Eyes of Authenticity: Document Forgery Detection using Siamese and OCR Techniques

Samiksha Chavan, Rajashree Kute, Prathamesh Rathi, Vaishnavi Narde, Prof.Rupal More
Artificial Intelligence and Data Science, K. K. Wagh Institute of Engineering Education and Research Nashik, Maharashtra, India.

Abstract- In an era marked by the widespread adoption of numerous digital tools and techniques, combating digital crimes like the fraudulent replication or forgery of official documents has emerged as a formidable task. This has become feasible due to technological progress, enabling easy alteration of documents. Every aspect, such as official seals, signatures, and text, can be tampered with. Consequently, both public and private sectors are experiencing significant losses in time and resources. In such a situation, it becomes crucial to employ different methodologies to identify and examine irregularities within the inherent characteristics of the document image. By using, different image processing approaches designed to identify instances of document forgery, with the goal of curbing these unlawful practices. The primary objective is forgery detection to optimize the retrieval of information from manipulated photographs and documents, especially in cases involving noisy and post-processedimage documents.

Keywords: Digital Crimes, Document Forgery, Official Documents, Technological Progress, Image Processing, Forgery Detection, Information Retrieval, Fraudulent replication, Tampering.

I. INTRODUCTION

In today's digital age, verifying the legitimacy of documents has become crucial due to the prevalence of online transactions and virtual engagements. The increasing reliance on digital services and electronic documents has introduced fresh challenges, especially in the context of document falsification. Deceptive practices like counterfeit signatures and incorrect portrayal of document types of present substantial risks to the credibility of diverse sectors, including finance, healthcare, and government institutions.[1]

In educational institutions, the transition to the digital era has indeed brought about numerous benefits, particularly in terms of streamlining administrative procedures through document uploads. However, a notable challenge that has emerged is the inadvertent or intentional uploading of incorrect documents, such as using a sibling's Aadhaar or PAN card instead of one's own. This practice can have significant consequences, compromising the accuracy of student records and introducing risks related to identity verification, scholarship eligibility, and official documentation[1].Comprehending the motivations and situations that lead to the submission of inaccurate documents is essential for devising effective preventive measures. Errors in document uploads can arise from sources such as confusion, a lack of awareness, or even technological constraints. Aadhaar, a distinct 12-digit identification number provided by the Government of India, is closely tied to an individual's biometric and demographic data. In contrast, PAN plays a pivotal role as an identifier in financial transactions and matters related to taxation. Implementing robust verification protocols and educating stakeholders about the importance ofaccurate document submission are critical steps towards safeguarding the integrity of digital records and mitigating the risks associated with fraudulent activities.[11]

II.LITERATURE SURVEY

Yusuke Kataoka, Takashi Matsubara, Kuniaki Uehara, "A CNN-based Architecture for Forgery Detectionin Administrative Documents" at IEEE

This research paper presents a Convolutional Neural Network (CNN)-based architecture tailored for the detection of forgeries in administrative documents, with a focus on enhancing accuracy and efficiency in identifying fraudulent alterations. Algorithms used include KGG, SVM. Pros noticed include robustness to variations, scalability, high potential for automated document screening. Cons include dependency on availability of large and diverse labeleddatasets, high computational resources required for training and testing.

Misbah Shaikh, Dr. Dipak Patil, “Image Forgery/Tampering Detection Using Deep Learning and Cloud” at IEEE
This research paper investigates image forgery and tampering detection using deep learning and cloud-based solutions, focusing on the scalability and real-time capabilities offered by cloud computing. Algorithms used include Azure Form Recognizer, CNN, Deep Learning, etc. Pros noticed include enhanced scalability, enabling analysis of large volumes of images, and real-time processing capabilities. Cons include inaccuracy for on-premise services, low latency or data transfer issues, and significant expenses when handling particularly large datasets.

Saleem Summra, GhaniUsmanM, AslamMuhammad, Martinez, “Supervised Neural Network for Offline Forgery Detection of Handwritten Signature” at IEEE
This research paper presents a supervised neural network approach for detecting forged handwritten signatures inoffline documents, leveraging deep learning techniques to analyze signature features and patterns. Algorithms used include Five and Three Layer CNN. Pros found include good accuracy, signature style variability, and multimodal efficiency. Cons found include the requirement for a large amount of labelled data, generalization challenges, andoverfitting.

Maamouli Khadidja, Benhamza Hiba, Djeffal Abdelhamid, Abbas Cheddad, “Mitigating Digital SignatureForgery Using Machine Learning” at IEEE
This research paper investigates image forgery and tampering detection using deep learning and cloud-based solutions, focusing on the scalability and real-time capabilities offered by cloud computing. Algorithm used wasCycleGAN. Pros found include the utilization of advanced technology (machine learning) for detection and the potential to reduce the risk of signature forgery. Cons include less accuracy, complexity in implementation.

III: PROBLEM FORMULATION AND PROPOSED METHOD

To validate document authenticity, our system verifies separate signature photos for forgery and detects incorrect document types when uploaded. Eg. PAN card instead of Aadhar Card or vice versa.
In this section, we discuss the problem formulation and present the proposed method or system, focusing on the utilization of Siamese Networks for Signature Verification and OCR Techniques for Document Classification. Additionally, we elaborate on the integration of these components in the forgery detection process.

