SmartDoc: AI Based Mobile Application for Document Scanning, Verification and Categorization

Pranav Kishorsingh Rajput[1], Pranay Avinash Dahake[2], Anjali Mohanrao Kaware[3], Vivek Vinodrao Deshmukh[4], Sumedh Pundlikrao Ingale[5], and Roshan Rajendrakumar Karwa[6]

[1,2,3,4,5,6]*Department of Computer Science and Engineering, Prof. Ram Meghe Institute of Technology and Research Badnera- Amravati, Maharashtra, India*

Abstract –

Management of academic documents in higher education institutions often depends on manual verification processes that are time-consuming, difficult to scale, and susceptible to human error. This paper presents SmartDoc, a mobile-based system designed to automate document scanning, classification, and verification within academic environments. The proposed system integrates image preprocessing, Optical Character Recognition (OCR), and rule-based validation to ensure accurate and structured extraction of document content. The system integrates OCR with a hybrid validation approach combining rule-based checks and Gemini AI for intelligent document verification. The preprocessing module performs orientation correction, noise reduction, contrast enhancement, and user-assisted cropping to improve recognition reliability under practical capture conditions. Google ML Kit is used for on-device OCR processing, while extracted data is validated against predefined institutional rules to detect inconsistencies and incomplete information. The system also incorporates Gemini AI integration to support intelligent categorization, structured interpretation of extracted text, and automated verification feedback. SmartDoc is implemented using a cross-platform mobile framework with cloud-based backend services for secure storage, authentication, and role-based access control, enabling controlled document lifecycle management for students, faculty, and administrators. Experimental observations indicate improved consistency in document handling and reduced administrative workload. The proposed system provides a structured and secure approach to digital academic document management.

Key Words*: *Academic Document Management, Optical Character Recognition (OCR), Image Preprocessing, Rule - Based Validation, Role-Based Access Control (RBAC), Mobile Application Framework.

1.INTRODUCTION

To manage its extensive documentation, the modern academic environment increasingly depends on electronic technologies. As numerous studies [1], [2], [3] have shown, traditional manual processes for obtaining, verifying, and validating academic documents are not only time-consuming but also prone to human mistakes and non-scalable. Administrative difficulties, extensive delays, and a lack of transparency on document status for both staff and students are the outcomes of such inefficiencies.

Academic document management in higher education institutions is still a crucial administrative task that includes student and faculty for record verification, storage, retrieval, and limited access. Many schools still use manual or semi-automated verification procedures, which are complicated, hard to scale for big student populations, and prone to human mistakes, even in the face of continuous digitization attempts. Therefore, accurate text extraction, structured validation, and secure access management within an integrated system are necessary for the efficient digital processing of academic materials.

Textual information can be extracted from scanned documents and photographs using optical character recognition, or OCR. For printed materials, early methods concentrated on automatic text extraction using common OCR engines [1], [4]. Later studies investigated text extraction from complex structures and small text areas in papers and charts [5], as well as connected component-based techniques for text

detection in color images [6]. OCR performance differs greatly between printed and handwritten texts, according to comparative studies, particularly when scanning in noisy or low-quality environments [3]. Significant gains in recognition accuracy have been shown in more recent work that combines machine learning methods with OCR pipelines, especially for handwritten and rotating inputs [8], [9].

In secure situations, document authenticity and fraud detection have become more crucial than text extraction. In automated document validation tasks, multi-domain AI-based verification frameworks that include OCR, convolutional neural networks, and traditional classifiers like Random Forests and Support Vector Machines have demonstrated good classification accuracy [5], [25]. Preprocessing and structural pattern learning can greatly enhance the quality of corrupted scanned texts before recognition, according to studies on document cleaning using generative probabilistic models [7], [17].

A mobile-based framework that integrates preprocessing, OCR-based extraction, structured validation, categorization, and secure role-based access control is still required for academic institutions, even though these methods help with specific aspects of document processing like extraction, cleaning, or verification. To meet this demand, this study suggests SmartDoc, a mobile application that ensures controlled document lifecycle management in institutional settings while automating the scanning, classification, and verification of academic documents

2. RELATED WORK

Recent research in document processing has primarily focused on improving text extraction, document understanding, verification, and automated digital record management through the integration of Optical Character Recognition (OCR), machine learning, and deep learning techniques.

Early work by Kaundilya et al. [1] demonstrated the use of OCR systems for automated text extraction from scanned images, establishing the foundation for document digitization. Mahmood and Okumus [2] further explored electronic document management systems, emphasizing the importance of structured storage, retrieval, and access control in digital environments. Masa [3] extended this domain by integrating machine learning and natural language processing (NLP) for intelligent document processing automation, enabling improved semantic understanding of extracted content.

