A

Project Report

on

Automatic Cheque Detection

submitted as partial fulfillment for the award of

BACHELOR OF TECHNOLOGY DEGREE

SESSION 2023-24

in

Computer Science and Engineering

By

Yogesh Kumar Singh(2000290100199)

Rishabh Jaiswal(2000290100114)

Prerna Choudhary(2000290100105)

Under the supervision of

Ms. Bharti Chugh

KIET Group of Institutions, Ghaziabad

Affiliated to

Dr. A.P.J. Abdul Kalam Technical University, Lucknow

(Formerly UPTU)

May, 2024

DECLARATION

We hereby declare that this submission is our own work and that, to the best of our knowledge and belief, it contains no material previously published or written by another person nor material which to a substantial extent has been accepted for the award of any other degree or diploma of the university or other institute of higher learning, except where due acknowledgment has been made in the text.

Date: 12/05/2024

Signature:	Signature:	Signature:
Name: Yogesh Kumar	Name: Rishabh Jaiswal	Name: Prerna Choudhary
Roll No.:2000290100199	Roll No.:20002901001	Roll No.:2000290100105

CERTIFICATE

This is to certify that Project Report entitled "**Automatic Cheque Detection**" which is submitted by **Yogesh Kumar, Rishabh Jaiswal , Prerna Choudhary** in partial fulfillment of the requirement for the award of degree B. Tech. in Department of Computer Science & Engineering of Dr. A.P.J. Abdul Kalam Technical University, Lucknow is a record of the candidates own work carried out by them under my supervision. The matter embodied in this report is original and has not been submitted for the award of any other degree.

Ms. Bharti Chugh
(Assistant Professor)

Dr. Vineet Sharma
(Head of Department)

Date: 12/05/2024

ACKNOWLEDGEMENT

It gives us a great sense of pleasure to present the report of the B. Tech Project undertaken during B. Tech. Final Year. We owe special debt of gratitude Prof. Bharti Chugh , Department of Computer Science & Engineering, KIET, Ghaziabad, for her constant support and guidance throughout the course of our work. Her sincerity, thoroughness and perseverance have been a constant source of inspiration for us. It is only his cognizant efforts that our endeavors have seen light of the day.

We also take the opportunity to acknowledge the contribution of Dr. Vineet Sharma, Head of the Department of Computer Science & Engineering, KIET, Ghaziabad, for his full support and assistance during the development of the project. We also do not like to miss the opportunity to acknowledge the contribution of all the faculty members of the department for their kind assistance and cooperation during the development of our project.

We also do not like to miss the opportunity to acknowledge the contribution of all faculty members, especially faculty/industry person/any person, of the department for their kind assistance and cooperation during the development of our project. Last but not the least, we acknowledge our friends for their contribution in the completion of the project.

Date:12/05/2024

Signature:	Signature:	Signature:
Name: Yogesh Kumar	Name: Rishabh Jaiswal	Name: Prerna Choudhary
Roll No.:2000290100199	Roll No.:20002901001	Roll No.:2000290100105

ABSTRACT

The integration of deep learning and image processing technologies has revolutionized the traditional process of verifying bank cheques. By leveraging innovative approaches, we've significantly minimized human intervention while maximizing efficiency in the cheque truncation mechanism. Our methodology focuses on extracting crucial details from cheque booklets, such as the bank branch code, cheque number, precise amount, account number, and unique signature patterns.

In our study, we capitalized on the IDRBT cheque dataset, employing convolutional neural networks (CNNs) to analyze handwritten components on bank cheques. This meticulous approach yielded an exceptional accuracy rate of 99.14% in identifying handwritten numeric characters, showcasing the robustness of deep learning in recognizing diverse handwriting styles.

Furthermore, our utilization of MATLAB's integrated optical character recognition (OCR) technique achieved an impressive accuracy of 97.7% in interpreting machine-printed text. This underscores the versatility and effectiveness of OCR technology in automating text extraction tasks with high precision.

For signature verification, we employed advanced techniques such as Scale Invariant Feature Transform (SIFT) for feature extraction and Support Vector Machine (SVM) for classification. This comprehensive approach resulted in a notable accuracy rate of 98.10% in authenticating signatures, demonstrating the efficacy of combining feature-based methods with machine learning algorithms.

TABLE OF CONTENTS

Page No.

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ANNs	Artificial Neural Networks
CNN	Convolutional Neural Network
CTS	Cheque Truncation System
DNNs	Deep Neural Networks
EMNIST	Extended Modified National Institute of Standards and Technology
GAN	Generative Adversarial Network
IDRBT	Institute for Development and Research in Banking Technology
IPV	Indian Place Value
KDE	Key Point Descriptor
KPL	Key Point Localization
MICR	Magnetic Ink Character Recognition
MNIST	Modified National Institute of Standards and Technology
OCR	Optical Character Recognition

CHAPTER 1
INTRODUCTION

1.1 INTRODUCTION

Bank cheques have maintained their significance in facilitating financial transactions globally, despite the rise of digital transactions. While digital methods are increasingly popular, cheques are still widely used. However, manual processing of cheques is laborious, expensive, and prone to errors. Automated bank cheque processing systems have emerged as a promising solution to address these challenges, prompting scholars to delve into this field.

Fig. 1.1. CTS-2010 compliant sample bank cheque

Automated bank cheque verification systems are essential tools in modern banking, offering efficiency and accuracy in processing financial transactions. These systems are designed to extract and verify crucial information from bank cheques, ensuring the integrity of transactions while minimizing errors and fraud risks. In this context, the extraction of specific information such as the date, account number, cheque number, account holder signature, courtesy amount, legal amount, and bank information (IFS Code) is paramount. Traditional manual processing methods are not only time-consuming but also prone to errors, highlighting the need for automated solutions to streamline the cheque verification process. By leveraging advanced technologies such as optical character recognition (OCR), deep learning, and image processing, automated systems can accurately extract and verify essential cheque details, enhancing the efficiency and reliability of banking operations. This paper explores the key features and functionalities required for an effective automated bank cheque verification system, aiming to address the challenges faced by the banking sector and improve the overall transaction processing experience.

Fig. 1.2. Scan of a sample cheque in its raw form

The Need for Automation:

The manual processing of bank cheques is cumbersome and error-prone, leading to inefficiencies and risks. Automated systems offer a solution by leveraging technologies such as computer vision, image processing, pattern recognition, machine learning, and deep learning. These technologies enable the extraction and verification of vital details from cheque images, such as bank branch codes, cheque numbers, amounts, and signature patterns.

Advancements in Fraud Detection:

Machine learning algorithms play a crucial role in detecting financial transaction fraud by analyzing features extracted from scanned cheque images. Modern algorithms like Generative Adversarial Networks (GANs) can even identify potential counterfeit products in real-time, enhancing fraud prevention measures.

The Role of Truncation Methods:

Research on truncation methods has also benefited from machine learning and artificial intelligence, streamlining the cheque clearance process. By automating tasks such as acquisition, preprocessing, segmentation, interpretation, and recognition, these techniques optimize efficiency and accuracy.

Contributions of the Study:

This research focuses on enhancing the precision of automated bank cheque verification through image processing and deep learning techniques. Key contributions include segmenting scanned cheque images to extract specific information and validating it against converted legal amount data using Convolutional Neural Networks (CNNs) and other technologies. The proposed model aims to improve the efficiency of financial transactions, reduce costs, and mitigate the risk of cheque fraud.

1.2 PROJECT DESCRIPTION

The project on automated bank cheque verification aims to streamline the authentication process through a series of essential steps executed in a specific sequence, ensuring precision and efficiency. The system begins by verifying the IFSC code and cheque number, ensuring that the cheque originates from the account holder's designated set of cheque booklets. This initial step is crucial as it establishes the legitimacy of the cheque's source, preventing potential fraud from unauthorized cheque use.

Following the verification of the cheque's origin, the system proceeds to authenticate the cheque issuer's signature(s) and the cheque amount. These verifications are cross-checked against the customer's account balance to confirm sufficient funds, preventing overdrafts and ensuring financial integrity. These steps are vital for all cheque-clearing processes, including withdrawals and transfers, thereby maintaining a secure banking environment.

To achieve these verifications, the project employs a range of accurate and effective methods. Optical Character Recognition (OCR) is utilized for extracting machine-printed text, allowing for the accurate reading of printed information such as the IFSC code and cheque number. For numerical and handwritten text recognition, deep learning-based Convolutional Neural

Networks (CNNs) are employed. CNNs are particularly adept at recognizing patterns and variations in handwritten text, making them ideal for this task.