A. Methodology

Siamese Network for Signature Verification
Finding instances of signature forgeries is an essential task in a number of domains, such as document verification,legal, and financial. Because handwriting is so intricate and varied, it might be difficult for traditional methods to discern between authentic and fake signatures. Neural network architectures called Siamese Networks have shown promise in the identification of signature forgeries. The purpose of Siamese Networks is to detect and quantify the similarity between input pair pairs. The network is taught to recognise the minute distinctions between authentic and falsified signatures in the context of signature forgery detection. Siamese networks function on the basic principle of mapping two input signatures into a shared feature space and calculating the degree of similarity or dissimilarity between these representations. [4]
Each subnetwork within the Siamese network conducts feature extraction from input signature images, utilizing convolutional layers to capture hierarchical and spatial features. Simultaneously, pooling layers may be incorporatedto reduce dimensionality while retaining crucial features.
The Siamese network calculates a similarity metric between the feature representations of two given input signatures, employing common metrics like Euclidean distance or cosine similarity. The primary objective isto minimize the distance between genuine signatures and maximize the distance between genuine and forged signatures. During the training phase, pairs of genuine signatures and genuine-forged signature pairs are utilized.The network is trained to position genuine signatures closely in the feature space and create a distinct separation for forged signatures. Training batches are curated with a combination of genuine and forged pairs, ensuring the network generalizes effectively across diverse signature characteristics. The contrastive loss function is

instrumental in training, penalizing similarity between pairs and encouraging dissimilarity. This function incorporates a marginparameter to enforce a minimum separation between feature representations. In the testing phase, a threshold for the similarity score is established, either based on a validation set or domain knowledge. Signatures with similarity scores below the threshold are categorized as genuine, while those above are deemed forged. Siamese networks exhibit adaptability for continuous learning, allowing the model to be updated with new genuine and forged signatures over time. This adaptation ensures the model remains robust in dynamic environments.[4]

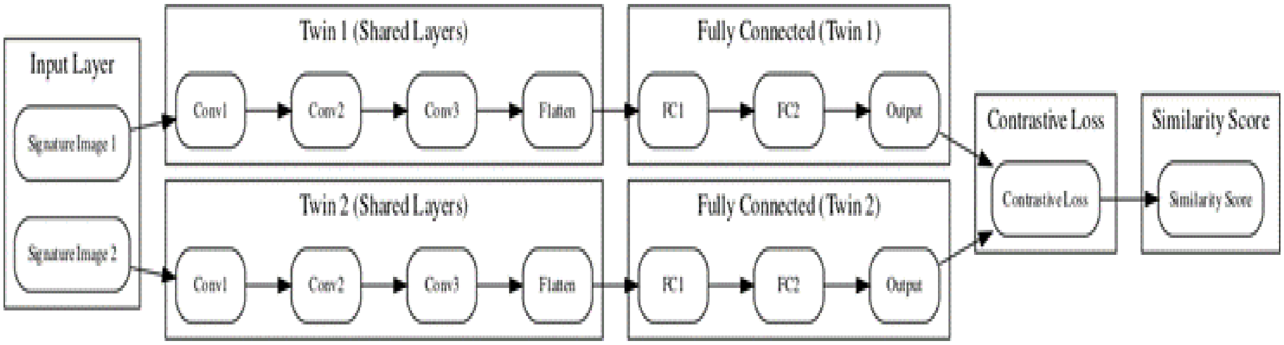


Figure 1: Siamese Working

OCR Technique for Document Classification

OCR techniques for document classification typically begin by identifying relevant regions within the document, focusing on areas likely to contain important information, such as names, dates of birth, and identification numbers. The process involves character recognition, where individual characters within the identified regions are recognized and interpreted. This fundamental OCR technique enables accurate character recognition even in the presence of noise or variations. Going beyond mere character recognition, OCR systems also aim to understand the context of the extracted text, interpreting relationships between different pieces of information for coherent and meaningful classification. In the case of specific document types like Aadhar and PAN cards, OCR systems may employ template matching techniques. Template matching involves comparing the extracted text with predefined templates for these cards, facilitating accurate classification based on predefined document structures. To further enhance accuracy and validation, OCR results can be cross-verified with information entered by the user during document upload. This additional step ensures consistency between the extracted document details

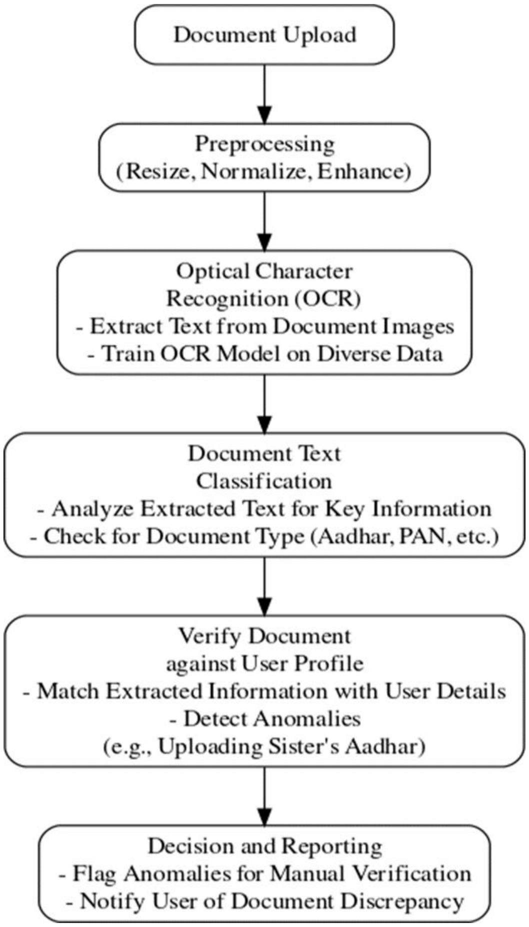

Document Upload
Preprocessing
(Resize, Normalize, Enhance)
Optical Character
Recognition (OCR)
- Extract Text from Document Images
- Train OCR Model on Diverse Data
Document Text
Classification
- Analyze Extracted Text for Key Information
- Check for Document Type (Aadhar, PAN, etc.)
Verify Document
against User Profile
- Match Extracted Information with User Details
- Detect Anomalies
(e.g., Uploading Sister's Aadhar)
Decision and Reporting
- Flag Anomalies for Manual Verification
- Notify User of Document Discrepancy

