonstrated that the tampering detection module achieved an average detection accuracy of approximately 91%. The combination of OCR-assisted semantic analysis and image forensic validation significantly improved fake certificate detection when compared with systems that rely only on blockchain hash verification [1], [5].

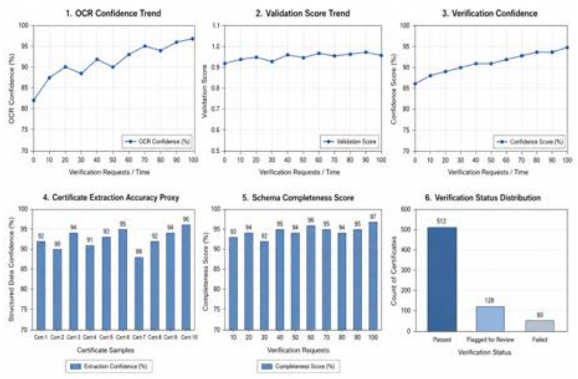

Fig. 2. OCR Accuracy Graph

C. Blockchain Verification Performance

The blockchain verification subsystem was evaluated to analyze the efficiency and reliability of Ethereum smart contract-based certificate validation. Certificate hashes generated from normalized certificate data were successfully stored on Ethereum smart contracts and later validated during verification requests [1], [10]. The system effectively detected modifications in certificate data by comparing uploaded certificate hashes with trusted blockchain records. Experimental results showed that the average time required for blockchain hash registration was approximately 2.1 seconds, while the average certificate verification time was less than 2 seconds. These results indicate that the blockchain layer provides transparent and tamper-resistant certificate validation while maintaining practical performance suitable for real-world deployment environments [1], [2], [10].

D. Overall System Performance

The complete system was evaluated using a mixed dataset containing authentic certificates, forged certificates, and poor-quality scanned documents. The proposed framework achieved an overall certificate verification accuracy of approximately 94% while processing and verifying certificates within an average time of 4.2 seconds per document.

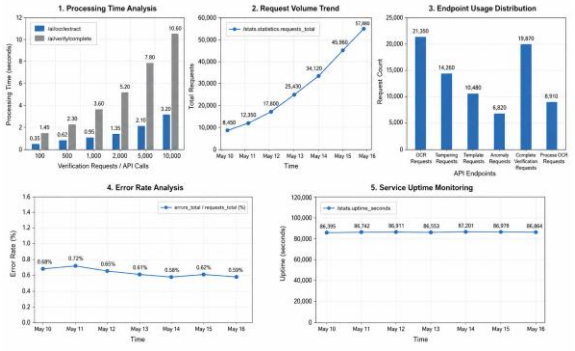

Fig. 3. Overall Performance Graph

The results demonstrate that the proposed system can efficiently automate certificate verification processes while maintaining high reliability and operational efficiency. The integration of OCR extraction, AI-assisted tampering detection, anomaly analysis, and blockchain validation enables the system to provide more comprehensive verification capabilities compared to traditional manual verification systems and blockchain-only verification frameworks [1], [2], [5].

E. Comparative Analysis

Compared to conventional manual verification approaches, the proposed system significantly reduces verification time, minimizes human intervention, and improves transparency in certificate validation. Traditional systems often require direct communication between institutions and employers, resulting in delays and increased operational complexity [1], [2]. Similarly, blockchain-only verification systems mainly focus on validating certificate hashes but may fail to identify visually manipulated certificates where the content has been altered before hash generation. In contrast, the proposed system combines blockchain verification with OCR-based semantic extraction, AI-assisted forensic analysis, and anomaly detection, thereby improving both tampering detection and overall verification reliability [1], [5].

F. Practical Advantages of the Proposed System

The experimental analysis also highlighted several practical advantages of the proposed architecture. The use of blockchain technology improves data integrity, decentralized trust management, and transparency in certificate verification. The OCR and forensic-analysis modules reduce dependency on manual inspection while improving automation and scalability [1], [3], [10]. The system also supports real-time verification notifications, audit logging, and transparent activity monitoring mechanisms, which improve traceability and accountability. Furthermore, the modular architecture allows future integration with advanced AI models, decentralized storage systems such as IPFS, and large-scale institutional verification ecosystems [4], [8].