The project also incorporates the Scale-Invariant Feature Transform (SIFT) method for signature authentication. SIFT is effective in identifying key features and outlines of signatures, ensuring that even minor variations are detected and authenticated accurately. The combination of these advanced technologies ensures that the verification process is both comprehensive and precise.

For optimal performance, the system requires noise-free images, which are achieved through a series of image preprocessing and segmentation techniques. Image preprocessing involves cleaning the images to remove any noise or distortions that could affect the accuracy of OCR and CNN analysis. Segmentation further isolates critical data points on the cheque, such as the signature, amount, and account details, ensuring that each element is analyzed independently and accurately.

Segmentation also plays a crucial role in identifying and extracting handwritten digits and letters. Through the use of transfer learning and OCR, the system can accurately recognize and interpret handwritten information, which is essential for processing cheques that may not adhere to standard printing formats.

In addition to these methods, a specialized CNN model is utilized for the identification of legal and courtesy amounts on the cheque. This model ensures that the amounts written in words (legal amount) and figures (courtesy amount) are accurately identified and matched, which is a critical requirement for cheque processing. This step ensures compliance with CTS-2010 criteria for Indian banks and international cheque regulations, guaranteeing that the cheques meet all necessary legal and procedural standards.

Signature verification is another critical aspect of the project. By extracting features and outlines using SIFT, the system can create a detailed map of the signature's unique characteristics. These features are then authenticated using Support Vector Machine (SVM) classifiers, which are highly effective in distinguishing between genuine and forged signatures. This dual approach of feature extraction and machine learning classification ensures that signature verification is both robust and reliable.

CHAPTER 2

LITERATURE REVIEW

Over the years, researchers have delved into the intricacies of bank cheque verification, striving to enhance the accuracy and efficiency of various processes involved. A cornerstone in this exploration has been the utilization of optical character recognition (OCR) techniques, which have demonstrated remarkable success in recognizing characters within scanned cheque images. Pioneering studies by Wankhede and Mohod (1983), Singh and Sachan (1977), and Ramanathan et al. (1970) have laid the foundation for OCR-based character recognition, showcasing its potential in the realm of picture character OCR.

Table 2.1 :Description of the main characteristics used for bank cheque verification

COMPONENT	DESCRIPTION
ACCOUNT NUMBER	Unique numerical identifier assigned to a specific bank account.
CHEQUE NUMBER	Sequential number assigned to each individual cheque for tracking and identification purposes.
ACCOUNT HOLDER SIGNATURE	Handwritten signature of the account holder for authentication and authorization.
COURTESY AMOUNT	The monetary amount written in words on the cheque as a friendly representation of the legal amount.
LEGAL AMOUNT	The monetary amount written in numerical digits on the cheque, legally binding.
BANK INFORMATION (IFS CODE)	Unique alphanumeric code identifying a specific bank branch in the electronic funds transfer network.
MICR NUMBER	Magnetic Ink Character Recognition number, a unique code printed at the bottom of cheques for automated processing.

Further innovations have emerged in segmentation techniques, addressing the challenge of isolating specific data within cheque documents. Chen's introduction of spectral clustering marked a significant breakthrough, achieving an impressive 97% accuracy in segmenting handwritten digits. Similarly, researchers have devised statistical techniques for recognizing text line Latin Language fonts with a commendable 97% recognition rate, demonstrating the efficacy of segmentation methodologies in handling diverse types of characters and fonts.

In parallel, the quest for accurate English character identification has spurred investigations into various methodologies. Notable among these is the successful application of support vector machines (SVMs), which achieved a notable accuracy of 93.54% in multi-knowledge categorization and feature extraction. The integration of convolutional neural networks (CNNs) with SVMs in a hybrid classifier, as demonstrated by Peres et al. (2018), further bolstered accuracy rates, underscoring the potential of machine learning approaches in character identification tasks.

Moreover, researchers have explored the impact of environmental factors such as background color, noise, and patterns on character recognition accuracy. Techniques such as binarization have been introduced to standardize amounts on Chinese bank cheques, while correlation coefficients have been utilized to estimate maximum amounts on Indian bank cheques. These endeavors have not only enhanced recognition accuracy but have also contributed to the development of robust segmentation algorithms capable of handling diverse cheque formats and backgrounds.

Additionally, efforts have been directed towards the extraction and verification of Magnetic Ink Character Recognition (MICR) codes, vital for secure cheque verification. Studies have addressed challenges such as shadow and spatial noise in low-quality cheque images, employing advanced image processing techniques to ensure accurate MICR code extraction. The integration of deep learning methods with traditional image processing techniques has further improved the accuracy and efficiency of MICR code extraction, contributing to the overall reliability of cheque verification systems.

Despite these advancements, there remains a gap in research pertaining to automated cheque handling systems and comprehensive cheque validation methodologies. This gap underscores the need for further research and development in this domain, with a focus on enhancing the efficiency and accuracy of cheque processing and reconciliation processes. By leveraging state-of-the-art OCR, segmentation, and signature verification techniques, this paper aims to address these challenges and pave the way for more robust and automated cheque handling systems.

Fig 2.1 .Images of sample cheques from various nations

CHAPTER 3

PROPOSED METHODOLOGY

The authentication process of bank cheques encompasses a series of indispensable steps, each meticulously executed in a predetermined sequence. Initially, the system undertakes the validation of the IFSC (Indian Financial System Code) present on the cheque, a crucial identifier ensuring the accuracy of financial transactions. Following this, the verification of the cheque number ensues, a pivotal step aimed at confirming its alignment with the designated set of cheque booklets assigned to the respective account holder. Subsequently, the process progresses towards the scrutiny of the signature(s) affixed by the cheque's issuer, coupled with a meticulous examination of the corresponding amount against the customer's account balance. The expeditious and precise completion of these verifications holds paramount importance in facilitating seamless cheque clearing operations, encompassing essential financial transactions such as withdrawals and transfers.

Throughout this intricate process, a diverse array of accurate and efficient methodologies comes into play to extract vital data from the cheque leaflet. Optical character recognition (OCR) emerges as a cornerstone, adeptly extracting information from machine-printed text with unparalleled precision. In tandem, deep learning-based Convolutional Neural Networks (CNNs) are deployed, demonstrating remarkable proficiency in handling both numerical data and handwritten text, thereby ensuring comprehensive coverage of diverse information formats. Meanwhile, the deployment of the Scale-Invariant Feature Transform (SIFT) algorithm serves to extract distinctive features crucial for authenticating signatures, bolstering the overall robustness of the verification process. Complementing this, Support Vector Machines (SVMs) are harnessed to classify extracted features, thereby enhancing the performance metrics and bolstering the system's efficacy.

Among these technologies, OCR, SIFT, SVM, and CNN stand out for their ability to extract, process, and analyze key information from bank cheques.

3.1 Key Components:

Before delving into the specifics of each key component, it's essential to understand the foundational elements that drive automated bank cheque verification. This section outlines the core components that form the backbone of the system, each playing a crucial role in ensuring the accuracy and reliability of the verification process.

3.1.1 CNN (Convolutional Neural Network):

CNNs are a subclass of deep neural networks that are mainly employed in the analysis of image data. Using a variety of building blocks, including convolutional layers, pooling layers, and fully connected layers, they are made to automatically and adaptively learn spatial hierarchies of features through backpropagation. CNNs can recognize both temporal and spatial dependencies in an image thanks to this architecture.

In tasks like image classification, object detection, and facial recognition, CNNs have achieved state-of-the-art results, revolutionizing computer vision. CNNs are indispensable for a variety of applications, from autonomous driving to medical image analysis, because they can efficiently identify complex patterns and objects in images by utilizing the hierarchical structure. CNN is a deep learning model that is specifically designed to process visual data in terms of functionality.