Figure 2: OCR Working

and user- provided information, adding an extra layer of verification to the classification process.[11]

Integration of Siamese Network and OCR in Forgery Detection

After the Siamese network completes signature verification, the OCR components are seamlessly incorporated into the forgery detection system. This integration is carefully designed to leverage the strengths of both systems, creating a unified and synergistic operation. The Siamese network excels in assessing visual similarities in signatures, while OCR specializes in extracting and comprehending textual content. By collaborating, the systemcan harness the complementary strengths of each component, resulting in a more comprehensive forgery detectioncapability. OCR plays a vital role in validating the accuracy of the text extracted from the document. For example, it verifies that the extracted name corresponds with the name entered by the user during document upload, ensuring precise identification of individuals. Utilizing OCR, the system verifies specific document details such as identification numbers, addresses, or dates. This verification process adds an extra layer of scrutiny, confirmingthat the content extracted from the document aligns with the expected and user-provided information. The fusionenables the forgery detection system to adapt to a wide range of forgery techniques. While the Siamese network excels in recognizing visual inconsistencies in signatures, OCR contributes by revealing discrepancies in textual content, making the system resilient against various forgery attempts.

B. Proposed System

The methodology of this project encompasses two primary components: Signature Verification and Document Type Identification and Classification, each addressing distinct challenges in document authenticity assessment. The review paper identifies the escalating complexities associated with detecting fraudulent documents and proposes a comprehensive approach for detecting signature forgery. Integrating multiple modalities such as text recognitionand image analysis, the suggested methodology aims to enhance the precision and reliability of the detection process.

In addition to the technical aspects, the project emphasizes the importance of an intuitive user interface to streamline the document verification process for users. By providing a user-friendly design with clear instructions, feedback mechanisms, and intuitive navigation, users can easily upload documents and signatures. Concise and unambiguous feedback on the verification outcomes, including document type and signature authenticity, is crucial for enhancing user confidence. To further improve user comprehension, the interface incorporates status indicators to provide visual cues throughout the verification process.

Moreover, ensuring the effectiveness of the document forgery detection system requires a rigorous validation process encompassing feature extraction, machine learning model training, and comprehensive data gathering. The methodology prioritizes these validation steps to validate the accuracy and robustness of the system. Addi- tionally, ongoing refinement and optimization of the system components are integral to continuously improving itsperformance and adaptability to evolving fraudulent techniques.

Expanding on the two main components of the approach, the project delves deeper into each stage, including advanced feature extraction techniques tailored to signature analysis and document layout, as well as the development of sophisticated machine learning models capable of detecting subtle patterns indicative of forgery.Furthermore, the integration of cutting-edge technologies such as optical character recognition (OCR) and deep learning algorithms enriches the system's capabilities, enabling it to handle a diverse range of document types andsignature styles with high accuracy and efficiency.[9]

Signature Verification

To achieve a thorough depiction of signature styles and variants, a broad dataset of real and fraudulent signatures was painstakingly created throughout the data collecting phase. Sincere signatures from people in a range of demographics were collected, capturing differences in writing velocities, styles, and angles. To test the system's resilience, a variety of forgery methods were used on faked signatures, including freehand, traced, and simulated forgeries. Next, different techniques were investigated to objectively quantify signatures through the extraction ofsignature features. This required examining relevant properties, curves, and stroke patterns. The need for carefully choosing characteristics that capture the distinctive and differentiating qualities of each trademark was underlined. Using cutting-edge image processing algorithms, both static and dynamic elements were extracted, guaranteeing a thorough depiction of the inherent qualities of the signatures.[6]