Busa et al. [4] investigated the challenges of extracting small text regions from documents and chart images, highlighting the limitations of traditional OCR systems when dealing with complex layouts. Kumar et al. [5] conducted a comparative study on OCR models for handwritten and printed documents, demonstrating that recognition accuracy is highly dependent on input quality, noise levels, and writing variability. These studies underline the necessity of robust preprocessing techniques before text recognition.

Hoh et al. [6] proposed a mobile-based OCR system that integrates machine learning to enhance recognition performance for rotated and handwritten text, showing the effectiveness of combining OCR with intelligent models in real-world scenarios. Jadhav et al. [7] introduced an AI-powered document verification framework that combines OCR, convolutional neural networks (CNNs), and classical classifiers such as Support Vector Machines and Random Forests to detect fraudulent documents with high accuracy.

Smith [8] provided a comprehensive overview of the Tesseract OCR engine, explaining how segmentation and preprocessing significantly affect recognition performance. Shafait et al. [9] emphasized the importance of document image binarization techniques, demonstrating that noise reduction and contrast enhancement are critical for improving OCR reliability in practical applications.

With advancements in deep learning, Shi et al. [10] proposed an end-to-end trainable neural network for image-based sequence recognition, combining convolutional neural networks (CNNs) for feature extraction and recurrent neural networks (RNNs) for sequence modeling. This approach significantly improved OCR accuracy by learning features directly from data without manual feature engineering.

Kim et al. [11] introduced an OCR-free document understanding transformer (Donut), which eliminates the need for explicit text extraction by directly generating structured outputs from document images. This approach represents a

paradigm shift from traditional OCR pipelines toward end-to-end document understanding.

Benchmark datasets such as MIDV-2019 [12] have played a crucial role in evaluating OCR performance in mobile-based document capture scenarios, highlighting challenges such as lighting variations, blur, and perspective distortion. Feng et al. [13] proposed DocTr, a transformer-based model for geometric correction of document images, which improves recognition accuracy by correcting distortions in captured documents.

Lerouge et al. [14] introduced the DocXPand dataset, enabling large-scale training and evaluation for identity document recognition systems. Similarly, Qiao et al. [15] developed DavarOCR, a comprehensive toolbox that integrates multimodal document understanding by combining visual, textual, and semantic features for improved extraction and interpretation.

Chowdhury et al. [16] proposed DeepReader, a system for extracting structured information from document images, demonstrating the importance of integrating layout understanding with text extraction. Baek et al. [17] introduced character region awareness techniques for accurate text detection, while Grüning et al. [18] proposed a two-stage method for text line detection, improving segmentation accuracy in complex documents.

Tensmeyer and Martinez [19] explored convolutional neural network (CNN)-based document image binarization, further enhancing preprocessing effectiveness. Zhang et al. [20] proposed TrOCR, a transformer-based OCR model that leverages attention mechanisms to achieve state-of-the-art performance in text recognition tasks.

In the context of document security, Pouliquen et al. [21] introduced weakly supervised learning techniques for detecting security features such as holograms in identity documents, improving fraud detection capabilities. Chen et al. [22] proposed hybrid document layout analysis techniques that enable accurate segmentation of document structures, such as tables and forms.

Bulatov et al. [23] extended benchmark datasets with MIDV-2020, providing improved evaluation frameworks for identity document analysis systems. Egwali et al. [24] developed an online academic document verification system, highlighting the need for automated and scalable verification frameworks in educational institutions. Saleh et al. [25] proposed a blockchain-based framework for secure educational certificate verification, ensuring immutability and trust in digital credentials. Finally, Grüning et al. [26] introduced the READ-BAD dataset for archival document analysis, supporting research in degraded and historical document processing.

Despite these advancements, most existing approaches focus on isolated aspects such as OCR, preprocessing, layout analysis, or verification. There remains a lack of integrated mobile-based systems that combine document scanning, intelligent extraction, structured validation, categorization, and secure role-based access control tailored for academic environments. The proposed SmartDoc system addresses this gap by providing a unified framework that integrates these components into a single mobile-cloud solution.

3. METHODOLOGY

The suggested SmartDoc system uses a structured mobile-based architecture to automate the processing and verification of academic documents. The methodology's main goal is to combine preprocessing, text extraction, document scanning, validation, and controlled access management into a single workflow. The system's modular architecture ensures dependability, transparency, and safe access to academic records by having each component contribute to the overall document lifecycle management.