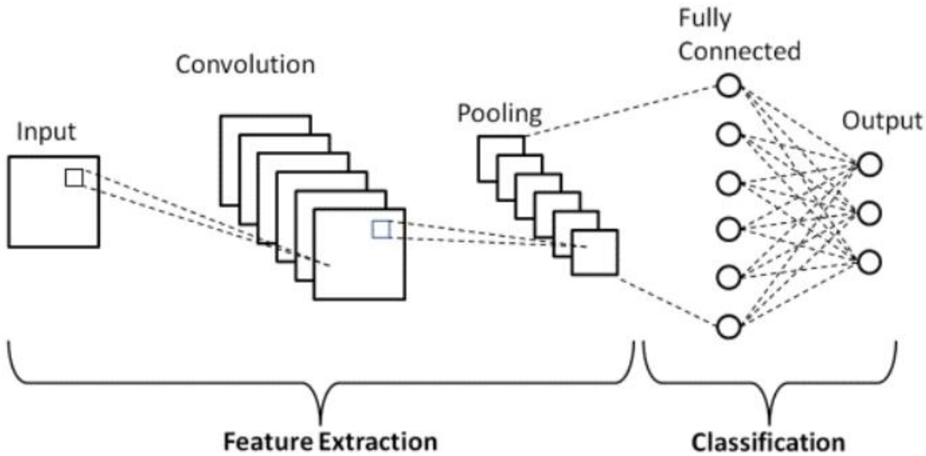

Fig 3.1 Basic CNN architecture

The main parts of a Convolutional Neural Network (CNN) are as follows:

a) Convolutional Layers: These layers use filters (kernels) to perform convolutional operations on input images in order to detect features such as textures, edges, and complex patterns. Convolutional operations improve the extraction of features by maintaining the spatial relationships among pixels.

b) Pooling Layers: These layers reduce network parameters and computational complexity by downsampling input spatial dimensions. One common operation that makes feature selection easier is max pooling, which takes the maximum value from a set of adjacent pixels.

c) Activation Functions: The model's non-linearity is introduced by non-linear activation functions, such as the Rectified Linear Unit (ReLU). This makes it possible to learn intricate relationships between data, which are essential for precise forecasting.

d) Fully Connected Layers: Responsible for making predictions based on high-level features learned in previous layers, these layers establish connections between every neuron in one layer and every neuron in the next. They play a vital role in synthesizing extracted features for final output

Functionality: CNN is a deep learning model specialized in processing visual data.

Usage in Bank Cheque Verification: CNNs are employed for character recognition and image analysis tasks in cheque verification.

Methodological Approach: Trained CNN models are applied to cheque images to accurately identify handwritten digits, detect text blocks, and perform image classification for various verification purposes.

3.1.2 OCR (Optical Character Recognition):

OCR is a technology used to turn various document types into editable and searchable data. Examples of these documents include scanned paper documents, PDFs, and digital camera images. Text recognition and extraction are made possible by the system's analysis of the character shapes in the image and comparison with pre-existing letter and number patterns.

OCR technology is commonly utilized in a range of applications, including converting printed text into electronic formats for storage, automating data entry tasks, and enabling text-to-speech functionality for accessibility tools. Modern OCR systems often integrate machine learning methods to enhance precision, particularly with challenging fonts and handwritten text, therefore playing a crucial role in optimizing workflows that require processing significant amounts of textual information..

Functionality: OCR converts scanned text from bank cheques into machine-readable data.

Usage in Bank Cheque Verification: OCR enables automated extraction of essential information such as account numbers, cheque numbers, and legal amounts from scanned cheques.

Methodological Approach: Utilizing OCR algorithms, scanned images of bank cheques are processed to identify and extract textual information with high accuracy.

3.1.3 SIFT (Scale-Invariant Feature Transform):

In computer vision, the SIFT algorithm is used to identify and characterize local features in images. It is especially resilient to illumination, rotation, and scale changes. In order to facilitate efficient image matching and object recognition, SIFT functions by locating important points in the image and generating descriptors based on the surrounding pixel patterns.

Because of its resilience, SIFT is perfect for tasks like 3D modeling, object recognition, and image stitching. SIFT can match objects in different images despite differences in viewpoint or scale by concentrating on distinguishing key points. This is important for tasks that demand high accuracy in feature matching across a range of visual conditions.

Functionality: SIFT extracts distinctive features from images, allowing robust object recognition and matching.

Usage in Bank Cheque Verification: SIFT is employed for signature verification and identifying key elements within cheque images.

Methodological Approach: The SIFT algorithm is applied to scanned cheque images to extract unique features, aiding in the verification of signatures and other critical components.

3.1.4 SVM (Support Vector Machine):

A supervised machine learning algorithm for regression and classification applications is SVM. Maximizing the margin between classes is achieved by identifying the hyperplane in the feature space that best divides the various classes. SVM is frequently utilized in applications like text classification, image recognition, and bioinformatics because it performs well in high-dimensional spaces.

Regression: Using input variables as a basis, regression forecasts a continuous result. Finding a correlation between independent variables and a dependent variable—usually represented by a numerical value—is the aim of regression analysis. To build a model that, when presented with new input data, correctly predicts the value of the dependent variable is the goal.

Understanding the relationship between variables and making predictions based on continuous data are crucial in many fields, including science, economics, and finance, where regression techniques are widely used. As an illustration, a classic regression problem is estimating house prices based on features like size, location, and amenities. The model's goal is to estimate a numeric value (the price) based on input features.

Classification: Data must be categorized into predetermined classes or categories in order to be classified. The goal of classification tasks is to label input data according to its attributes or features. Classification deals with discrete outcomes as opposed to regression's continuous outcome. The objective is to create a model that can correctly categorize newly collected data into pre-established classes. When data points need to be divided into different categories, classification is frequently used in fields like image recognition, spam detection, and medical diagnosis. For example, the model assigns each incoming email to one of two predefined classes for email spam detection, based on features like keywords, sender information, and email content, and classifies the emails as spam or non-spam.

SVM is prized for its resilience and capacity to process non-linear data by means of kernel functions that raise the dimensions of the input data. This feature ensures high accuracy and generalization in a variety of predictive modeling tasks, making SVM a potent tool for complicated classification problems where the decision boundary is not obvious.

SVM encompasses two main types:

a) Linear SVM: This is utilized for linearly separable data, where classes can be distinguished using a single straight line. It's suitable for datasets that exhibit clear class division along a linear boundary.

b) Non-linear SVM: This is employed for data that isn't linearly separable, meaning a single straight line cannot adequately classify the dataset. Non-linear SVMs accommodate more complex datasets by utilizing non-linear boundaries, enabling effective classification in scenarios where linear separation is not feasible.

Functionality: SVM is a classification algorithm that identifies optimal hyperplanes to separate data into distinct categories.

Usage in Bank Cheque Verification: SVM is utilized for pattern recognition and classification tasks in cheque verification, such as signature authentication.

Methodological Approach: SVM classifiers are trained on extracted features from cheque images to distinguish between genuine and forged signatures, enhancing the security of the verification process.

3.2 Integration of Technologies:

Data Preprocessing: Before applying OCR, SIFT, SVM, and CNN, scanned cheque images undergo preprocessing steps such as noise removal and image enhancement to ensure optimal performance.

Feature Extraction: Each technology extracts specific features from cheque images, focusing on different aspects such as text recognition, signature verification, and object detection.

Classification and Verification: Extracted features are fed into SVM and CNN classifiers for pattern recognition and verification, enabling the system to validate the authenticity of cheque contents and detect any anomalies.

Automation and Efficiency: The integrated approach streamlines the bank cheque verification process, automating repetitive tasks and improving efficiency while maintaining high levels of accuracy and security.

Crucially, to uphold the efficacy of these methodologies, it becomes imperative to work with noise-free images, a feat accomplished through the meticulous implementation of image segmentation techniques. By selectively extracting pertinent information while discarding extraneous noise, these segmentation methodologies ensure the integrity and reliability of the data under scrutiny, thereby fortifying the overall authentication process of bank cheques.

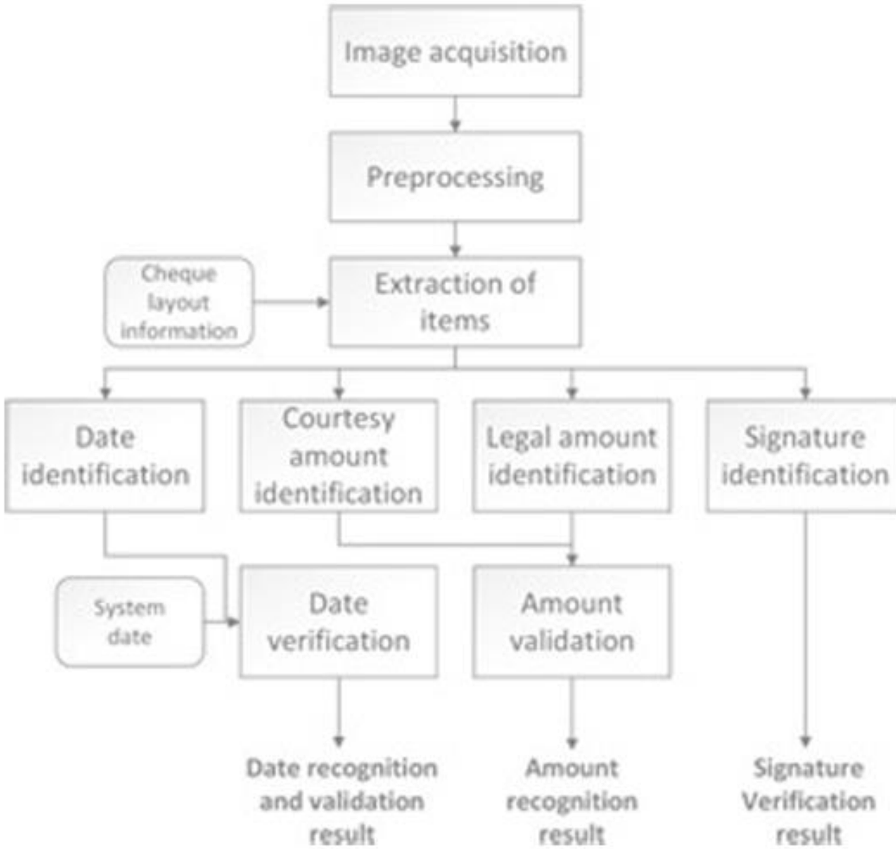

Fig. 3.2. Schematic depicting the intended work process

The flowchart depicts the process of automated cheque processing using image recognition technology. The process can be broken down into several steps. The first step is Image Acquisition, which involves capturing an image of the cheque. This is followed by Preprocessing, where the captured image undergoes enhancements to improve quality and accurately extract relevant data.

