Signature Verification and Fraud Detecting Using Opencv and Machine Learning

Tasmiya Anjhum H N[1], Sachin R Gowda[2], Sai Chethana S P[3], Sangeetha R[4], Sathvik R Bharadwaj[5],

[1,2,3,4,5]Information and Science and Engineering, Malnad College of Engineering Hassan

Abstract— In the evolving landscape of financial transactions, ensuring the security and authenticity of signature verification on bank cheques is vital to prevent fraud. This paper presents a survey of techniques used in signature verification, covering traditional image processing methods such as feature extraction, edge detection, and shape analysis, alongside modern deep learning approaches like Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and hybrid models. The survey explores key challenges, including handwriting variations, image noise, and the need for reliable real-world systems. Solutions such as pre-processing techniques, data augmentation, and transfer learning are discussed to improve accuracy. Commonly used datasets for training and evaluating models are reviewed, highlighting their features and limitations. The paper also considers ethical concerns such as fairness, transparency, and the broader implications of deploying such systems in financial institutions. By integrating insights from both traditional and advanced methods, this survey provides a valuable reference for researchers and practitioners aiming to enhance the accuracy, robustness, and reliability of signature verification systems. Strengthening these systems contributes to the overall integrity and trust in digital financial transactions.

Key words: Bank Cheques, CNN, OCR, Line Sweeping.

I. INTRODUCTION

The security and authenticity of financial transactions, particularly those involving bank cheques, are critical in preventing fraud. Traditional manual verification methods are slow, error-prone, and subjective, making them unsuitable for handling the large volume of transactions in today's financial landscape. Advanced technologies, such as Optical Character Recognition (OCR) and machine learning, are increasingly used to enhance accuracy and efficiency in signature verification.

Handwritten signatures remain a widely accepted method of individual authentication for legal and financial documents. However, verifying these signatures on a large scale presents significant challenges due to variations in writing styles and the potential for forgeries. Automatic signature verification systems aim to address these issues by reducing fraud and streamlining financial processes.

This study explores a multi-layered approach to signature classification and verification of bank cheques. The methodology combines OCR to extract textual information, line sweep techniques to analyse spatial features, and Convolutional Neural Networks (CNNs) to classify and authenticate signatures. These advanced technologies collectively create a robust system capable of distinguishing genuine signatures from forgeries, ensuring greater reliability and security in financial transactions.

By automating the verification process, the proposed system minimizes errors and establishes a higher level of trust in financial exchanges. The subsequent sections delve deeper into the methodology, supported by empirical evidence showcasing the system's effectiveness and potential.

II. RELATED WORK

[1]A signature verification tool combining image processing and deep learning techniques has demonstrated impressive accuracy in verifying key cheque components, such as branch identifiers, cheque serials, monetary sums, account numbers, and signature styles. By leveraging the IDRBT cheque dataset and convolutional neural networks (CNNs), the system achieved a precision of 99.14% for handwritten numeric characters and 97.7% for machine-printed script using MATLAB's OCR. Additionally, the

integration of Scale Invariant Feature Transform (SIFT) for signature feature extraction and Support Vector Machine (SVM) classification resulted in a 98.10% accuracy for signature verification. Despite these results, the tool faces challenges when handling illegible handwriting, smudged or damaged cheques, and atypical signature patterns. Its performance is also influenced by image quality and variations in cheque design, necessitating ongoing updates to maintain effectiveness in real-world applications.

[2]Another study introduced a recurrent attention optical character recognition network (RA_OCRN) to extract legal information from cheque images, aiming to enhance verification accuracy. This approach incorporated data pre-treatment and augmentation techniques, achieving superior results on the Kaggle cheque detection dataset compared to previous methods. The RA_OCRN model, when combined with data augmentation, demonstrated the highest recognition accuracy for its application. However, its scope is limited to the recognition of legal information and does not address broader challenges like damaged or altered cheques. Furthermore, the model's performance with handwritten or non-standard fonts remains unexamined, and the computational resources required for real-world deployment could pose additional challenges.