A. System Architecture Overview

Scalable document processing is supported by SmartDoc's mobile-cloud architecture. To enable accessibility across multiple devices, a cross-platform framework is used in the development of the mobile application interface. Database management, secure storage, and authentication are all handled by cloud-based services. The architecture consists of three basic layers: the mobile application layer, the processing layer, and the cloud storage layer. The mobile application manages document capture, preprocessing, and user interaction, while the processing layer does text extraction, validation, and categorization. To guarantee safe data handling, the cloud

architecture supports role-based access management, document storage, and user authentication.

B. Image Acquisition and Preprocessing

The process of processing documents starts with taking pictures with the camera on a mobile device. Preprocessing is done before text extraction since real-world photos can include noise, skew, or uneven illumination. To enhance the clarity of textual content, the preprocessing stage carries out tasks like orientation correction, contrast adjustment, noise reduction, and region-based cropping. These procedures improve the acquired document quality and boost the dependability of the OCR processing that follows.

C. OCR-Based Text Extraction

Following preprocessing, the document image is transformed into machine-readable text using Optical Character Recognition (OCR). SmartDoc detects and extracts textual information from scanned academic publications using an on-device OCR engine. The OCR module creates digital representations of the retrieved content by processing printed and structured text fields. This stage makes it possible to handle document data automatically and serves as the foundation for additional validation and classification.

D. Document Validation Pipeline

The validation process is applied on the extracted text to establish whether there is any contextual and structure accuracy. In this case, the validation process will take into account both rule-based approaches in validating structured documents and Gemini AI in validating context-based documents developed in accordance with the organizational standards as opposed to keyword matching. The validation process will entail an analysis of the features of the document such as field requirements, formatting standards, and logical relationships in the extraction of the documents.

E. Gemini AI Integration and Verification

The system incorporates the Gemini API as its core intelligence engine to evaluate extracted text and verify documents using predetermined academic categories. This integration enhances document interpretation by understanding the context of the extracted data, rather than relying on simple keyword matching.

F. Secure Storage and Workflow Management

A cloud-based database environment safely stores all processed documents and related metadata. Before allowing users to access the system, authentication procedures are performed to confirm their identities. The workflow management component keeps track of document status, which enables users to keep an eye on how verification procedures are going. This organized procedure facilitates effective administrative operations and increases document processing transparency.

4. SYSTEM ARCHITECTURE OF SMARTDOC

The SmartDoc system's modular mobile-cloud architecture allows for safe storage, limited access to academic records, and effective document processing. To facilitate document scanning, text extraction, verification, and management on a single platform, the architecture combines a mobile application interface with cloud-based backend services. The mobile application layer, the processor layer, and the cloud service layer make up the three main parts of the overall system architecture.

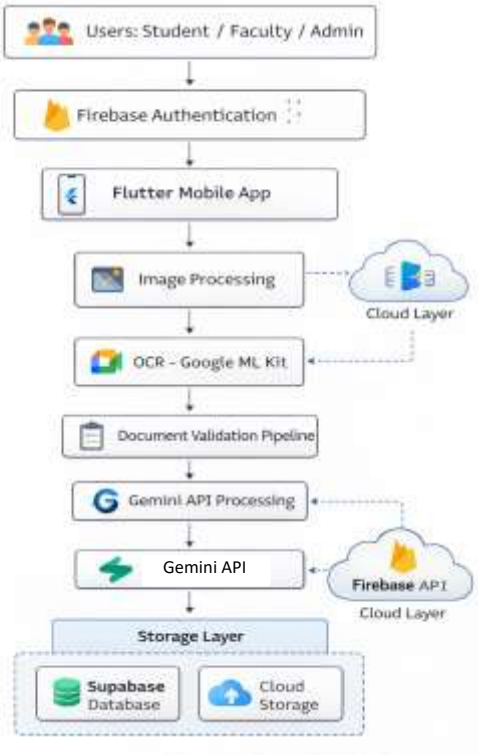

Figure - 1. SmartDoc system architecture

Moreover, the design of this application facilitates smooth communication and synchronization of information among the mobile and cloud elements by leveraging the use of APIs and

real-time information sharing protocols. The processor layer functions as an intelligent intermediary that manages processes like image processing, OCR, and validation using AI before transferring information to the cloud layer. The cloud services layer focuses on managing persistent data storage, access control based on roles, and scalable processing.

A. Mobile Application Layer

The main interface customers use to communicate with the SmartDoc system is the mobile application layer. A cross-platform mobile framework that enables deployment across many mobile operating systems was used in the development of the application. This layer oversees taking pictures of documents, carrying out preliminary preprocessing tasks, and enabling user interactions including document submission and status monitoring.