Next, the process moves to Extraction of Items, identifying and extracting key elements from the preprocessed image, such as the layout information of the cheque. This leads to the Identification of Cheque Components: detecting the date printed on the cheque (Date Identification), recognizing the numerical amount written on the cheque (Courtesy Amount Identification), extracting the amount written in words (Legal Amount Identification), and verifying the presence and authenticity of the signature (Signature Identification).

Following identification, the Verification Processes take place. This includes comparing the identified date against the system's current date to ensure the cheque's validity (Date Verification), ensuring that the numerical and written amounts match (Amount Validation), and confirming that the signature is valid and matches the authorized signature on file (Signature Verification).

The final step is Final Validation, which involves confirming that the date is correct and valid (Date Recognition and Validation Result), ensuring the amount is correct and matches across different formats (Amount Recognition Result), and validating the signature (Signature Verification Result). This automated process leverages image recognition technology to streamline the validation of cheques, reducing the risk of errors and fraud by accurately identifying and verifying critical data.

3.3 Procedure

In order to ensure accurate and efficient cheque verification and processing, a systematic procedure is followed, encompassing various essential steps. These steps are meticulously designed to uphold industry standards and regulatory requirements while enhancing the reliability of the authentication process.

3.3.1 Image Acquisition:

For precise processing and analysis, it is essential to obtain high-quality scanned images of the cheques. The authentication process's later steps are built upon these pictures. We were able to obtain a wide range of cheque images, representing different formats and styles that are frequently seen in banking transactions, by working together with the assigned source. Following industry best practices and guidelines, we implemented standardized scanning procedures to guarantee consistency and reliability. This required minimizing any potential distortions or artifacts that might impair image quality, choosing the right resolution settings, and maximizing lighting conditions. After the scanned images were received, we went through a rigorous pre-processing process to make them more suitable for additional analysis. In this first step, common issues that may come up during the scanning process are addressed, including skewness, noise, and background clutter. Our goal was to enhance the visual clarity and fidelity of the cheque images by utilizing sophisticated image processing techniques such as noise reduction, rotation, and background removal. Accurate feature extraction and segmentation in later stages of the authentication process are made possible by these preparatory steps.

3.3.2 Image Pre-processing:

The image pre-processing phase focuses on refining the scanned cheque images to ensure uniformity and clarity across the dataset. Central to this process is the correction of any distortions or irregularities that may impact subsequent analysis. We initiated this phase by identifying prominent visual cues within the cheque images, such as the date field, which serves as a reference point for rotation. By aligning the images based on this reference, we aimed to standardize their orientation and minimize potential alignment errors.

In order to maintain uniformity and clarity throughout the dataset, the pre-processing stage of the image processing process is essential. Adjusting for any distortions or irregularities that might influence further analysis is the main focus of this phase. As a solid starting point for rotation, the date field is one of the most conspicuous visual cues that we first locate within the cheque images. Standardizing the orientation of the images and reducing alignment errors during processing are made possible by aligning them based on this reference.

When images are processed before being used as an input or output, they are subjected to a number of basic operations, with intensity images serving as both the input and the output. The original sensor data, which is usually displayed as a matrix of brightness values, is preserved in these images in the same format. The main goal of pre-processing is to improve the quality of image data by correcting unwanted distortions and highlighting important aspects of the image that are necessary for further analysis. The fact that geometric transformations like rotation, scaling, and translation, although involving different techniques, are regarded as pre-processing techniques because they help prepare images for additional examination

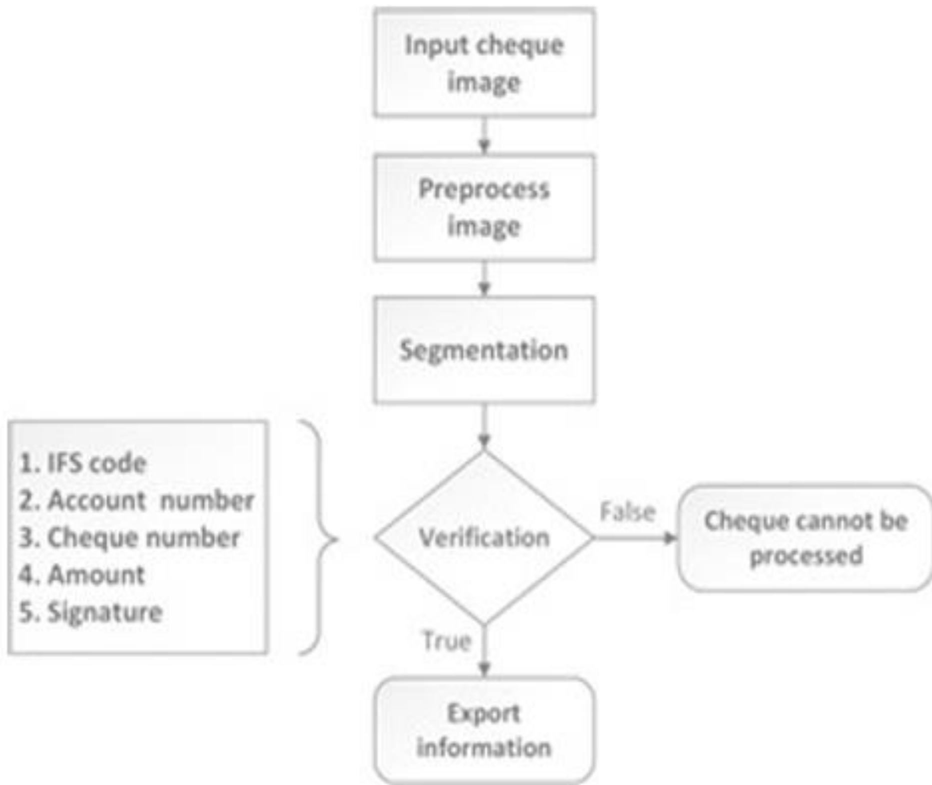

Fig. 3.3. A schematic showing the steps to verify a bank cheque

Subsequently, we employed advanced image processing algorithms to remove noise and irrelevant elements from the scanned images. This involved applying filters and morphological operations to enhance contrast, sharpen edges, and eliminate extraneous artifacts. By optimizing the visual fidelity of the cheque images, we sought to improve the accuracy of parameter identification and feature extraction in subsequent stages of analysis. Additionally, we ensured compliance with industry standards andregulatory requirements governing cheque processing and authentication.

3.3.3 Segmentation:

Image segmentation plays a pivotal role in isolating relevant regions of interest within the cheque images, facilitating targeted analysis and feature extraction. Leveraging state-of-the-art segmentation algorithms, we partitioned the scanned images into distinct components corresponding to critical data elements such as the account number, cheque number, and signature. This segmentation process involved the delineation of boundaries and contours within the images, guided by predefined templates and reference points.