[3]A hybrid method combining dynamic and static features was proposed for handwritten signature verification, enhancing authenticity and safeguarding customer assets in the banking sector. Traditional manual approaches to signature verification, prone to errors and subjectivity, often result in financial losses and strained customer relationships. The proposed digital system improves customer service, strengthens staff-client interactions, and secures business transactions. Despite its potential, the system faces challenges, including vulnerabilities to impersonation, forgery, and electronic signature counterfeiting. Moreover, its reliance on machine learning and image processing introduces technological complexities and potential risks.

[4]Signature verification plays a pivotal role in biometric identification, particularly in banking and financial transactions. Techniques such as the Hopfield neural network, Back-propagation algorithm, Support Vector Machine (SVM), Hidden Markov Models, and Artificial Neural Networks (ANNs) have been widely utilized. The adoption of ANN for signature validation is a prominent trend due to its adaptability and accuracy. However, these systems face challenges such as vulnerability to skilled forgeries, lack of signature standardization, computational complexity, and dependency on high-quality training data.

[5]A three-layer signature verification system employing writer-independent and offline verification was proposed for bank cheque images. The system integrates graph metrical and FAST feature extraction methods, with classification achieved through Artificial Neural Networks, Gaussian Mixture Models, and image matching techniques, attaining a 99% accuracy rate. While effective, the system struggles with inconsistencies in signature image quality, resource-intensive real-time processing, and vulnerabilities to adversarial attacks. Handling complex or overlapping signatures also presents significant challenges.

[6]Offline signature verification has been improved with a CNN model developed in Python, achieving a testing accuracy of 99.70%. The model addresses inefficiencies in large-scale document verification. Static and dynamic verification types were compared, with static verification focusing on post-signature validation, while dynamic methods, including biometric approaches like fingerprint or eye scans, are gaining traction. However, dynamic verification faces obstacles such as high infrastructure costs, resistance due to privacy concerns, and difficulties for individuals with physical limitations. Moreover, ensuring secure storage and preventing misuse of biometric data remain critical concerns.

[7]A device-independent online handwritten signature verification system (ASSV) has been developed, leveraging acoustic signals to verify paper-based signatures. Unlike traditional systems, ASSV does not require specialized hardware like sensor-equipped pens or tablets, enhancing accessibility. The system combines a chord-based phase estimation method with discrete cosine transform (DCT) for feature extraction, followed by a CNN model for signature verification. Testing revealed strong performance, with an Area Under the Curve (AUC) of 98.7% and an Equal Error Rate (EER) of 5.5%. However, environmental noise or variations in acoustic conditions could impact reliability. Additionally,

challenges such as variability in signing styles and inaccuracies in phase-related changes or frequency-domain features extraction may affect its effectiveness in complex scenarios.

[8]Online attacks have advanced rapidly, outpacing traditional two-factor authentication. Signature verification remains vital for legal and financial transaction authentication. Handwritten signatures are unique to individuals and challenging to duplicate, making them a reliable security method. However, skilled forgeries can undermine their integrity. Automated signature verification systems must address the difficulty of distinguishing authentic signatures from forgeries. Integrating user involvement in attack prevention through signature verification requires robust security measures, technological advancements, and user education. The success of such approaches hinges on user compliance and system resilience.

[9]Automation in handwritten bank cheque processing aims to streamline verification and expedite the identification of bounced cheques. This involves pre-processing, information extraction, recognition, and verification of cheque details. Automation enhances efficiency and reduces manual effort, allowing for quicker actions on invalid cheques. However, the system may face challenges with handwritten style variations, image quality, and handling non-standard cheques. Initial investments in technology and training, as well as integration with existing banking systems, may pose significant hurdles. Ensuring system security and fraud prevention is critical for widespread adoption.