By using the camera on their mobile device or by choosing pre-existing files from local storage, users can upload academic documents. To enhance the quality of the input image before it is sent for additional processing, the program also supports image preprocessing techniques like cropping, orientation correction, and contrast modification. Additionally, the mobile interface offers various features based on user responsibilities, allowing administrators, teachers, and students to engage with the system in accordance with the rights that have been granted to them.

B. Document Processing Layer

The fundamental tasks needed for automated document analysis are carried out by the document processing layer. Preprocessing, optical character recognition (OCR), validation, and document categorization modules are all included in this component. By lowering noise and adjusting orientation to increase text visibility, the preprocessing module enhances image quality.

The OCR module extracts text from the document image and transforms it into a machine-readable format after preprocessing is finished. After that, the extracted material is sent to the validation module, where rule-based checks are carried out to guarantee the accuracy and consistency of document data. These checks confirm that the document structure complies with institutional rules and that the necessary fields are present.

A document categorization component, which groups documents into predetermined categories such certificates, transcripts, and identifying records, is another component of the processing layer. The system's effective indexing and retrieval of documents is made possible by this classification.

C. Cloud Service Layer

The cloud service layer oversees system scalability, authentication, and safe storage. Database management, document storage, and user authentication are all provided by backend services. Every uploaded document and the metadata that goes with it - such as user details, document type, and verification status - are safely preserved.

To govern user permissions, a role-based access control system is put in place at the cloud level. Administrators oversee system-level functions including user administration and database maintenance; faculty members can examine and validate supplied records, and students are allowed to upload and track their documents. Only authorized individuals will be able to access sensitive academic data thanks to this controlled access approach.

The cloud infrastructure also enables real-time synchronization between the mobile application and the backend database, allowing users to monitor document verification status and receive updates regarding document processing. This integration improves transparency and reduces delays associated with traditional manual verification procedures.

5.IMPLEMENTATION DETAILS

The technology stack, data management infrastructure, and functional processes created for various user roles are all covered in this section's practical implementation of the SmartDoc system. The system is built using a mobile-cloud architecture, where backend services handle authentication, storage, and document data management while the mobile application handles user interaction and document collection. A systematic procedure for document submission, verification, and administrative management is the main goal of implementation.

A. Technology Stack and Data Management:

A cutting-edge cross-platform and cloud-based technological stack that facilitates scalability, safe authentication, and effective data management is used to develop the SmartDoc application.

1) Front-end Layer:

The mobile application was developed using Flutter, a cross-platform framework that enables the creation of responsive user interfaces across multiple operating systems. Flutter's unified interface provides a consistent and accessible experience for students, faculty, and administrators. To maintain a reactive and synchronized user experience, the application implements the provider's package for state management. At the root level, a UserProvider manages the global state, ensuring smooth interface updates and seamlessly handling dynamic role transitions between Students, Faculty, and Administrators based on their active authentication status.

2) Backend Services:

Firebase supports the backend infrastructure by offering integrated services for cloud storage, database administration, and authentication.

a) *Firebase Authentication:*

This module controls user authentication and facilitates the system's Role-Based Access Control (RBAC) structure. Access to permissions is issued in accordance with predetermined authorization rules, and each user is given a certain role (administrator, faculty, or student).

b) *Supabase Storage:*

Supabase is used for file storage and integration support, while Firebase Firestore serves as the primary database for storing structured document metadata such as user profiles, categories, and verification status.

c) *Firebase Firestore(Database):*

Firebase Firestore serves as the primary NoSQL database for storing structured document metadata, including user profiles, document categories, and verification status. It provides real-time data synchronization, ensuring that updates are reflected instantly across all user roles.

d) *Firebase Cloud Messaging (FCM):*

For the timely communication aspect, the system uses Firebase Cloud Messaging, which enables the sending of push notifications to the students asynchronously, allowing the students to receive instant notifications when the verification status of the documents they submitted has been updated, or when certain feedback has been broadcasted from the side of the faculty members.

The mobile interface can communicate effectively with the backend infrastructure using Flutter technology, which ensures the safe handling of the documents submitted by the students.

e) *System Security and Data Integrity:*

The system uses Firebase App Check and ReCaptchaV3 to make sure that the backend infrastructure is safe from any kind of unauthorized access and API abuse. This makes sure that every network request is checked to make sure it comes from the real SmartDoc app. This stops any kind of spoofing and abuse of the backend infrastructure. To make sure that privacy is always protected, the system also uses granular-level Firestore security rules that are enforced directly at the database level. The rules are set up so that students can only read and write their own submitted documents. Faculty and administration, on the other hand, have the higher permissions they need to check documents submitted by more than one user.