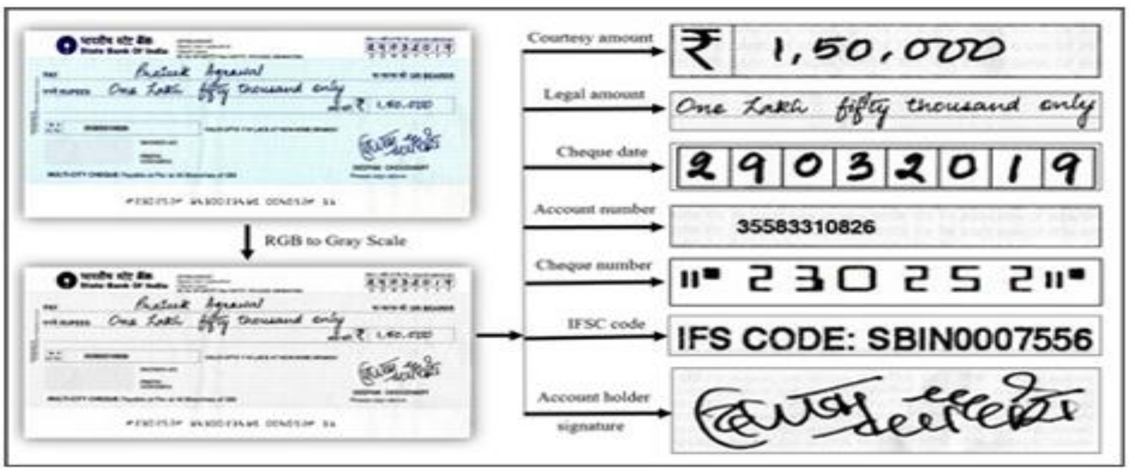

Fig. 3.4. Financing image segmentation sample

To enhance the accuracy and efficiency of segmentation, we integrated machine learning techniques such as transfer learning and support vector machines (SVM). These methods enabled the automatic identification and classification of handwritten digits, typographic letters, and signature features within the segmented regions. Additionally, operator intervention was occasionally required to validate segmentation results and address any deviations from standard dimensions or formats. Overall, segmentation served as a crucial preparatory step for subsequent stages of cheque authentication and verification.

3.3.4 CNN Legal and Courtesy Amount Identification Model:

The development of a specialized Convolutional Neural Network (CNN) represents a key component of our methodology for interpreting handwritten numbers representing the courtesy amount on cheques. CNNs are well-suited for image recognition tasks, particularly in scenarios involving complex visual data such as handwritten characters. Leveraging the Deep Learning Toolbox in MATLAB, we designed a CNN architecture comprising multiple convolutional layers and max pooling operations to extract meaningful features from the cheque images.

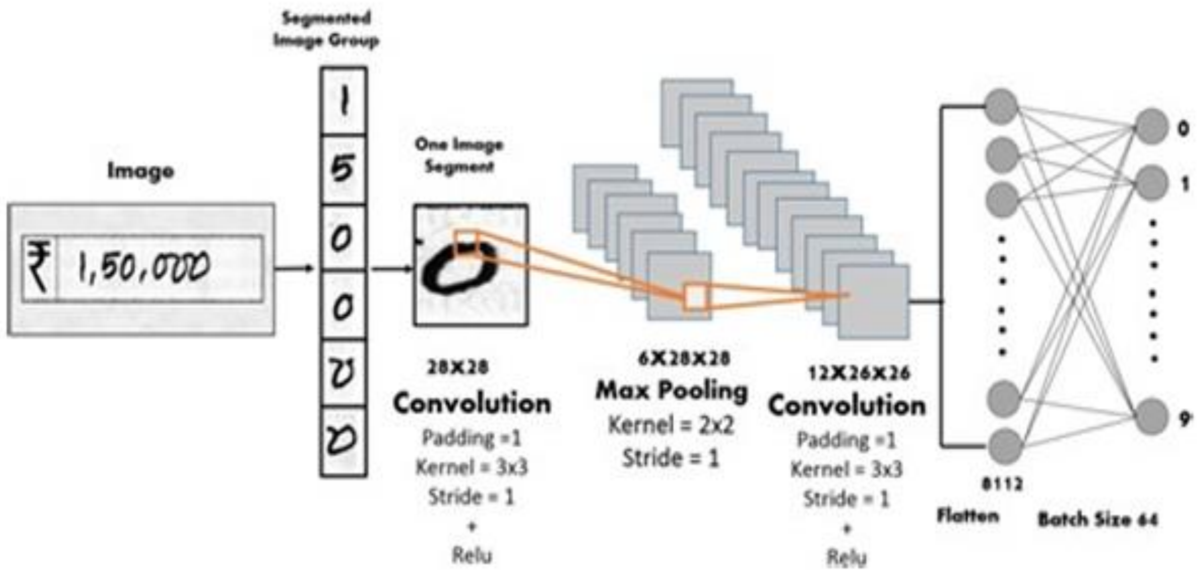

Fig. 3.5. CNN architecture proposed for recognizing handwritten digits in courteous amounts

Training and evaluation of the CNN model were conducted using established datasets such as MNIST for digit identification and EMNIST for English letter recognition. By leveraging transfer learning techniques, we aimed to adapt the CNN to recognize variations in handwritten number sequences commonly encountered in cheque processing. Furthermore, we implemented string verification mechanisms to ensure consistency between the legal amount extracted from the cheque and its textual representation. This comprehensive approach not only enhances the accuracy and reliability of cheque verification but also ensures compliance with regulatory standards and industry norms.

3.3.5 Signature Verification:

Signature authentication represents a critical aspect of cheque verification, requiring robust methodologies to detect and validate handwritten signatures accurately. To this end, we assembled a diverse dataset comprising offline handwritten signatures sourced from a broad spectrum of individuals. This dataset encompassed genuine signatures as well as simulated fraudulent signatures to facilitate rigorous testing and evaluation.

Fig. 3.6. Classification of account holder signatures for bank cheque verifi cation

Utilizing advanced feature extraction techniques such as Scale-Invariant Feature Transform (SIFT), we systematically extracted distinctive features and outlines from the signature images. These features were then input into Support Vector Machine (SVM) classifiers to distinguish between genuine and fraudulent signatures. By uniformly dividing the dataset for training and testing, we assessed the performance and reliability of the signature verification process across different scenarios and conditions.

Our methodology adheres to established criteria and guidelines governing cheque processing, ensuring compatibility with the Cheque Truncation System (CTS) standards for Indian banks. Additionally, our approach is adaptable to accommodate unique cheque dimensions and formats prevalent in international banking contexts, underscoring its versatility and applicability across diverse settings.

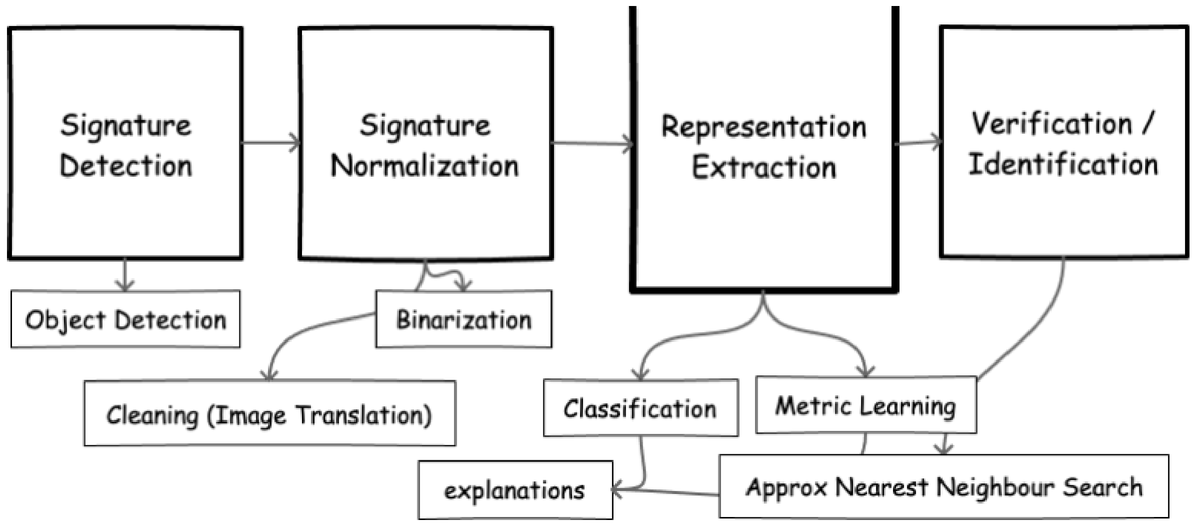

Fig 3.7 Diagrametic Representation of Signver module

The "signver" module consists of multiple interconnected sub-modules. In order to isolate the signature for clearer analysis, the first submodule, dubbed the "Cleaner," removes text and background lines from signature images. This submodule improves the quality of the signature images and sets the stage for precise verification by eliminating unnecessary components.

Once the signature images have been cleaned, the "Extractor" submodule is activated and turns them into vector representations. By standardizing the format, this conversion makes comparison and analysis more effective and guarantees consistency and dependability throughout the verification process. The Extractor submodule contributes to a more reliable and efficient matching process by extracting important features from the signature images.