G. Limitations and Challenges

Despite the promising results, certain limitations were observed during system evaluation. OCR extraction performance decreased for distorted, handwritten, low-resolution, or highly compressed certificate documents. Additionally, the tampering-analysis module occasionally generated false positives for certificates containing legitimate template variations or inconsistent printing artifacts [5], [13], [14].

Blockchain verification also introduced minor latency depending on Ethereum network conditions and smart contract execution overhead. These limitations can potentially be addressed in future work through deep-learning-based OCR enhancement, AI-driven forensic analysis models, adaptive template-learning mechanisms, and optimized blockchain deployment strategies [2], [10].

H. Overall Analysis

Overall, the experimental results demonstrate that the proposed system effectively combines blockchain technology, OCR extraction, AI-assisted tampering detection, smart contracts, and decentralized verification into a unified certificate authenticity validation framework. The proposed architecture provides a secure, transparent, tamper-resistant, and automated solution suitable for educational institutions, recruiters, government agencies, and digital credential verification platforms [1], [2], [3], [5].

VII. DISCUSSION

The implementation of OCR and image preprocessing techniques further improves the efficiency of certificate verification by automating information extraction from scanned certificates and PDF documents. Automated OCR extraction significantly reduces manual effort and operational delays associated with traditional verification procedures. The system can rapidly extract certificate information such as student name, institution name, certificate ID, issue date, and academic details while simultaneously performing authenticity validation. This automation enables faster processing and improves scalability for institutions handling large volumes of certificate verification requests [5], [13], [14].

Another important advantage of the proposed system is its decentralized and transparent architecture. Traditional certificate verification systems often rely on centralized databases maintained by single authorities, creating single points of failure and increasing vulnerability to cyberattacks or unauthorized modifications. By leveraging blockchain technology and decentralized validation mechanisms, the proposed framework improves transparency, trustworthiness, and auditability within credential verification ecosystems. Real-time verification notifications, audit logs, and immutable blockchain transaction records further enhance accountability and traceability [1], [2], [10].

The proposed system also demonstrates practical applicability in real-world academic and recruitment environments. Experimental results indicate that the system achieves high OCR extraction accuracy, reliable tampering detection performance, efficient blockchain validation, and fast verification response times. These results suggest that integrating blockchain technology with OCR and AI-assisted analysis provides a practical and scalable solution for secure academic credential verification [1], [2], [5].

Despite its advantages, the proposed system still faces certain limitations and challenges. OCR extraction performance decreases for handwritten, distorted, low-resolution, or highly compressed certificate documents. Similarly, heuristic tampering-analysis techniques may occasionally generate false positives when certificates contain legitimate template variations, inconsistent printing patterns, or document degradation artifacts. Blockchain operations may also introduce minor latency depending on Ethereum network conditions, gas fees, and smart contract execution overhead [5], [10], [14].

These limitations highlight several opportunities for future improvement. Advanced deep-learning-based OCR models and transformer-based document understanding techniques can improve extraction accuracy for degraded or complex certificate documents. AI-driven forensic analysis models based on convolutional neural networks and computer vision algorithms can further enhance tampering detection accuracy while reducing false positives. Adaptive template-learning mechanisms and self-supervised anomaly detection models can improve robustness against variations in certificate layouts and institutional formats [5], [13], [14].

Future research can also focus on integrating decentralized storage systems such as IPFS and BigchainDB for secure certificate storage and distributed document sharing. Multi-blockchain interoperability frameworks and consortium blockchain architectures may improve scalability and reduce operational costs for large-scale institutional deployments. Additionally, integrating the proposed system with national educational infrastructure, recruitment platforms, and international credential verification ecosystems could further enhance interoperability and global trust management [4], [8], [12]. The proposed framework represents a significant step toward modernizing academic certificate verification systems. By combining blockchain technology, OCR extraction, AI-assisted tampering detection, smart contracts, and decentralized trust mechanisms, the system provides a secure, transparent, scalable, and automated solution for validating academic credentials. The architecture reduces dependency on manual verification, minimizes certificate fraud, improves operational efficiency, and strengthens trust within educational and professional ecosystems [1], [2], [3], [10].