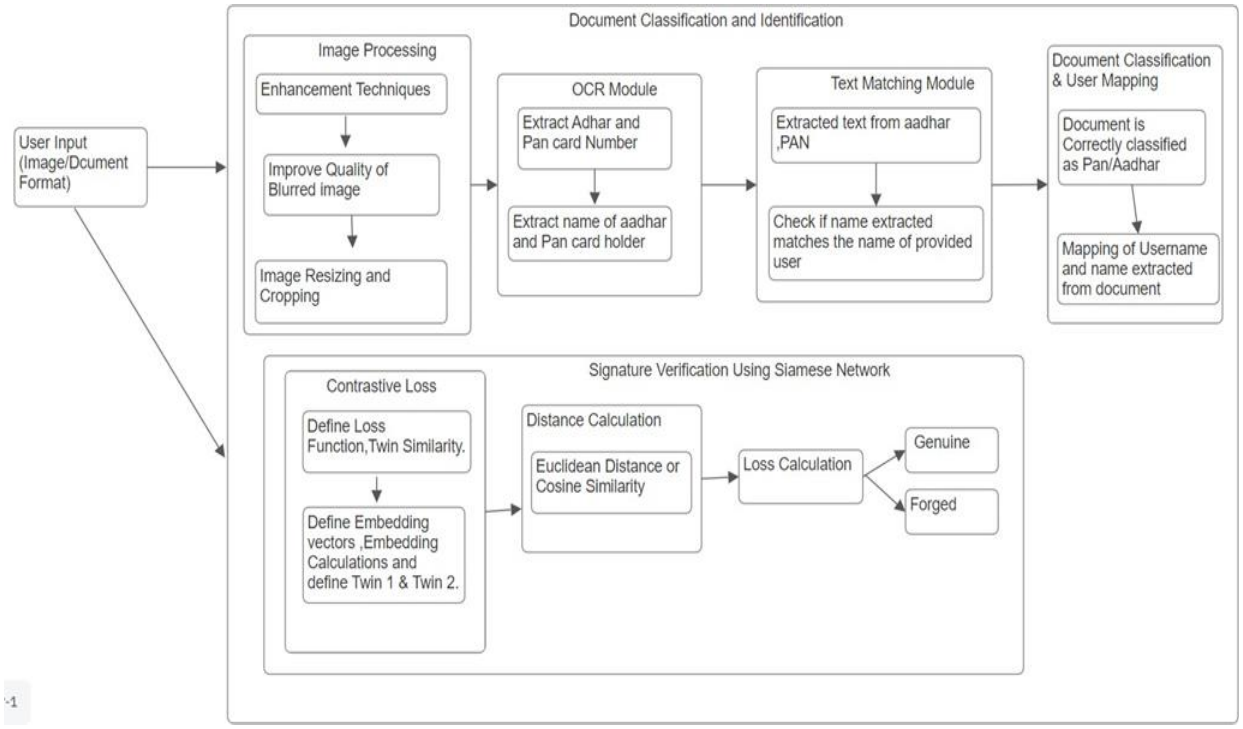

Document Classification and Identification
User Input (Image/Dcument Format)
Image Processing
Enhancement Techniques
Improve Quality of Blurred image
Image Resizing and Cropping
OCR Module
Extract Adhar and Pan card Number
Extract name of aadhar and Pan card holder
Text Matching Module
Extracted text from aadhar ,PAN
Check if name extracted matches the name of provided user
Dcoument Classification & User Mapping
Document is Correctly classified as Pan/Aadhar
Mapping of Username and name extracted from document
Signature Verification Using Siamese Network
Contrastive Loss
Define Loss Function,Twin Similarity.
Define Embedding vectors ,Embedding Calculations and define Twin 1 & Twin 2.
Distance Calculation
Euclidean Distance or Cosine Similarity
Loss Calculation
Genuine
Forged
-1

Figure 3: Proposed System

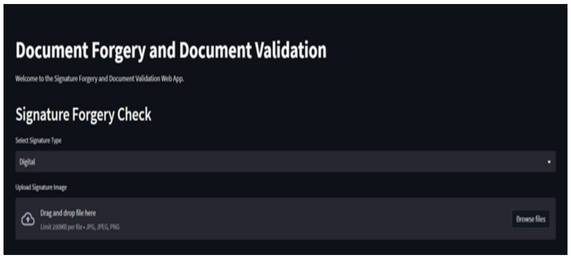
Siamese Network: The offline signature verification system leverages a Siamese network architecture to enhance its resilience and accuracy. Siamese networks, characterized by two identical sub-networks sharing parameters and architecture, specialize in tasks requiring similarity comparison, such as signature verification. During training, this network learns to discern meaningful features from pairs of input samples, calculating their similarity or dissimilarity. By minimizing the distance between similar pairs and maximizing the distance between dissimilar ones within a learned feature space, the Siamese network adeptly distinguishes between genuine and fraudulent signatures. Notably, this architecture excels in scenarios with limited labeled data, effectively generalizing from small training sets. Moreover, the Siamese network's capability to capture both static and dynamic signature elements facilitates comprehensive analysis and comparison of signature styles, bolstering the system's resilience against fraudulent attempts. This integration of Siamese network design significantly strengthens the offline signature verification system's capacity to identify subtle differences and similarities, particularly valuable in situations with minimal labeled data, thereby advancing its effectiveness in ensuring document authenticity.[5]

Document Classification

This study proposes a complete approach for an integrated multi-modal biometric verification system that combines document type identification with signature verification. To determine the document type (such as a PAN card orAadhar card), text-based categorization, document layout analysis, and optical character recognition (OCR) are used in conjunction with text extraction. To reliably identify document kinds, a specialized document matching algorithm is designed that considers many factors such as document format, text patterns, and keywords. Carefully chosen datasets are used to train machine learning models for both document type detection and signature verification. The system architecture is built for smooth integration, enabling the components for document type identification and signature verification to function together. To make user interactions easier, an intuitive interface is designed,assisting users in navigating the signing and uploading documents, all the while giving concise feedback on the verification outcomes. To ensure that the system is accurate and reliable, extensive testing is carried out on a wide range of datasets that include counterfeit and authentic documents and signatures. With the help of this technique,a reliable and user-friendly biometric verification system that can simultaneously validate signatures and check document types for increased security applications will be developed. Expanding upon the system's capabilities, OCR technology is employed to extract PAN and Aadhar numbers from the uploaded documents, thereby enabling more granular classification based on extracted information. Additionally, users are afforded the convenience of inputting their name, which the system autonomously extracts from the uploaded document, subsequently cross-referencing it with the user-input name to ensure consistency and accuracy. In summary, through the amalgamation of advanced technologies such as OCR, machine learning, and meticulous attention to user experience, this study endeavors to develop a robust, reliable, and user-friendly biometric verification system. This system not only verifies signatures but also adeptly classifies document types, thereby catering to the heightened security demands of modern applications.[11]

C. Results

Siamese Network Implementation for Signature Verification



Document Forgery and Document Validation

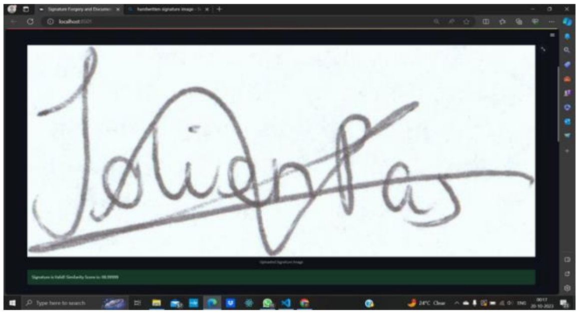
Welcome to the Signature Forgery and Document Validation Web App.