B. Student Document Submission Workflow:

The student interface is made to ensure accurate data capture while simplifying the document submission procedure. Following authentication, students can upload academic documents for verification on the Student Dashboard.

1) *Image Acquisition:*

Students can upload documents by choosing an existing image from the device gallery or by taking a new one with the camera on their mobile device.

2) *Image Processing:*

The google_mlkit_document_scanner API is built into the system to make sure that it captures high-quality data from physical papers. This specialized machine learning module automatically finds the edges of documents, corrects perspective, and lets the user approve or change the cropping

area exactly while taking the picture. These automated preprocessing and enhancement methods make text much easier to see and make sure that the document is set up perfectly for the next step, which is to extract text from it using OCR.

3) *Metadata Selection:*

The document category (such as marksheet, certificate, or identity document) is chosen by the student. To help with document structure, this data is saved with the uploaded image.

4) *Document Upload:*

The processed document image is uploaded to the cloud storage system, while the associated metadata is stored in the Firestore database.

5) *Status Update:*

After a successful upload, the AI verifies the document and marks the status as *Rejected* or *Approved* and if not possible to verify then system assigns an initial status of *Pending* to the document. The student dashboard displays the uploaded document along with its verification status, allowing the student to monitor the review process.

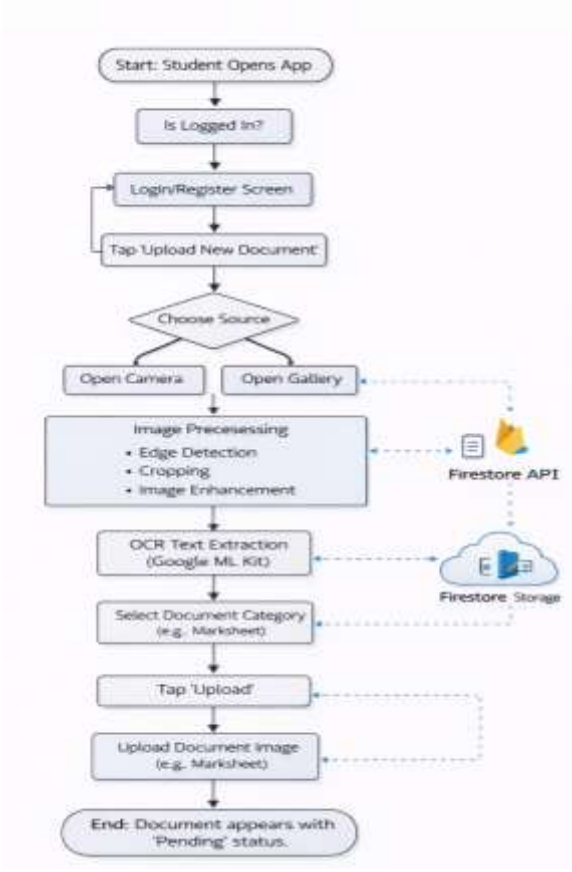

Figure - 2. Student Document Submission Workflow

C. Process for Faculty Review and Notification:

Faculty members are responsible for reviewing submitted documents and verifying their authenticity. The review process is supported by centralized document data stored in the Firestore database.

1) Review Dashboard:

After logging into the system, faculty members access the Faculty Review Dashboard, which displays a list of documents awaiting verification.

2) Document Inspection:

Faculty members select a document from the list to review the submitted image and associated student information.

3) Approval Process:

Upon determining that a document meets all institutional criteria, the faculty member selects 'Approve'. The Firestore database updates the document status to 'Approved', triggering a real-time notification to the student's dashboard.

4) Rejection Process:

If a document is non-compliant, the faculty member selects 'Reject' and provides a specific justification. This feedback, which can be supplemented by Gemini AI-generated analysis regarding missing fields or formatting errors, is sent instantly to the student to facilitate a corrected resubmission.

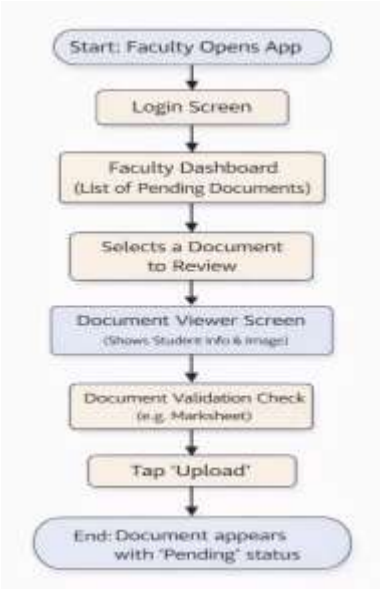

Figure - 3. Faculty Review and Notification Workflow

D. Admin Management Functionality:

The administrator interface provides system-level management capabilities, enabling monitoring and configuration of users and document categories.