Finally, by cross-referencing signatures with database records already in existence, the "Matcher" submodule is essential to verifying the legitimacy of signatures. This submodule compares the signature on the input cheque to the reference signature from the database using a distance measure. A match that satisfies predetermined standards counts as a verified signature; if not, it's marked as a fake or forgery. The "signver" module guarantees complete and accurate signature verification through the coordinated efforts of these sub-modules, improving the security and integrity of cheque processing.

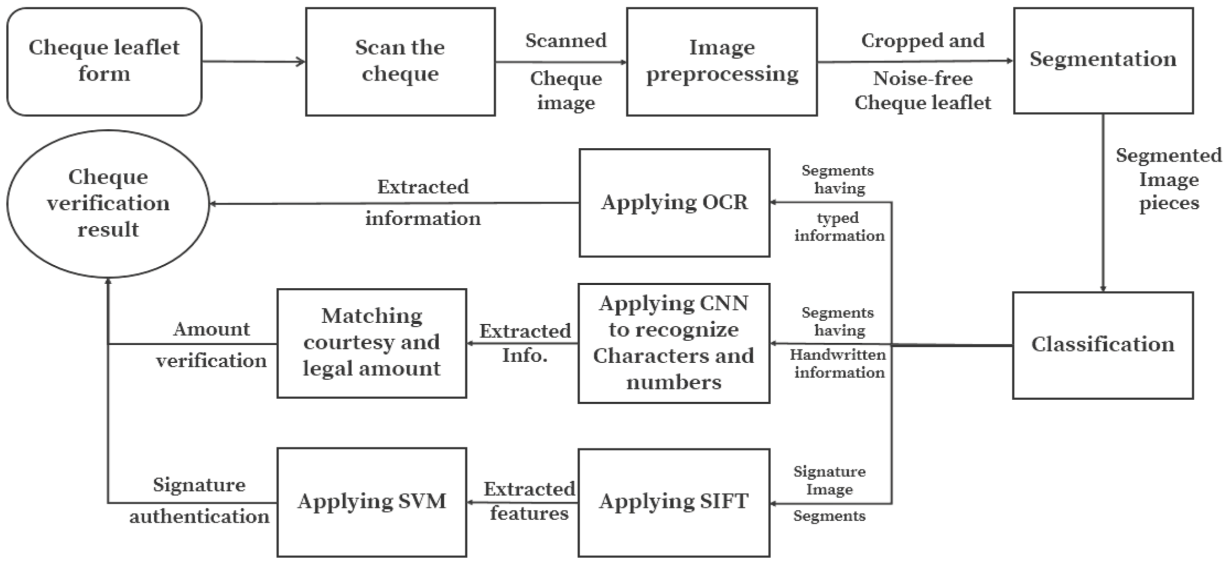

Fig. 3.8. Block diagram of bank cheque verification process

With each crucial step clearly defined to guarantee accuracy and efficiency, the flowchart provides a methodical process for verifying and processing cheques in a banking setting. It starts with scanning the actual cheque leaflet, which is the main input for processing that comes next. The flow splits off into various paths from there, each of which represents a different step in the verification procedure. Using different techniques to improve the quality of the scanned image of a cheque is image preprocessing, which is the first path. To create a standardized and refined image that is appropriate for further analysis, this involves steps to reduce noise, modify contrast, and enhance clarity.

Optical Character Recognition (OCR) is another method that is used concurrently to extract text information from the image of the cheque. Enabling automated verification and processing based on this data, OCR algorithms read machine-printed text, including the IFSC code and cheque number. In addition, the flowchart shows segmentation procedures designed to isolate particular areas of interest—like the signature area—within the cheque image. Dividing an object into segments makes it easier to analyze with focus and verify important elements precisely, like signature authenticity. As the process proceeds, the different routes come together to reach the crucial stage of signature authentication. Support Vector Machine (SVM) a is applied in this case.

CHAPTER 4

RESULTS AND DISCUSSION

In our study, we utilized MATLAB to investigate the performance and accuracy of a system designed for a specific issue related to bank cheque processing. We gathered a dataset comprising 114 bank cheque images, with 112 obtained from the Institute for Development and Research in Banking Technology (IDRBT) dataset and 2 independently sourced. These images were used for training and testing our system, focusing on the major parameter sections of the bank cheque booklet.

A key aspect of our system was the implementation of a Convolutional Neural Network (CNN) to identify handwritten digits present on the cheques. Our CNN exhibited remarkable performance, achieving a digit identification accuracy of 99.14%. Through iterative training over 850 iterations, we trained the character recognition networks, achieving a minimal mini-batch loss of 0.0077 and an impressive accuracy of 99.94%.

To convert English numerals to Indian place-value (IPV) system, we utilized courtesy photos. This conversion process ensured consistency and accuracy in representing numerical values, aligning with the standards required for bank cheque processing.

Following digit identification, our system proceeded to verify the signatory on the cheques and compared it to the approved sum. We employed a Support Vector Machine (SVM) classifier to detect and categorize patterns after standardizing the segmented images and extracting features using Scale-Invariant Feature Transform (SIFT) algorithm. The SVM classifier demonstrated high accuracy, achieving a validation accuracy of 98.1%.

Despite advancements in forensic technology, handwritten signatures remain a reliable tool for authenticity verification due to their unique writing styles. In our study, we explored both online and offline signature verification methods, each with its own set of advantages and limitations. Through rigorous analysis, we evaluated the performance of our system across various modules and achieved significant success.

Our study contributes to the field of automated bank cheque processing by demonstrating the effectiveness of machine learning algorithms, particularly CNNs and SVMs, in accurately identifying digits, verifying signatures, and categorizing patterns. The high accuracy rates obtained in our experiments underscore the potential of these technologies in streamlining the cheque processing workflow, reducing errors, and enhancing overall efficiency.

Fig.4.1. An actual scanned copy of the cheque, a cropped copy of the cheque booklet, a grayscale image of the cheque, an IFS code

Furthermore, our research highlights the importance of rigorous testing and evaluation methodologies to ensure the reliability and effectiveness of automated cheque processing systems. By thoroughly assessing the performance of each module and identifying areas for improvement, we can refine and optimize these systems for real-world applications.

In conclusion, our study demonstrates the feasibility and effectiveness of employing machine learning algorithms for automated bank cheque processing. The high accuracy rates achieved in digit identification, signature verification, and pattern categorization validate the efficacy of our approach. Moving forward, further research and development in this area could lead to the widespread adoption of automated cheque processing systems, offering numerous benefits to financial institutions and customers alike.

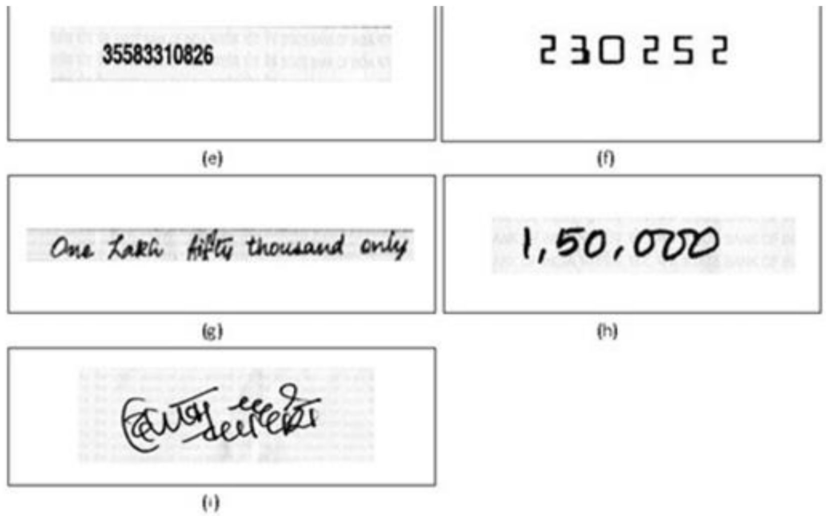

Fig 4.2 an account number, a cheque number, a legal amount, a courtesy amount, and a signature are the components that make up an SBI bank cheque.

CHAPTER 5

CONCLUSION AND FUTURE SCOPE

5.1 Conclusion

In conclusion, our complete bank cheque verification model, integrating Optical Character Recognition (OCR), Convolutional Neural Networks (CNN), Scale-Invariant Feature Transform (SIFT), and Support Vector Machines (SVM), represents a significant advancement in the automation of cheque clearance processes. By leveraging these cutting-edge technologies, we have achieved remarkable improvements in both efficiency and accuracy, offering numerous benefits to financial institutions and customers alike.

The utilization of OCR technology facilitated the extraction of data from machine-printed cheques with impressive accuracy, matching typographic characters at a rate of 97.7%. This machine-printed cheque data proved to be invaluable for subsequent processing stages. Our CNN model, trained with various databases, demonstrated exceptional performance in identifying handwritten numbers, achieving an accuracy rate of 99.14% after intensive training and testing. Additionally, CNNs significantly enhanced digit recognition, surpassing a recognition rate of 99.05%.