Signature Forgery Check

Select Signature Type: Digital

Upload Signature Image

Figure 4: Signature Verification



Signature is Valid! Similarity Score is: 98.99999

Figure 5: Signature Verification

The accuracy of our Siamese network model was evaluated through both training and validation phases. During training, the model achieved an impressive accuracy of approximately 95, indicating its capability to effectivelylearn and distinguish between similar and dissimilar pairs of data points. This high training accuracy underscoresthe robustness of the model in capturing intricate patterns and features within the dataset. the validation accuracy, obtained through rigorous testing on unseen data, yielded a commendable result of approximately 94.5. This validation accuracy reinforces the generalization ability of the model, demonstrating its capacity to perform well

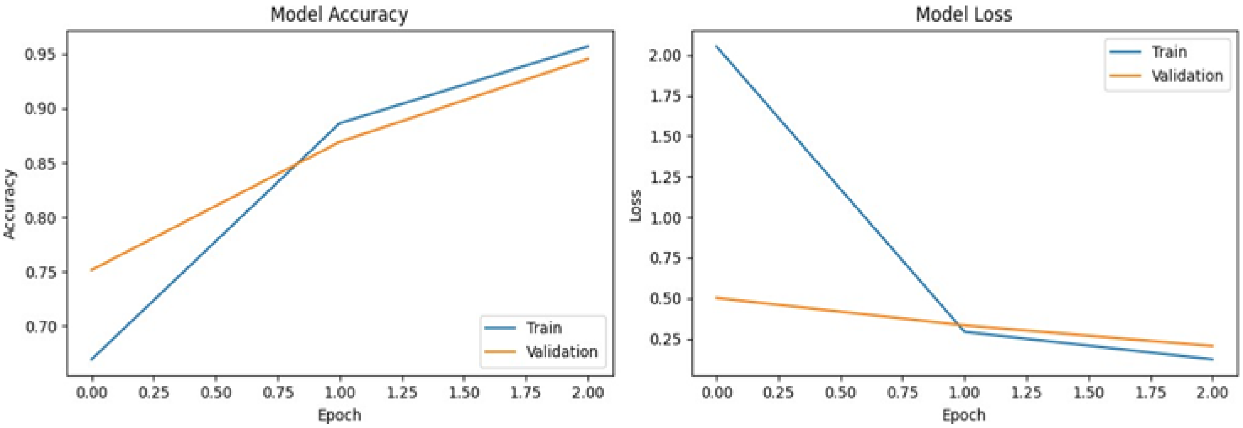
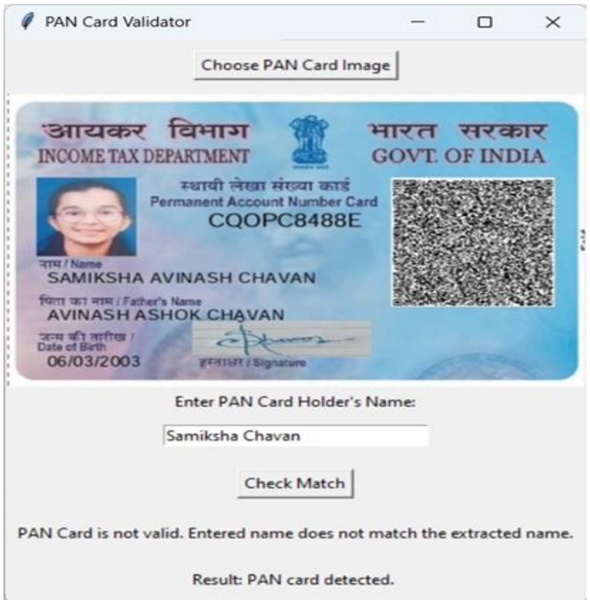


Figure 6: Model Accuracy



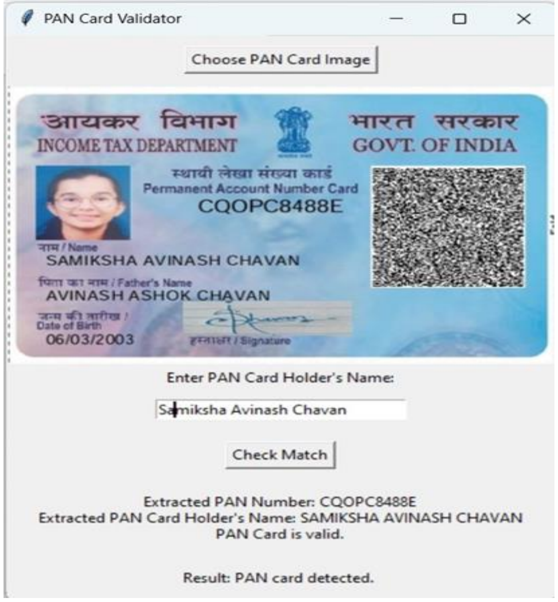
Enter PAN Card Holder's Name:

Samiksha Chavan

PAN Card is not valid. Entered name does not match the extracted name.

Result: PAN card detected.

Figure 7: Invalid Case



Enter PAN Card Holder's Name:

Samiksha Avinash Chavan

Extracted PAN Number: CQOPC8488E
Extracted PAN Card Holder's Name: SAMIKSHA AVINASH CHAVAN
PAN Card is valid.

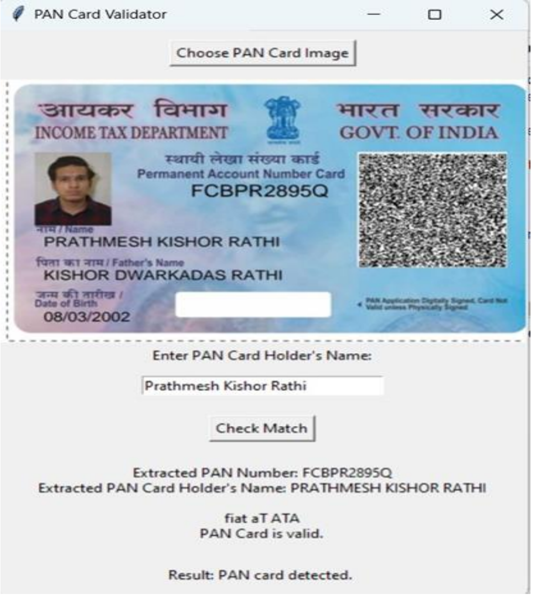
Result: PAN card detected.