1) User Management:

By generating new user accounts in Firebase Authentication and allocating the proper roles, administrators can add or remove system users. Administrators can manage the user base and restrict access to system functionalities according to this capability.

2) Category Management:

Document categories maintained in the Firestore database can be added, changed, or removed by the administrator using a specific interface.

3) Analytics and Monitoring:

An overview of system activity, including data on pending, approved, and rejected documents, is given via the administrative dashboard. Administrators can keep an eye on system performance and document processing progress with the use of these analytics.

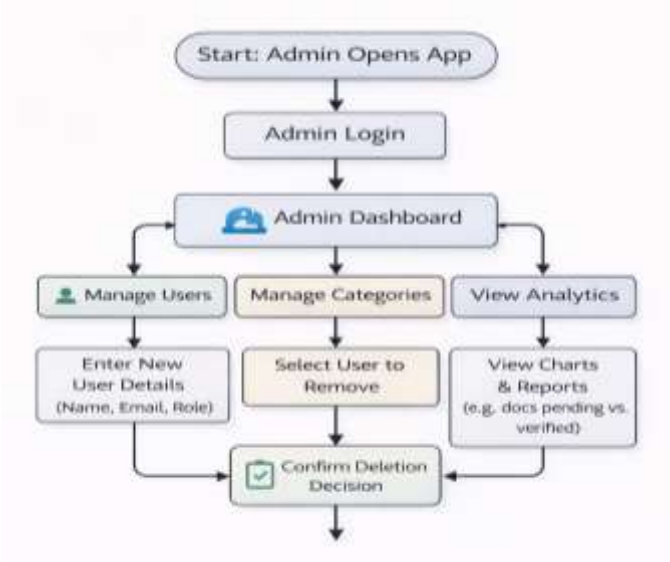

Figure - 4. Admin Login and Workflow Management

6. RESULTS AND ANALYSIS

Based on the installed application modules and user interaction procedures, this section assesses the SmartDoc system's operational efficiency and functional performance. Validating the role-based architecture, the document management process's usability, and the system's capacity to facilitate open communication between administrators, instructors, and students are the main objectives of the assessment. The system's capacity to effectively handle academic materials while upholding secure access control is demonstrated by the established application interfaces and workflow results.

A. Functional Verification of Multi-Role Architecture

The SmartDoc system implements a Role-Based Access Control (RBAC) framework that ensures secure and structured access to application functionalities for three user roles: student, faculty, and administrator.

1) Secure Role-Specific Authentication

Before they login, users choose their responsibilities via the application role selection interface. Administrators, instructors, and students each have their own login interfaces. Firebase Authentication manages user authentication, guaranteeing role-based access control and safe credential validation.

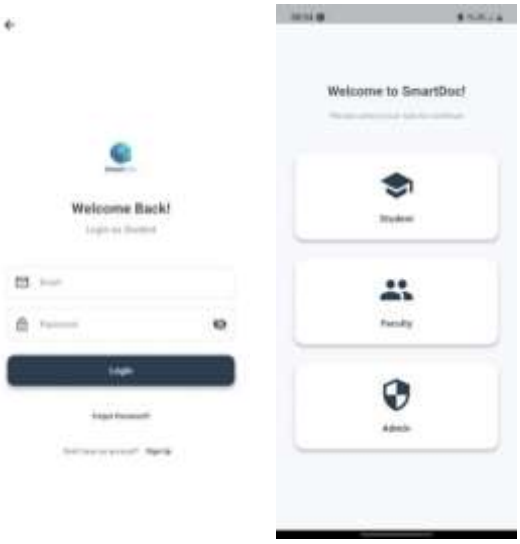

Figure - 5. Role selection and login interfaces, and Administrator users.

2) User Management and Data Privacy

To safeguard sensitive data, the system keeps user privileges clearly separated. The student dashboard allows students to access and control their uploaded documents and profile information. Faculty members are prevented from accessing unrelated user data by limiting their access to the list of student documents designated for verification.

Complete management over user accounts, including the ability to examine enrolled students and faculty members, manage user responsibilities, and delete inactive accounts, is made possible by the administrator interface. This centralized management guarantees controlled access and appropriate maintenance of the system's user database.

Figure - 6. Administrative interface for managing student and faculty user accounts.