[10]An automated cheque processing system combines Optical Character Recognition (OCR) and deep learning to verify details like payee name, amount, date, and bank information. Signature verification employs feature extraction and principal component analysis, with user signatures securely stored as hash values. The system, trained on the IAM dataset, improves processing speed and reduces costs. Despite its advantages, challenges include handwriting variability leading to false positives or negatives, inconsistent data extraction accuracy due to poor cheque quality, and the significant initial investment required for implementation. Privacy and security implications of storing hashed signatures and adhering to regulatory standards are essential considerations for deployment.

[11]This systematic review explores the cutting-edge machine learning-driven models for offline signature verification (OfSV) systems are examined, assessing five key aspects: datasets, preprocessing methods, feature extraction techniques, machine learning-based verification models, and performance evaluation criteria. This review encompasses articles released from January 2014 to October 2019, encompassing 56 meticulously chosen studies. It unveils that deep learning-powered neural networks exhibit promising outcomes for OfSV systems on publicly available datasets. The assessment consolidates the performance of OfSV systems across five public datasets (CEDAR, GPDS, MCYT-75, UTSig, and BHSig260) and pinpoints fifteen unresolved research matters for future advancement. However, it's noteworthy that this review is restricted to articles published between 2014 and 2019, potentially disregarding recent advancements. Moreover, it might not encompass unpublished or non-peer-reviewed research within the domain.

[12]Handwritten signature authentication plays a critical role in distinguishing authentic signatures from forgeries, especially in legal and financial contexts. Significant progress has been made in the past decade, with machine learning techniques improving accuracy through advancements in feature extraction and classification. A review of over 20 studies highlights the comparative effectiveness of various datasets and methodologies. However, challenges remain, including the need for extensive and diverse datasets, difficulties in handling writing style variations, and vulnerabilities to advanced forgery techniques. Additionally, ensuring computational efficiency and interpretability in feature extraction and classification models remains a key concern.

[13]The Cheque Truncation System (CTS) has introduced automation and standardization to cheque verification through image processing and pattern recognition techniques. By reducing manual intervention, CTS enhances efficiency and security. However, it is not impervious to sophisticated forgeries, which can mimic genuine cheques. The study highlights that the system's limitations include gaps in detecting all forgery types,

dependency on forensic expertise, and the evolving nature of counterfeit methods. Regular updates and continuous assessment are necessary to maintain its effectiveness.

[14]A robust signature verification tool employing image processing and machine learning has been proposed to address fraud in financial transactions. Key techniques include edge detection and feature extraction using SURF and SIFT algorithms, with extracted features compared against stored signatures in a database. The tool demonstrates high accuracy and operational speed. Nonetheless, it faces challenges such as sensitivity to image quality, potential susceptibility to adversarial attacks, and dependency on comprehensive signature databases. These factors necessitate ongoing updates to address emerging fraud techniques and maintain performance across large datasets.

[15]The integration of the Tesseract framework with Convolutional Neural Networks (CNNs) facilitates automated cheque data extraction, focusing on retrieving details like bank name, payee information, and transaction date. Enhancements in CNN-based models improve recognition accuracy and streamline processes. However, the approach struggles with handwritten or damaged text, non-standard cheque formats, and real-time processing demands. Implementation challenges include high costs, significant infrastructure requirements, and ensuring security and privacy in data handling, which could hinder adoption in smaller institutions.

[16]A comprehensive forgery detection system combines image processing, OCR, and machine learning to identify counterfeit signatures and ensure transaction security. Practical applications span banking, legal, and administrative domains. While the system significantly reduces fraud, limitations include its reliance on input image quality, computational resource demands, and the need for regular updates to counter increasingly sophisticated forgery methods. Ensuring robustness against determined fraudulent attempts remains a priority.

[17]Signatures, a universal method for identity verification, face inherent vulnerabilities, including potential for forgery, variations in appearance over time, and challenges posed by physical limitations or disabilities. E-signatures, while addressing some issues, introduce concerns related to security and authenticity, particularly across different platforms and devices. Moreover, variations in legal standards globally create inconsistencies, complicating cross-border transactions and document verification processes.