Figure 8: Valid Case

Figure 9: Digital Pancard validation

Document classification and Pancard validation

In the context of our system, users are prompted to select the type of document they intend to upload, for instance,Aadhar card or PAN card. Upon selecting the desired document type, users proceed to upload the corresponding document file. Our system employs sophisticated document verification to scrutinize the uploaded document andverify its validity against the selected document type. If the uploaded document matches the selected document type, the system acknowledges its validity and proceeds with the intended task. However, if the uploaded documentdoes not correspond to the selected document type, the system promptly identifies it as an invalid document. In such cases, users are promptly notified of the discrepancy and prompted to upload the correct document type. The selected document type was a pancard but uploaded document is an addhar card hence it shows result as ”Not a Pancard”.



Enter PAN Card Holder's Name:

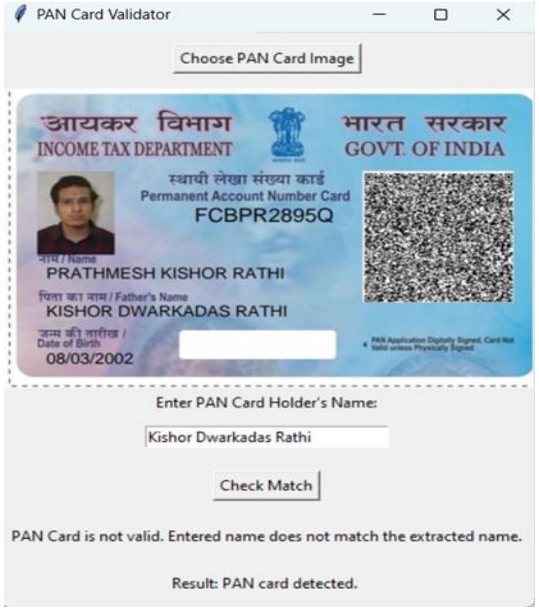
Prathmesh Kishor Rathi

Extracted PAN Number: FCBPR2895Q
Extracted PAN Card Holder's Name: PRATHMESH KISHOR RATHI

fiat aT ATA
PAN Card is valid.

Result: PAN card detected.

Figure 10: Valid Case



Enter PAN Card Holder's Name:

Kishor Dwarkadas Rathi

PAN Card is not valid. Entered name does not match the extracted name.

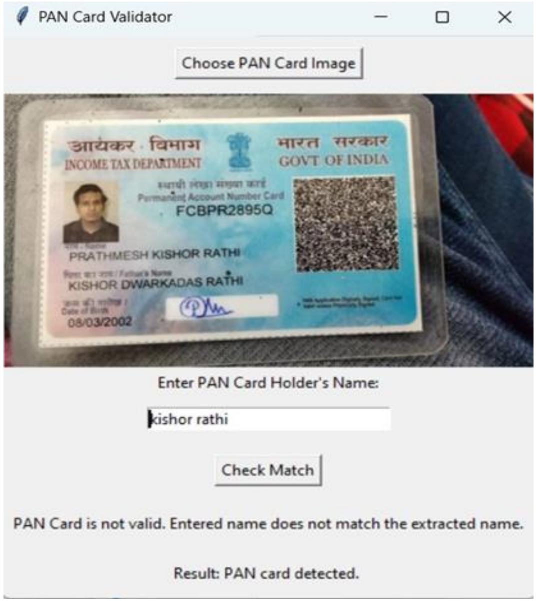
Result: PAN card detected.

Figure 11: InValid Case

Figure 12: Digital Pancard Validation

The system extracts text from the uploaded PAN card image, with particular emphasis on retrieving the full name. Following the extraction, users are prompted to input their full name for validation. The system then compares the provided full name with the extracted full name from the PAN card. Test success is contingent on anexact match between the provided and extracted full names, with any deviation signaling a test failure.

In the realm of identity verification systems, the detection of a PAN card, despite an incorrect name, under- scores the critical importance of stringent validation processes. While the system may successfully identify the document, discrepancies in essential details like the individual's name highlight the need for thorough verification mechanisms. Enhancing the accuracy of such systems is paramount to ensuring the integrity and reliability of the information they handle.

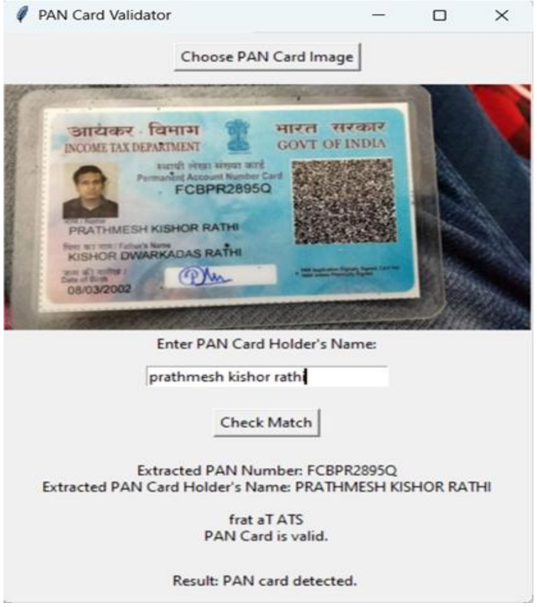


Enter PAN Card Holder's Name: kishor rathi

PAN Card is not valid. Entered name does not match the extracted name.