Recognizing the importance of handwritten digit detection for cheque processing, our CNN-based character recognition system excelled with a recognition accuracy of 99.94%. The successful identification of cheque characters by the CNN further improved the verification process. Furthermore, the combination of SIFT and SVM proved highly effective in identifying signatures, achieving a validation accuracy of 98.1% for cheque-clearing signatures. SIFT revealed critical elements of signature images, which were then accurately classified by SVM for verification.

The adoption of our automated cheque clearance system offers numerous advantages to banks, including time savings and simplification of central cheque transaction tracking. By streamlining the verification process and ensuring the accuracy of cheque clearance, financial institutions can enhance operational efficiency and customer satisfaction.

Looking ahead, our methodology holds promise for broader applications beyond English cheques. Future research endeavors could explore the extension of our approach to other languages, thereby expanding its utility and impact on global cheque processing systems.

In summary, our revolutionary bank cheque verification model, leveraging OCR, CNN, SIFT, and SVM technologies, represents a significant step forward in the automation of cheque clearance processes. By combining state-of-the-art machine learning algorithms, we have developed a robust and efficient system capable of accelerating cheque processing while maintaining high levels of accuracy and security.

5.2 Future Scope

The future scope for automatic bank cheque check detection is vast, driven by advancements in artificial intelligence, machine learning, and computer vision technologies. One significant area of development is the integration of more sophisticated image recognition algorithms, such as deep learning-based models, which can further enhance the accuracy of identifying and verifying handwritten elements and signatures on cheques. This improvement can lead to near-perfect detection rates, reducing the risk of fraud and errors.

Additionally, the implementation of blockchain technology for secure and immutable record-keeping could revolutionize cheque processing. By integrating automated cheque cheque detection systems with blockchain, financial institutions can ensure greater transparency and security in transactions. Moreover, advancements in real-time processing capabilities and cloud computing will enable the handling of large volumes of cheques efficiently, providing instant validation and clearing services. These innovations will contribute to a seamless, highly secure, and efficient banking experience for customers and institutions alike.

REFERENCES

[1] Dhanva, K., Harikrishnan, M., & Babu, P. U. (2018). Cheque image security enhancement in online banking. In Second International Conference on Inventive Communication and Computational Technologies (ICICCT) (pp. 1256–1260).

[2] Singh, H. K., & Tigga, A. E. (2012). Impact of information technology on Indian banking services. In 1st International Conference on Recent Advances in Information Technology (RAIT) (pp. 662–665).

[3] Wang, P., & Shiau, C. R. (1972). Machine recognition of printed Chinese characters via transformation algorithms. Pattern Recognition, 5(4), 303–321.

[4] Zramdini, A., & Ingold, R. (1998). Optical font recognition using typographical features. IEEE Transactions on Pattern Analysis and Machine Intelligence, 20(8), 877–882.

[5] Zhu, Y., Tan, T., & Wang, Y. (2002). Font recognition based on global texture analysis. IEEE Transactions on Pattern Analysis and Machine Intelligence, 23(20), 1192–1200.

[6] Saenthon, A., & Sukkhadamrongrak, N. (2014). Comparison of training methods of neural network for English and Thai character recognition. In Signal and Information Processing Association Annual Summit and Conference (APSIPA) (pp. 1–4).

[7] Ramanathan, T. T., & Sharma, D. (2017). Multiple classifications using SVM based multi-knowledge-based system. In International Conference on Advances in Computing & Communications (ICACC'17) (pp. 307–311).

[8] Stewart, S., Pinto, L., & Barrett, B. (2018). Segmentation and stitching improve handwriting recognition on datasets with few samples. In 16th International Conference on Frontiers in Handwriting Recognition (ICFHR) (pp. 465–470).

[9] Kajale, R., Das, S., & Medhekar, P. (2017). Supervised machine learning in intelligent character recognition of handwritten and printed nameplate. In International Conference on Advances in Computing, Communication and Control (ICAC3) (pp. 1–5).

[10] Feng, W., & Gao, S. (2010). A vehicle license plate recognition algorithm in the night based on HSV. In 3rd International Conference on Advanced Computer Theory and Engineering (ICACTE) (Vol. 4, pp. 53–56).

[11] Singh, S., Kariveda, T., Gupta, J. D., & Bhattacharya, K. (2015). Handwritten words recognition for legal amounts of bank cheques in English script. In Eighth IEEE International Conference on Advances in Pattern Recognition (ICAPR) (pp. 1–5).

[12] Miah, M. B. A., Yousuf, M. A., Mia, M. S., & Miya, M. P. (2015). Handwritten courtesy amount and signature recognition on bank cheque using neural network. International Journal of Computer Applications, 118(5), 21–26.

[13] Jayadevan, R., Pal, U., & Kimura, F. (2010). Recognition of words from legal amounts of Indian bank cheques. In 12th International Conference on Frontiers in Handwriting Recognition. IEEE, pp. 166–171.

[14] Raghavendra, S. P., & Danti, A. (2009). A novel recognition of Indian bank cheques based on invariant geometrical features. In International Conference on Trends in Automation, Communications and Computing Technology (I-TACT'15) (pp. 1–5).

[15] Shirai, K., Akita, M., Okamoto, M., Tanikawa, K., Akiyama, T., & Sakaguchi, T. (2012). Removal of background patterns and signatures for magnetic ink character recognition of cheques. In 10th IAPR International Workshop on Document Analysis Systems (pp. 190–194).

[16] Lu, H., Guo, B., Liu, J., & Yan, X. (2017). A shadow removal method for Tesseract text recognition. In International Congress on Image and Signal Processing Biomedical Engineering and Informatics (CISP-BMEI) (pp. 1–5).

[17] Wiraatmaja, C., Gunadi, K., & Sandjaja, I. N. (2017). The application of deep convolutional denoising autoencoder for optical character recognition preprocessing. In

International Conference on Soft Computing Intelligent System and Information Technology (ICSIIT) (pp. 72–77).

[18] Aronoff, J., Simske, S., & Sturgill, M. (2011). Automated optimization of void pantograph settings. HP Laboratories, 690–693.

[19] Shenoy, S. S., Nayak, S. U., & Neelima, B. (2016). SWIFT—a performance accelerated optimized string matching algorithm for Nvidia GPUs. In International Symposium on Parallel and Distributed Computing (pp. 80–87).

[20] Joshi, O. S., Upadhvay, B. R., & Musica, S. (2017). Parallelized advanced Rabin-Karp algorithm for string matching. In International Conference on Computing, Communication, Control, and Automation (ICCUBEA) (pp. 1–5).

APPENDIX 1

Verifying bank checks using deep learning and image processing

Yogesh Kumar Singh
KIET Group of Institution
Delhi-NCR,Ghaziabad
yogesh.2024cse1151@kiet.edu

Rishabh Jaiswal
KIET Group of Institution
Delhi-NCR,Ghaziabad
rishabh.2024cse1145@kiet.edu

Prerna Choudhary
KIET Group of Institution
Delhi-NCR,Ghaziabad
choudharyprerna493@gmail.com

Bharti Chugh
KIET Group of Institution
Delhi-NCR,Ghaziabad
bharti.cse@kiet.edu

***Abstract*—Our innovative approach completely revamps the process of verifying bank checks by leveraging deep learning and image processing. To make the cheque truncation mechanism far more efficient, we reduce the amount of human interaction. Our approach efficiently retrieves important details from the check booklet, including the bank branch code, check number, precise amount, account number, and unique signature patterns.**

We utilized the IDRBT cheque dataset and convolutional neural networks based on deep learning for our study. This enhanced assessment of handwritten components on bank cheques achieved 99.14 percent accuracy in identifying handwritten numeric characters.

We achieved an impressive 97.7 percent accuracy while utilizing MATLAB's integrated optical character recognition (OCR) technique to interpret machine-printed text. Utilizing Scale Invariant Feature Transform (SIFT) for feature extraction and Support Vector Machine (SVM) for classification resulted in a notable 98.10 percent accuracy in signature verification.

***Index Terms*—Image processing,Deep learning,Signature patterns,Convolutional neural networks,OCR (Optical Character Recognition)**

I. Introduction

However, even with the growing popularity of digital transactions, bank checks continue to play a vital role in facilitating financial transactions on a global scale. This remains valid even as digital transactions are becoming more and more widespread. Nevertheless, the manual processing of these tests by people is not only laborious but also expensive, prone to errors, and time-consuming. Furthermore, errors may also occur. Moreover, there are inherent risks associated with this. Moreover, it increases the likelihood of detecting and correcting errors. Hence, the advent of automated bank cheque processing systems has spurred scholars to investigate this potentially promising field of study. This can be attributed to one of the variables stated above.