Overall, the discussion demonstrates that the proposed system provides substantial improvements over conventional verification methods and existing blockchain-only solutions. Although certain technical and operational challenges remain, the integration of advanced AI techniques, decentralized technologies, and intelligent verification mechanisms offers strong potential for future development and large-scale

adoption in global academic credential verification environments [1], [2], [5], [10].

VIII. CONCLUSION

The rapid growth of digital education systems, online recruitment platforms, and credential-sharing ecosystems has significantly increased the challenges associated with fake academic certificates and document tampering. Traditional certificate verification methods are largely manual, time-consuming, centralized, and inefficient for large-scale verification environments. These limitations reduce trust in academic credentials and increase the risk of fraud within educational institutions, recruitment processes, and professional qualification systems.

To address these challenges, this research proposed a blockchain-based certificate authenticity validation system integrated with Optical Character Recognition (OCR), AI-assisted tampering detection, smart contracts, and decentralized verification mechanisms. The proposed framework combines blockchain technology, OCR-based information extraction, image forensic analysis, anomaly detection, and automated verification workflows to provide a secure, transparent, scalable, and intelligent certificate verification solution.

The system was implemented using a modern hybrid architecture consisting of a React.js-based frontend, Node.js and Express.js backend services, a FastAPI-based AI verification microservice, MongoDB database integration, and Ethereum blockchain smart contracts. OCR and image-processing techniques were used to automatically extract important certificate information such as student name, institution name, certificate ID, issue date, and academic details. The AI-assisted tampering detection module improved verification reliability by identifying suspicious modifications, inconsistencies, and document alterations. The blockchain layer provided immutable and decentralized certificate validation by securely storing certificate hashes through Ethereum smart contracts. During verification, uploaded certificate hashes were compared against trusted blockchain records to ensure authenticity and integrity. Smart contracts also automated certificate registration, issuer validation, and verification workflows, thereby improving transparency and reducing dependency on centralized authorities.

Experimental evaluation demonstrated that the proposed system achieved high OCR extraction accuracy, effective tampering detection performance, reliable blockchain validation, and fast certificate verification response times. The integration of blockchain technology with OCR extraction, AI-assisted analysis, and decentralized trust management significantly improved fake certificate detection and operational efficiency when compared to traditional manual verification systems and blockchain-only frameworks. The research findings also highlighted the advantages of integrating multiple verification mechanisms within a unified framework. Features such as real-time verification notifications, audit logging, role-based access control, modular architecture, and automated certificate processing improved the practicality, usability, and scalability of the proposed system. The modular implementation further enables future integration with advanced AI models, decentralized storage systems, and large-scale institutional verification infrastructures.

Despite the promising results, several limitations were identified during system evaluation. OCR performance decreased for poor-quality, handwritten, or highly compressed documents, while heuristic tampering-analysis methods occasionally generated false positives for certificates with legitimate template variations. Blockchain verification latency was also influenced by network conditions and smart contract execution overhead. These limitations can be addressed in future work through advanced deep-learning-based OCR enhancement, AI-driven forensic analysis models, adaptive template-learning mechanisms, and optimized blockchain deployment strategies.

Overall, the proposed blockchain-based certificate authenticity validation system provides a secure, decentralized, scalable, and intelligent solution for academic credential verification. By integrating blockchain technology, OCR extraction, AI-assisted tampering detection, and smart contracts, the system significantly reduces certificate fraud, improves transparency, and strengthens trust within educational and professional ecosystems. The proposed framework has strong potential for future adoption in educational institutions, recruitment platforms, government verification systems, and digital credential management infrastructures.

ACKNOWLEDGMENT

The authors would like to express their sincere gratitude to the project guide for continuous guidance, encouragement, and valuable suggestions throughout this research work. The support, technical expertise, and constructive feedback provided were instrumental in the successful completion of the project. The authors also thank the faculty members and department staff for their assistance and cooperation during the research and development process.