[18]Biometric authentication systems, which utilize physical or behavioural traits for identity verification, have gained prominence as a secure alternative to traditional methods such as passwords or tokens. Signature-based biometrics offer enhanced security but are affected by variations in individual signatures over time. Privacy concerns related to data collection and storage, susceptibility to spoofing, and the high costs of deployment present significant barriers to widespread adoption. Despite their potential, these systems require advancements in robustness and efficiency for practical large-scale implementation.

[19]A proposed system leveraging OCR and signature verification automates cheque validation by extracting details like payee name, amount, and transaction date. Utilizing OpenCV and Easy OCR for text extraction, the system demonstrates promising results when tested on custom datasets. However, its performance heavily depends on the quality and consistency of input cheque images. Real-time validation and integration with diverse banking technologies pose additional challenges, necessitating further optimization to enhance applicability and reliability in practical settings.

III. TECHNIQUES USED IN SIGNATURE AUTHENTICATION

A. OCR with LineSweep methods

Optical Character Recognition (OCR) with line sweep is a technique used to extract text information from images by systematically scanning the content along horizontal or vertical lines. This process involves analyzing the image pixel by pixel, recognizing patterns that form characters, and reconstructing the text. Below are the key details of OCR with line sweep:

1) Line Sweep Algorithm:

Directional Scanning: Line sweep involves scanning the image along lines, typically horizontally or vertically. This directional scanning simplifies the recognition process.

Pixel-by-Pixel Analysis: The algorithm processes each pixel along the scanning lines. It examines pixel intensity,

color, and neighboring pixels to identify features that correspond to characters.

Sequential Processing: The image is processed sequentially along each line, allowing the algorithm to reconstruct the text in a systematic manner.

2) Preprocessing:

Noise Reduction: Before line sweep, preprocessing techniques may be applied to reduce noise in the image. Common preprocessing steps include blurring, thresholding, and morphological operations.

Binarization: Converting the image to binary form simplifies the subsequent analysis. This step involves separating the text from the background by setting a threshold for pixel intensity.

3) Character Recognition:

Feature Extraction: As the line sweep progresses, the algorithm identifies features such as edges, corners, and connected components that correspond to characters.

Pattern Matching: Recognized features are compared to pre-trained character templates or models. Pattern matching algorithms, such as template matching or machine learning-based methods, can be employed.

4) Text Reconstruction:

Sequential Assembly: As characters are recognized along the lines, they are sequentially assembled to reconstruct words and sentences.

Contextual Analysis: Some OCR systems incorporate contextual analysis to improve accuracy by considering the surrounding characters and words.

5) Postprocessing:

Error Correction: Postprocessing steps may include error correction techniques to enhance the accuracy of the OCR results. This can involve spell-checking, context-based corrections, and machine learning algorithms.

6) Challenges and Considerations:

Text Orientation and Skew: OCR with line sweep may face challenges with skewed or rotated text. Techniques to handle text orientation and skew correction may be incorporated.

Font and Style Variations: Different fonts and writing styles can pose challenges. Robust OCR systems often include a diverse set of training data to handle variations.

7) Applications:

Document Scanning: OCR with line sweep is commonly used in document scanning applications to convert printed or handwritten text into editable and searchable formats.

Automated Data Entry: It is employed in automated data entry systems to extract information from forms, invoices, and other structured documents.

OCR with line sweep provides a systematic and efficient approach to text extraction from images, making it a fundamental component in various applications that involve text recognition and document processing.

B. Data Augmentation:

The subsequent code presents the implementation of the data augmentation process employed to generate images and artificially expand the dataset's size. In this context, the Image Data Generator is utilized to create novel images. The rotation range is configured to 360 degrees since the images, being spiral in nature, can be rotated to various angles without altering the inherent meaning of the image. It is advisable to explore other image transformations within the Image Data Generator class; however, caution is advised while applying augmentations as certain transformations might potentially reduce the CNN model's overall accuracy. Post augmentation, the distribution of data is adjusted.

