VeriDoc - AI-Powered Conditional Document Fraud Detection System

1. Problem Statement

Organizations lack an automated, scalable, and explainable mechanism to verify the authenticity and integrity of digital documents across heterogeneous formats. Existing solutions typically fall into one of two categories:

1. **Rule-based or metadata-driven systems** that are fast but shallow and easily bypassed.
2. **Black-box AI classifiers** that provide probabilistic results without technical justification.

Neither approach offers a balance of accuracy, transparency, and forensic reliability. Furthermore, the rise of digitally signed documents creates a new vector for fraud where cryptographic signatures are manipulated or stripped, a challenge existing visual classifiers fail to address. The absence of a unified, adaptive verification framework leads to inconsistent decisions, high manual workload, and limited auditability.

2. Proposed Solution Overview

VeriDoc introduces a **Conditional Forgery Detection Architecture** that dynamically selects the optimal verification strategy based on document structure, content type, and cryptographic status. Rather than applying a uniform pipeline to all inputs, VeriDoc routes documents through specialized forensic, cryptographic, and AI-driven modules to maximize detection accuracy while maintaining computational efficiency.

At the core of the platform is a **FastAPI-based orchestration service** that manages:

- Secure document ingestion
- Format detection and pipeline selection
- **Digital Signature & Certificate Validation**
- AI model inference
- Forensic result aggregation
- Report generation and delivery

The system outputs a structured verification report containing:

- Tampering confidence score
- **Cryptographic validity status**
- Evidence-based findings (highlighted suspicious regions or elements)
- Logical consistency assessment

3. System Architecture

3.1 High-Level Flow

1. **User uploads document** via web interface.
2. **Secure ingestion layer** validates file type, size, and integrity.
3. **Orchestrator identifies document category** (Native PDF, Scanned Image, or Digitally Signed Document).
4. **Routing:** Document is sent to specialized pipelines (Forensic, Visual, or Cryptographic).
5. **Results are aggregated** and evaluated by the AI reasoning layer.
6. **Final report is generated** and delivered to the user.

4. Technical Approach

4.1 Pipeline A - Structural Forensics for Native PDFs

This pipeline targets digitally generated PDFs containing internal object structures and selectable text. Instead of visual inspection, it analyzes the file at the binary and object-reference level.

Key techniques include:

- **Incremental Update Detection:** Identification of multiple end-of-file markers that indicate post-creation append operations.
- **Cross-Reference Table Traversal:** Parsing of Prev pointers to reconstruct hidden revision histories.
- **Orphaned Object Analysis:** Detection of unreferenced objects that remain in the file after content removal or modification.

These methods provide deterministic evidence of tampering that can be reproduced and verified independently.

4.2 Pipeline B - Visual & Statistical Analysis for Raster Documents

This pipeline processes scanned PDFs, JPEGs, PNGs, and screenshots where structural metadata is unavailable.

Core components:

- **JPEG Double Quantization Analysis:** Detection of recompression patterns using DCT coefficient distribution.
- **Enhanced Error Level Analysis (ELA):** Measurement of compression noise variance across image regions.
- **Deep Learning Splice Detection:** Pixel-level segmentation using a SegFormer-based architecture with frequency-aware feature extraction.

The output is a probability heatmap that visually highlights regions of potential manipulation.

4.3 Pipeline C - Digital Signature & Cryptographic Verification

This pipeline handles documents claiming legal validity through digital signatures (e.g., e-signed contracts, government certificates). It moves beyond visual inspection to mathematically prove authenticity.

Core components:

- **PAdES Standard Compliance:** Verifies adherence to PDF Advanced Electronic Signatures (PAdES) standards to ensure long-term validity.
- **Chain of Trust Validation:** Extracts X.509 certificates and traces the chain back to a trusted Root Certificate Authority (CA). This flags self-signed certificates or those issued by untrusted entities.
- **Integrity Checks:** Cryptographically confirms that the document hash has not changed since the signature was applied. Any byte-level alteration post-signing triggers an immediate invalidation.
- **Revocation Status (OCSP/CRL):** Real-time checks against Certificate Revocation Lists to ensure the signing certificate was valid at the time of signing.

4.4 Multi-Agent Forensic Reasoning Layer

If a document passes low-level structural, visual, and cryptographic checks, VeriDoc activates a semantic validation stage.

- **Audit Agent:** Extracts key entities, numerical values, and logical relationships from the document.
- **Verify Agent:** Cross-validates extracted data to detect inconsistencies, impossible values, or logical contradictions.

This agent-based loop improves precision and reduces false positives caused by OCR errors or ambiguous formatting.

5. Cloud Infrastructure (Google Cloud Platform)

VeriDoc leverages a fully managed, serverless cloud stack:

- **Frontend Hosting:** Firebase Hosting
- **Backend Orchestration:** Cloud Run (Gen 2) for containerized FastAPI services
- **AI Model Serving:** Vertex AI Prediction with GPU acceleration
- **Agent Reasoning:** Vertex AI Agent Builder
- **Key Management:** Cloud Key Management Service (KMS) for secure handling of verification keys.
- **Storage:** Cloud Storage with lifecycle and access control policies
- **Security:** IAM-based role separation and encrypted object storage

This architecture ensures horizontal scalability, fault tolerance, and compliance-ready data handling.

6. Expected Impact

VeriDoc reduces document verification time from manual review cycles to near real-time automated analysis. It enhances institutional trust by providing explainable, reproducible forensic results and **mathematical certainty regarding signatures**, instead of opaque AI predictions.

Key benefits include:

- Reduced fraud-related financial and operational risk
- **Non-repudiation** through cryptographic verification
- Increased auditability and compliance readiness
- Improved efficiency in high-volume verification workflows

The platform directly supports secure digital governance and enterprise digitization, contributing to the vision of a transparent and technology-driven Viksit Bharat.

7. Scalability & Future Roadmap

- Integration with institutional and government verification systems (e.g., DigiLocker).
- Multilingual OCR and regional document format support.
- Public and private API access for enterprise workflows.
- Blockchain-backed document hash anchoring for long-term integrity.
- Continuous learning pipeline for adaptive fraud pattern detection.

8. Conclusion

VeriDoc represents a shift from reactive document screening to proactive, forensic-grade digital trust infrastructure. By unifying deterministic file analysis, deep learning, **cryptographic authentication**, and cloud-native AI reasoning, it offers a scalable, transparent, and future-ready solution to one of the most critical challenges in the digital ecosystem: trust in documents.