The authors acknowledge the institution for providing the necessary infrastructure, technical resources, laboratory facilities, and a supportive learning environment required for conducting this research. These facilities greatly assisted in the implementation, testing, and evaluation of the proposed system. The authors also gratefully acknowledge the contributions of researchers and scholars whose published work on blockchain technology, Optical Character Recognition (OCR), smart contracts, decentralized verification systems, artificial intelligence, and certificate authentication frameworks served as valuable references for this study. Special thanks are extended to friends, classmates, and family members for their constant encouragement and support. Generative AI tools were used for language refinement and formatting assistance, while all technical content, implementation, experimentation, and analysis are the sole intellectual contributions of the authors.

REFERENCES

[1] A. K. C., D. Bhandari, R. Priyadarshini, and P. K. Meher, "Detection of Fake Physical Certificates using a Blockchain-Based Certificate Verification and Issuer Validation System," in 2024 IEEE International Conference on Blockchain and Distributed Systems Security (ICBDS), 2024, pp. 1–6, doi: 10.1109/ICBDS61829.2024.10837263.

[2] R. Priyadarshini et al., "A Faster, Integrated, and Trusted Certificate Authentication and Issuer Validation System Based on Blockchain," IEEE Access, vol. 13, pp. 1–15, 2025, doi: 10.1109/ACCESS.2025.3539180.

[3] M. J. H. Faruk, J. Basney, and J. Q. Cheng, "Blockchain-Based Decentralized Verifiable Credentials: Leveraging Smart Contracts for Privacy-Preserving Authentication Mechanisms to Enhance Data Security in Scientific Data Access," in 2023 IEEE International Conference on Big Data (BigData), 2023, pp. 5493–5502, doi: 10.1109/BigData59044.2023.10386360.

[4] D. Babrekar, D. Patel, S. Patkar, and V. B. Lobo, "Blockchain-based Digital Locker using BigchainDB and InterPlanetary File System," in 2021 International Conference on Communication and Electronics Systems (ICCES), 2021, pp. 948–953, doi: 10.1109/ICCES51350.2021.9489028.

[5] I. Marín-Aguilar, L. A. Chavarría-Zamora, and L. Araya-Martinez, "A Practical Approach to Validate the Authenticity of Identity Documents," in 2022 IEEE Latin America Electron Devices Conference (LAEDC), 2022, pp. 1–4, doi: 10.1109/LAEDC54796.2022.9908240.

[6] M. Turkanović, M. Hölbl, K. Košič, M. Heričko, and A. Kamišalić, "EduCTX: A Blockchain-Based Higher Education Credit Platform," IEEE Access, vol. 6, pp. 5112–5127, 2018, doi: 10.1109/ACCESS.2018.2789929.

[7] Z. A. Lux, T. Grechenig, and S. Feld, "Cerberus: A Blockchain-Based Accreditation and Degree Verification System," in Proceedings of the International Conference on Blockchain Technology, 2019, pp. 1–6.

[8] A. Shrestha and J. Vassileva, "Blockchain-Based Research Data Sharing Framework using Smart Contracts and Decentralized Verification," Future Generation Computer Systems, vol. 112, pp. 1–10, 2020.

[9] S. Nakamoto, "Bitcoin: A Peer-to-Peer Electronic Cash System," 2008. [Online]. Available: Bitcoin Whitepaper.

[10] G. Wood, "Ethereum: A Secure Decentralised Generalised Transaction Ledger," Ethereum Project Yellow Paper, 2014.

[11] W3C, "Verifiable Credentials Data Model 1.0," World Wide Web Consortium (W3C), 2019. [Online]. Available: W3C Verifiable Credentials Data Model.

[12] Hyperledger Foundation, "Hyperledger Fabric Documentation," 2023. [Online]. Available: Hyperledger Fabric Documentation.

[13] G. Bradski, "The OpenCV Library," Dr. Dobb's Journal of Software Tools, 2000.

[14] R. Smith, “An Overview of the Tesseract OCR Engine,” in Proceedings of the Ninth International Conference on Document Analysis and Recognition (ICDAR), 2007, pp. 629–633.