Result: PAN card detected.

Figure 13: Invalid Case



Enter PAN Card Holder's Name: prathmesh kishor rathi

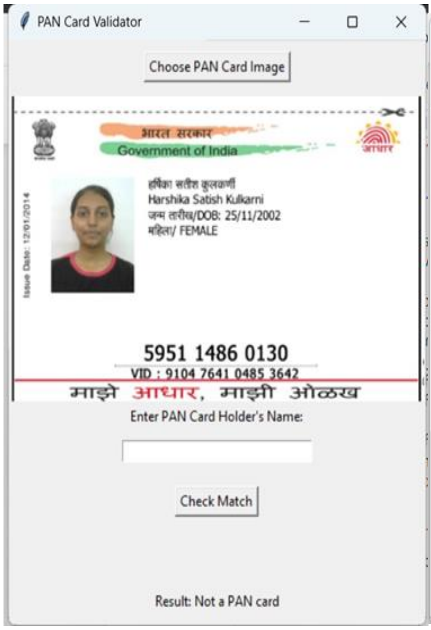
Extracted PAN Number: FCBPR2895Q
Extracted PAN Card Holder's Name: PRATHMESH KISHOR RATHI

frat aT ATS
PAN Card is valid.

Result: PAN card detected.

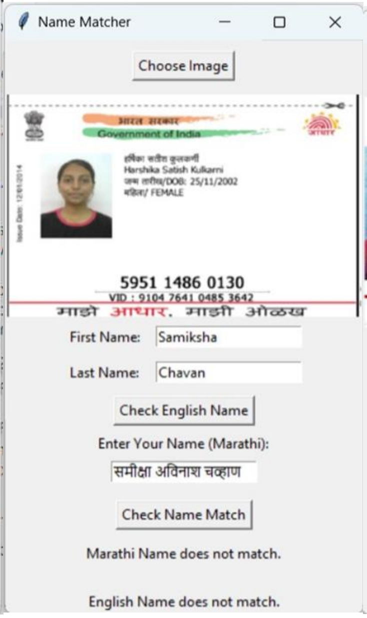
Figure 14: Valid Case

Figure 15: Scanned Pancard Validation



Enter PAN Card Holder's Name:

Result: Not a PAN card



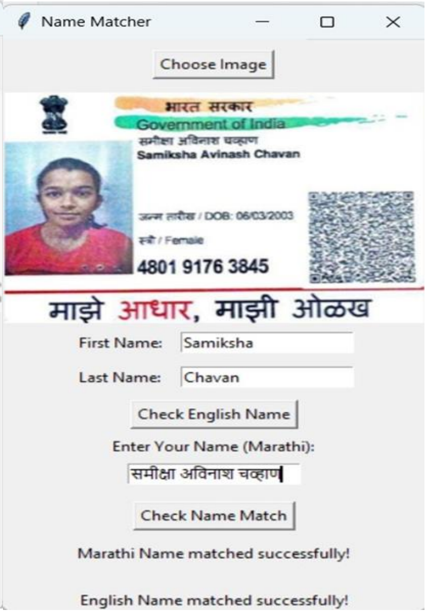
First Name: Samiksha

Last Name: Chavan

Enter Your Name (Marathi): समीक्षा अविनाश चव्हाण

Marathi Name does not match.

English Name does not match.



First Name: Samiksha

Last Name: Chavan

Enter Your Name (Marathi): समीक्षा अविनाश चव्हाण

Marathi Name matched successfully!

English Name matched successfully!

Figure 16: Invalid Document Upload Figure 17: InValid Case Figure 18: Valid Case

Figure 19: Aadhar Card Validation

Document classification and Addhar validation

The below figures shows how it works for addhar card its authenticity in accordance with the selected documenttype. If the uploaded document aligns with the chosen document type, the system recognizes its validity and allows the intended process to continue. However, if the uploaded document does not match the selected document type, the system promptly flags it as invalid.It incorporates Optical Character Recognition (OCR) technology to accurately recognize and match names written in both English and Marathi languages. This feature enables usersto seamlessly authenticate their identity regardless of whether their name is written in English or Marathi on their Aadhar card.scans the text on the uploaded Aadhar card, extracts the name details, and then meticulously compares them against the user-provided input. This meticulous matching process guarantees precise verification, irrespective of whether the name is in English or Marathi script.

Applications

The methodology presented offers versatile solutions for document forgery detection, with applications span-ning industries such as finance, government, legal, insurance, education, corporate, e-commerce, healthcare, notary services, and law enforcement.

1. Financial Institutions: Banks, credit unions, and financial service providers benefit from authenticating signatures on financial agreements, including checks, loan documents, and investment contracts, thereby reducing the risk of fraudulent transactions and financial losses.
2. Government Agencies: Departments such as immigration, law enforcement, and border control enhance national security through identity verification. This involves verifying signatures on passports, visas, identity documents, and official government forms, ensuring the integrity of legal proceedings and safeguarding against identity theft and fraudulent activities.
3. Legal Firms: Law firms and legal professionals validate signatures on legal documents such as contracts, deeds, wills, and court filings. This ensures the authenticity of signatures, preventing contract disputes, upholding the validity of legal agreements, and maintaining the integrity of the judicial system.
4. Insurance Companies: Providers mitigate fraudulent claims by authenticating signatures on insurance policies, claim forms, medical records, and legal affidavits. This assists in detecting fraudulent activities like insurance fraud, false claims, and identity theft, ultimately reducing financial losses and ensuring fair compensation for policyholders.
5. Educational Institutions: Universities, colleges, and academic institutions safeguard academic credentials by authenticating signatures on academic transcripts, diplomas, certificates, and enrollment documents. This maintains the integrity of academic records, prevents credential fraud, and upholds the reputation of educational institutions.
6. Corporate Enterprises: Large corporations, businesses, and enterprises authenticate signatures on business contracts, procurement documents, partnership agreements, and employee records. This helps prevent contract disputes, ensures compliance with legal regulations, and protects sensitive corporate information.
7. E-commerce Platforms: Online retailers and e-commerce platforms ensure the security of online transac- tions by authenticating signatures on shipping documents, delivery receipts, purchase orders, and customeragreements. This fosters trust among online shoppers, reduces the risk of fraudulent transactions, and safeguards sensitive financial information.
8. Healthcare Providers: Hospitals, clinics, and healthcare facilities comply with patient privacy regulationsby authenticating signatures on medical records, consent forms, treatment plans, and healthcare directives. This ensures patient confidentiality, prevents medical identity theft, and fosters trust in the healthcare system.
9. Notary Services: Notary publics and notarial services maintain the integrity of notarial acts by authenticating signatures on notarized documents, including affidavits, sworn statements, deeds, and powers of attorney. This upholds the legality and validity of legal documents, prevents fraud, and ensures the rule of law.
10. Law Enforcement Agencies: Police departments, law enforcement agencies, and investigative bodies aid criminal investigations by authenticating signatures on arrest warrants, search warrants, court orders, and evidentiary documents. This assists in identifying suspects, gathering evidence,