3) Administrative Monitoring

The administrator dashboard uses real-time analytics to summarize system activity. Important data is shown in a visual fashion, including the overall number of faculty members, registered students, and document verification status. Administrators can track the status of document processing throughout the organization with the aid of a graphical depiction of the document status categories (Approved, Pending, and Rejected).

Figure - 7. Administrative dashboard displaying system statics and document verification status.

A. Workflow Validation for Efficiency and Real-Time Feedback

The SmartDoc system was also evaluated based on its ability to streamline document submission, enable transparent verification workflows, and provide timely feedback to users

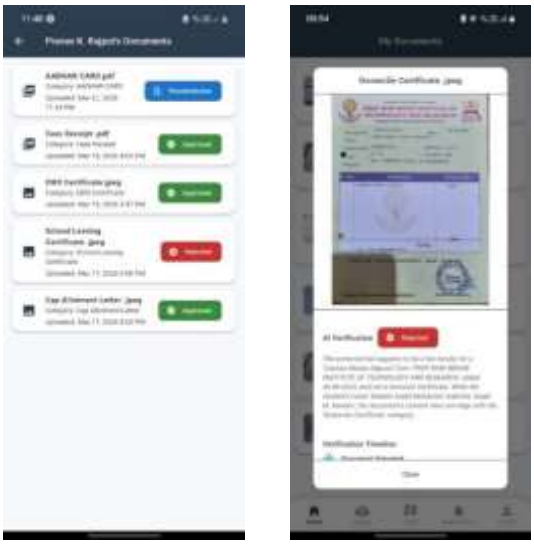

Figure - 8. Student dashboard displaying uploaded documents and verification status tracking.

4) Streamlined Document Submission

Students can use the mobile application to instantly upload academic documents. Both manual file selection and document scanning using the mobile camera are supported by the document upload interface. Before uploading files in JPG, PNG, or PDF formats, the application lets users choose document categories. Both academics and students can handle

documents more easily thanks to this standardized submission method.

Figure - 9. Student interface for document scanning and uploading through the mobile application.

5) Transparent Verification and Notification Mechanism

The system incorporates a notification mechanism that notifies students of document verification results in order to guarantee transparency in the verification process. Faculty members have the option to accept or reject uploaded documents after reviewing them. Faculty members give a detailed explanation if they are rejected. Through the application's notification interface, the student receives the notification instantly.

This function enables students to promptly address problems with document submission and guarantees clear communication between students and faculty members.

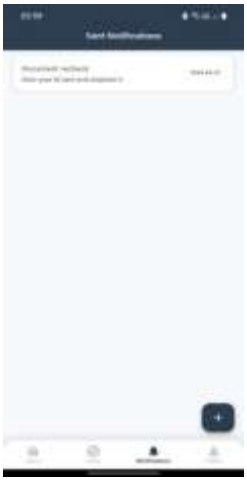

Figure - 10. Faculty interface of notification to send students notices about their documents

6) Processing Time Reduction

The SmartDoc system greatly reduces the administrative burden connected with manual document verification by automating document submission, verification, and status tracking. Without having to go to administrative offices, students may keep an eye on the status of their documents right from their dashboards. A centralized verification portal allows faculty members to examine and verify documents.

The automated workflow minimizes delays in document processing and improves the efficiency of academic record management.

A. Implementation Validation and System Scalability

The system implementation confirms the technical feasibility of deploying SmartDoc in academic institutions.

1) Technology Implementation

Firebase services manage cloud storage, database administration, and authentication, whereas Flutter is used for cross-platform mobile development. The administrative interface allows for the organized classification of submitted documents by managing document types including SSC Marksheet, Aadhaar Card, Bank Passbook, Income Certificate, College ID Card, etc.

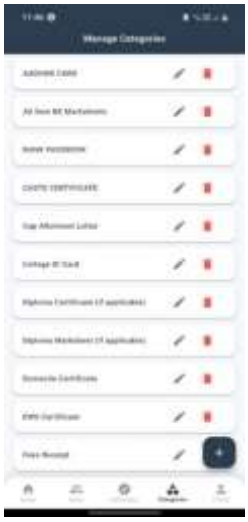

Figure - 11. Administrative interface for managing academic document categories

2) System Scalabilty

Large numbers of students and papers can be accommodated by the SmartDoc system thanks to its cloud-based architecture. While cloud storage guarantees safe and scalable document storage, Firebase Firestore allows users to synchronize document status updates in real-time.

Based on the implemented workflow automation, the system has the potential to significantly reduce document processing time compared to manual verification methods typically used in academic institutions.