The images are also resized to a uniform dimension of (128, 128, 1), and subsequent to resizing, image normalization is executed before incorporating the dataset into the model.

C. CNN Model Architecture:

The implementation adopts a CNN model architecture characterized by the ensuing attributes:

- The model encompasses four Convolutional Layers housing 128, 64, 32, and 32 filters, respectively.
- Diverse filter sizes are employed across the convolutional layers.
- Each convolutional layer is succeeded by a MaxPool2D layer.
- A pair of Fully Connected layers ensue after the convolutional block.

Specification of the Model using Keras:

D. Model Training:

The model undergoes training with a learning rate of 3.15e-5 using the Adam optimizer. The number of epochs is set at 70, and the batch size is established as 128.

E. Model Performance:

To assess the model's performance, various metrics are employed, including Loss and Accuracy Plots, Classification Report, and Confusion Matrix. These

evaluations provide insights into the model's effectiveness and accuracy.

IV. CONCLUSION

In conclusion, this survey has provided a comprehensive overview of signature verification on bank cheque images, amalgamating insights from both traditional image processing and cutting-edge deep learning approaches. The synthesis of these methodologies has demonstrated a synergistic potential for addressing the challenges inherent in signature verification systems, such as variations in writing styles and the need for robustness in real-world scenarios. The emergence of hybrid models, combining the interpretability of image processing with the automatic feature learning of deep learning, marks a significant stride toward achieving more accurate and resilient signature verification systems.

Looking ahead, the survey emphasizes the ethical considerations surrounding the deployment of these systems in financial institutions and underscores the importance of fairness, transparency, and accountability. As technology continues to advance, future research directions may involve the exploration of novel deep learning architectures, the expansion of diverse training datasets, and the development of standardized evaluation metrics. By fostering a deeper understanding of the current state-of-the-art and charting potential pathways for future enhancements, this survey serves as a valuable resource for researchers, practitioners, and stakeholders seeking to bolster the security and reliability of signature verification processes in the ever-evolving landscape of financial transactions.

REFERENCES

[1] Prateek Agrawal, Deepak Chaudhary, Vishu Madaan, Anatoliy Zabrovskiy, Radu Prodan, Dragi Kimovski and Christian Timmerer. "Automated bank cheque verification using image processing and deep learning methods". (2021) Multimedia Tools and Applications. 80. 1-32. 10.1007/s11042-020-09818-1.

[2] Hitesh Chaitanyaswami, and Ashwin Dobariya. "Deep Learning Recurrent Attention Optical Character Recognition Network with Data Augmentation for Cheque Data Extraction." *2023 International Conference on Computer, Electronics & Electrical Engineering & their Applications (IC2E3)* https://doi.org/10.1109/IC2E357697.2023.10262478

[3] Dr.Thangavel V, "Use of Digital Signature Verification System (DSVS) in Various Industries: Security to Protect against Counterfeiting: Research". (April 24, 2023). SSRN: https://ssrn.com/abstract=4461014 or http://dx.doi.org/10.2139/ssrn.4461014

[4] Vijaya Dhopte1 , Prof. Shaila P. Kharde, "Signature Verification of Bank Cheque" IJIRSET Volume 10, Issue 10, October 2021 DOI:10.15680/IJIRSET.2021.1010093

[5] Bramara Neelima K and S Arulselvi "Signature Extraction and Recognition from Bank Cheque Image" IJITEE 2019 DOI: 10.35940/ijitee.B7766.129219

[6] Eman Alajrami, Belal A. M. Ashqar, Bassem S. Abu-Nasser, Ahmed J. Khalil, Musleh M. Musleh, Alaa M. Barhoom, Samy S. Abu-Naser "Handwritten Signature Verification using Deep Learning" International Journal of Academic Multidisciplinary Research (IJAMR) ISSN: 2643-9670 Vol. 3 Issue 12, December – 2019,