and ensuring the integrity of legal proceedings.

Overall, the methodology offers a versatile solution for document forgery detection, contributing to enhanced security, compliance, and trust across various industries and sectors in the digital age.

CONCLUSION

In conclusion, this research has presented a comprehensive methodology for Document Forgery Detection, focus- ing on two main components: Signature Verification and Document Classification. Siamese networks for signature analysis and OCR techniques for document classification has been proposed to enhance the overall efficiency and accuracy of the forgery detection system. The Signature Verification component leverages a Siamese network architecture, which proves to be robust in discerning between genuine and fraudulent signatures. Through metic- ulous data collection and advanced image processing techniques, the system achieves a thorough understanding of signature styles, making it resilient against various forgery attempts. The adaptability of Siamese networks for continuous learning ensures the system remains effective in dynamic environments. On the other hand, the Document Classification component integrates OCR techniques, document layout analysis, and machine learningmodels to identify and classify document types. This multi-modal verification system not only validates signatures but also checks document types, offering enhanced security applications. The incorporation of OCR technology for text extraction from documents further improves the system's granularity in document classification. The proposed methodology emphasizes user experience by incorporating a user-friendly interface with clear instructions and feedback mechanisms. The validation process ensures the accuracy and reliability of the system, with ongoing refinement to adapt to evolving fraudulent techniques. Through the amalgamation of advanced technologies, including Siamese networks, OCR, and machine learning, this study strives to develop a robust, reliable, and user-friendly system capable of handling a diverse range of document types and signature styles. In summary, the proposed methodology holds promise in addressing the challenges associated with document forgery detec- tion, offering a holistic approach that combines visual and textual analysis for enhanced authenticity assessment.The continuous improvement and adaptation of the system contribute to its effectiveness in real-world scenarios, catering to the heightened security demands of modern applications.

REFERENCES

[1] Yang, Piaoyang, Wei Fang, Feng Zhang, Lifei Bai, and Yuanyuan Gao. ”Document Image Forgery Detection Based on Deep Learning Models.” Proceedings of the [Conference Name], 2022: [Page Range].ry Detection Based on Deep Learning Models.”

[2] Y. Qi, Y. Z. Song, H. Zhang, J. Liu, ”Sketch-based image retrieval via siamese convolutional neural network,” in: ICIP, 2016, pp. 2460–2464.

[3] G. Koch, R. Zemel, R. Salakhutdinov, ”Siamese neural networks for one-shot image recognition,” in: ICML, 2015, pp. 1–8.

[4] Bayar, Belhassen, and Matthew C. Stamm. ”A deep learning approach to universal image manipulation detection using a new convolutional layer.” Proceedings of the 4th ACM workshop on information hiding and multimedia security. 2016: 5-10.

[5] Subramaniam, Manjula, Teja E, and N Arpith Mathew. ”Signature Forgery Detection Using Machine Learn- ing.” Proceedings of the , 2018.

[6] A Krizhevsky, A., Sutskever, I., Hinton, G. E. ”Imagenet classification with deep convolutional neural networks.” In: Proceedings of the Neural Information Processing Systems (NIPS), 2012: 1097–1105.

[7] Haziq Idrose, Nouar AlDahoul, Hezerul Abdul Karim, Rehan Shahid, and Manish Kumar Mishra, "An Evaluation of Various Pre-trained Optical Character Recognition Models for Complex License Plates," 2023.

[8] A Text Recognition Data Generator, April 2022, [online] Available: https://github.com/Belval/ TextRecognition DataGenerator.

[9] Mr. Ramdas Bagawade, Rohit Kharat, Kishor Matsagar, Shravani Kharade, "AI BASED OCR FOR TEXT RECOGNITION FROM HANDWRITTEN DOCUMENTS," Volume:05/Issue:05/May-2023.

[10] M. Sonkusare and N. Sahu, "A survey on handwritten character recognition (hcr) techniques for English alphabets," Advances in

Vision Computing: An International Journal, vol. 3, pp. 1–12, March 2016.

[11] T. K. Hazra, D. P. Singh and N. Daga, “Optical character recognition using KNN on custom image dataset,” 2017 8th Annual Industrial Automation and Electromechanical Engineering.

[12] A Detailed Review on Text Extraction Using Optical Character Recognition, ICT Analysis and Applications- Lecture Notes in Networks and Systems, 2022, pp. 719-728. Author(s): Chhanam Thorat, Aishwarya Bhat, Padmaja Sawant, Isha Bartakke, Swati Shirsath. DOI: 10.1007/978-981-16-5655-2 69.