B. Performance Metrics and Comparative Analysis

The SmartDoc system was evaluated based on the accuracy of its Gemini-powered categorization and the speed of its OCR-extraction pipeline.

1) Categorization Accuracy

During initial prototype testing on a limited dataset of academic documents (marksheets, ID cards, and certificates), the system demonstrated high accuracy in document categorization using the Gemini API. The results indicate that the model is effective in identifying document types under standard conditions.

2) Processing Latency

The system demonstrates fast processing performance from image capture to OCR-based text extraction and validation, enabling near real-time document verification.

3) Administrative Efficiency

The integration of automated validation and structured feedback reduces the manual effort required for document verification. Faculty members can review and process documents more efficiently through a centralized dashboard, leading to improved workflow management.

7. CONCLUSION AND FUTURE SCOPE

SmartDoc offers a simple and effective way to manage and verify academic documents using a mobile app. Built with Flutter and supported by Firebase, it allows students to upload and track their documents, faculty to review and verify them, and administrators to manage users and system activities in an organized way. By using role-based access, the system keeps everything structured and easy to follow. The implementation shows that it reduces manual effort, improves transparency, and keeps everyone updated in real time. Its cloud-based setup also makes it reliable for institutions with a large number of students. In the future, SmartDoc can be improved by adding smarter document checks using machine learning, exploring blockchain-based decentralized identifiers (DIDs) for immutable credential storage and integrating Multi-Modal Gemini models to support video-based document liveness detection, connecting with existing ERP systems, and enhancing features like OCR, document classification, and performance tracking to make the system even more useful and efficient.

REFERENCES

1] C. Kaundilya, D. Chawla, Y. Chopra, "Automated Text Extraction from Images Using OCR System," IEEE INDIACom, 2019.

[2] Mahmood, A., Okumus, I. T., "Electronic Document Management System," 2017.

[3] Masa, B. A., "AI-Based Document Processing Automation using ML & NLP," University of Bologna, 2023.

[4] R. Busa, S. K. C., Lijiya, "Small Text Extraction from Documents and Chart Images," IEEE INDICON, 2022.

[5] A. Kumar, P. Singh, K. Lata, "Comparative Study of OCR Models on Handwritten and Printed Medical Reports," IEEE, 2023.

[6] W. S. Hoh et al., "ScanNote: A Mobile Application for Enhanced Text Recognition Using Machine Learning-Driven OCR," Journal of Advanced Research Design, 2025.

[7] A. Jadhav et al., "AI Powered Document Verification: A Behavioral and Multi-Domain Approach," IEEE ICTBIG, 2025.

[8] R. Smith, "An Overview of the Tesseract OCR Engine," ICDAR, IEEE.

[9] F. Shafait, D. Keysers, T. Breuel, "Efficient Implementation of Document Image Binarization Techniques," ICDAR.

[10] B. Shi, X. Bai, C. Yao, "An End-to-End Trainable Neural Network for Image-Based Sequence Recognition," IEEE TPAMI, 2017.

[11] G. Kim et al., "OCR-Free Document Understanding Transformer (Donut)," 2021.

[12] K. Bulatov et al., "MIDV-2019: Challenges of Mobile-Based Document OCR," 2019.

[13] H. Feng et al., "DocTr: Document Image Transformer for Geometric Correction," 2021.

[14] J. Lerouge et al., "DocXPand-25k Dataset for Identity Document Recognition," 2024.

[15] L. Qiao et al., "DavarOCR: A Toolbox for OCR and Multimodal Document Understanding," 2022.

[16] S. Chowdhury et al., "DeepReader: Information Extraction from Document Images," 2018.

[17] J. Baek et al., "Character Region Awareness for Text Detection," 2019.

[18] T. Grüning et al., "A Two-Stage Method for Text Line Detection," 2018.

[19] C. Tensmeyer, T. Martinez, "Document Image Binarization using CNN," 2017.

[20] H. Zhang et al., "TrOCR: Transformer-based Optical Character Recognition," 2021.

[21] G. Pouliquen et al., "Weakly Supervised Training for Hologram Verification in Identity Documents," 2024.

[22] Y. Chen et al., "Hybrid Document Layout Analysis in Document Images," 2024.

[23] K. Bulatov et al., "MIDV-2020: Benchmark Dataset for Identity Document Analysis," 2021.

[24] Egwali, A. O. et al., "Online Academic Document Verification System," 2025.

[25] Saleh, O. S. et al., "Blockchain-Based Framework for Educational Certificate Verification," 2020.

[26] T. Grüning et al., "READ-BAD Dataset for Archival Document Analysis."