[7] Feng Ding, Dong Wang, Qian Zhang, and Run Zhao. 2019. "ASSV: Handwritten Signature Verification Using Acoustic Signals". Proc. ACM Interact. (September 2019), https://doi.org/10.1145/3351238

[8] Yash Borse, Anjali Patil, Sumeet Shah, Anand Gharu "Signature Verification Using Deep Learning" IJCRT 2023 Volume 11, Issue 4 April 2023 ISSN: 2320-2882

[9] Phalguni Nikam, Prof. Pramod Patil, Meetali Patidar, Aishwarya Nanoskar, Pranav Parmar and Jayesh More. "Checkue Bounce Detection System Using Image Processing". IRJET-volume 7 issue:01 Jan 2020 DOI https://doi.org/10.1007/s10032-011-0170-8.

[10] Mukesh Jha, Madhur Kabra, Sahil Jobanputra and Rupali Sawant "Automation of Cheque Transaction using Deep Learning and Optical Character Recognition." *International Conference on Smart Systems and Inventive Technology (ICSSIT)* (2019): DOI https://doi.org/10.1109/ICSSIT46314.2019.8987925

[11] M. Muzaffar Hameed, Rodina Ahmad, Miss Laiha Mat Kiah and Ghulam Murtaza (2021). "Machine learning-based offline signature verification systems:

A systematic review." ResearchGate (2021) DOI http://dx.doi.org/10.1016/j.image.2021.116139

[12] Zainab Hashim, Hanaa M.Ahmed and Ahmed Hussein Alkhayyat, "A Comparative Study among Handwritten Signature Verification Methods Using Machine Learning Techniques" Hindwai vol.2022, DOI https://doi.org/10.1155/2022/8170424

[13] Vishnu Dev Yadav, Khushboo Phore and Loganathan Lingan. "Bank Cheque Fraud and Cheque Truncation System to Illuminate the Fraudulent Transaction: A Case Study" J Forensic Res 14 (2023), DOI: 10.37421/2157-7145.2023.14.562

[14] Walid Hussein, Mostafa A. Salama and Osman Ibrahim "Image Processing Based Signature Verification Technique to Reduce Fraud in Financial Institutions"MATEC Web Conf. Volume 76, 2016 https://doi.org/10.1051/matecconf/20167605004

[15] Prajwal V Suranagi, and Dr T R Arunkumar "Automated Cheque Image Processing AI Driven Recognition For Efficient banking" irjmets Volume:05/Issue:09/September-2023 DOI : https://www.doi.org/10.56726/IRJMETS44875

[16] Navya V K, Abhilasha Sarkar, Aditi Viswanath, Akshita Koul, and Amipra Srivastava "Signature Verification And Forgery Detection System" IJCRT Volume 11, Issue 6 June 2023 ISSN: 2320-2882

[17] M. B. Sudhan, Abhilasha Sarkar, Aditi Viswanath, Akshita Koul, and Amipra Srivastava "A Survey on Signature Verification and Forgery Detection System" International Journal of Research Publication and Reviews, Vol 4, no 4, pp 1735-1738, April 2023 https://doi.org/10.55248/gengpi.2023.4.4.35359

[18] Shraddha Jagtap, Sejal Kalyankar, Tejal Jadhav and Ashwini Jarali "Signature Verification And Detection System" International Journal of Recent Scientific Research Vol. 13, Issue, 06 (A), pp. 1412-1418, June, 2022 http://dx.doi.org/10.24327/ijrsr.2022.1306.0298

[19] P. Kunekar, K. Vayadande, O. Kulkarni, K. Ingale, R. Kadam and S. Inamdar, "OCR based Cheque Validation using Image Processing," 5th Biennial International Conference on Nascent Technologies in Engineering (ICNTE), 2023, pp. 1-5,doi : 10.1109/ICNTE56631.2023.10146687.

.