Several disciplines are essential to the automation of bank check processing, including computer vision, image processing, pattern recognition, machine learning, and deep learning. Gathering images, processing them, extracting them, and finally recognizing them are all steps in the process. Important details like bank branch codes, check numbers, exact amounts, account numbers, and signature patterns may be found and verified using a combination of image processing and deep learning techniques. Financial transaction fraud may be easily detected by machine learning algorithms. By examining various features collected from scanned pictures of checks, modern algorithms such as the Generative Adversarial Network (GAN) may potentially identify possible counterfeit products in real time. Research on truncation methods has also made use of machine learning and artificial intelligence.

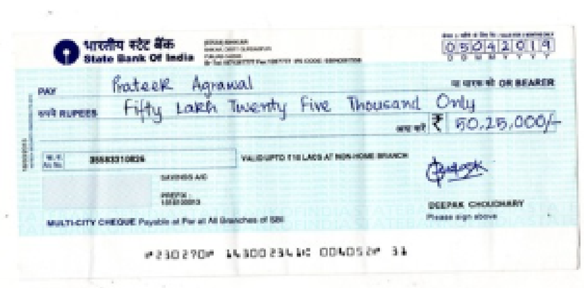

Fig. 1. CTS-2010 compliant sample bank cheque

Checks are confirmed to have cleared clearance by the Cheque Truncation System (CTS) after a two-step verification process. Following the advice of Deepak and colleagues (2010), the first thing to do is make sure the check booklet is real. Next, the information in the brochure is examined. Some of the innovative approaches to automated bank check verification showcased in this article include extracting IFSC codes, check numbers, account numbers, legal amounts, courtesy amounts, and the signatures of the check issuer. The retrieved characteristics are displayed using bounding boxes for validation purposes. Various researchers[1,2] have studied approaches for identifying, authenticating, and verifying bank cheques, but their effectiveness is low. OCR automates the process of recognizing text in images, including identifying machine-printed or handwritten characters and numbers. Image processing involves analyzing digitized images and doing augmentation, compression, segmentation, and editing. Image processing includes acquisition, preprocessing, segmentation, interpretation, and recognition. CNNs with multiple layers achieve high accuracy in recognizing patterns in images. Convolutional Neural Networks require a large dataset for efficient training. These techniques facilitate precise algorithmic execution, reducing the time and resources required for banks' cheque clearance procedures. The investigation using MATLAB-2018a. This research primarily focuses on discovering bank cheque verification techniques and applying image processing and deep learning technologies to enhance the precision of automated verification. Summarizing the primary contributions: Segmenting scanned bank check images to extract specific information such as date, account number, cheque number, account holder signature, courtesy amount, legal amount, and bank information (IFS Code). Detecting specific numerical values with a Convolutional Neural Network (CNN) and validating them against the converted legal amount data using the proposed method. This research demonstrated a more efficient and curated bank cheque verification approach compared to others. The work proposes an image processing and deep learning approach to automate bank cheque verification. The proposed model improves precision and effectiveness by extracting crucial data from check leaflets with OCR, CNN, SIFT, and SVM technologies. Banking benefits from research that accelerates processing and automates check clearance. An essential component of this study is its potential to enhance the efficiency of financial transactions, resulting in time and cost savings, as well as reduced risk of cheque fraud. This study provides a valuable resource for future research on this topic by addressing the persistent challenges currently encountered by the sector and proposing a viable research methodology. It provides a comprehensive comparison of many models and their unique accuracy rates in recognizing payee names, dates, digit identification, and signature detection.

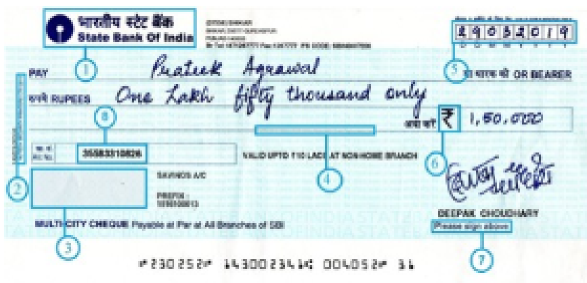

Fig. 2. Scan of a sample check in its raw form

II. LITERATURE REVIEW

Visual input requires character recognition to understand patterns. Wankhede and Mohod (1983), Singh and Sachan (1977), and Ramanathan et al. (1970) demonstrated picture character OCR. Division of digits, a major character recognition challenge, has been solved in several ways. Chen created spectral clustering to segment handwritten digits 97% accurately. Using global typographical features, other researchers [3,4] devised a statistical technique for text line Latin Language font recognition The recognition rate was 97%. Researchers [5,6] created popular and cost-effective English character identification systems. Multiple methods correctly detected text block fonts. With 93.54% accuracy, [7] used SVMs for multi-knowledge categorization and feature extraction.

Image processing still favors OCR or segmentation for character recognition. With small training datasets, researchers [8] demonstrated handwriting recognition without segmentation. Manual or automatic segmentation and sewing. A 2018 Peres et al. hybrid CNN-SVM classifier had 96.7 percent MNIST accuracy. Researchers [9] detected alpha digits and special characters with over 95% accuracy using supervised machine learning.

Key feature	Description
IFS code	–Indian financial system code. –It is an eleven-character code where first four characters represent bank name, and last six characters represents the branch. The fifth character is kept reserved for future use.
Account number	–A unique number assigned to the user of the bank services, used for monetary transaction related services.
Cheque number	–A six-digit number on the bottom left corner of a cheque, it helpsto identify a cheque and verify if the cheque is from the list of issued cheques to the user.
Legal amount	–The amount mentioned on the cheque leaflet in handwritten / printed words.
Courtesy amount	–The amount mentioned on the cheque leaflet in handwritten / printed numeric digits.
Signature	–Signature of the account holder. –Every person has a unique signature which depends on the static or dynamic features, and it is hard to copy it. –It is used as a key feature in authenticating a user's identity.

Fig. 3. Description of the main characteristics used for bank check verification

Character text formatting vs. dynamic templates. Lines, schematics, watermarks, and patterns influence character recognition differently. From colored or noisy backdrops, researchers [10] collected data. Changing backdrop color or adding intricate patterns and numbers won't slow productivity. Jia and Wei (2014) offered binarization using global and local criteria to standardize Chinese bank check amounts. With a correlation coefficient, researchers [11] estimated the maximum Indian bank check amount. Word recognition accuracy was 76.4% with 19,215 words and 61 keywords. Few researchers [12] found that ANNs identified handwritten Bangladesh bank check amounts 93.4% properly. Jayadevan et al. [13] successfully detected authorized monetary amounts on English bank checks with 97.20% accuracy using the Modified Quadratic Discriminant Function (MQDF). Functional lexicon segmentation. Few researchers [14]say removing background noise helps uncover cheque bank data. Researchers used invariant geometrical properties to find 93.8% accurate checks. Secure cheque verification with MICR codes. Bank checks require MIR, but the pattern and legal amount detection can be done prior. Some researchers [15]. developed three multimedia tools and application methodologies. Character color, backdrop, pitch, noise, stroke tracing, and sign fragment elimination are used to extract MICR codes.

Few researchers [16] addressed shadows and spatial noise in low-quality copied photos. To capture text contours, researchers converted the gray-scale image to binary using a local adaptive threshold. Dual filtering eliminated text-character interference. Few researchers [17] found that deep convolutional de-noising auto-encoders beat OCR systems by 26.78%. Authenticating check images requires a Void Pantograph. Initial prints featured an invisible pantograph. It stands out when printed, therefore security printing toolkits need it. Aronoff et al. automate void pantograph parameter tuning. Some researchers [18] automated void pantograph parameter optimization.

Numerous studies have compared handwritten and machine-generated characters across languages. Arabic handwritten and machine-generated characters were categorized with 68% accuracy in an experiment. Image processing and computer vision were employed for object and edge detection in other investigations. KNN and SVM classifiers achieved 99% accuracy in offline handwritten digit recognition. CNNs identified scratchy and non-scratchy handwriting. Automatic grab-cut image segmentation. DNNs classify real-world data. We employed watershed and distance transform-enhanced cell image segmentation. Performance-based string matching on

Fig. 4. Images of sample checks from various nations

Nvidia GPUs using Quick Search, Horspool, and Brute Force algorithms was developed by a few researchers[19,20] and offered Rabin- Karp parallelization.

Researchers enhanced multi-pattern matching and cursive English handwriting recognition. Two-hash multi-pattern matching was accelerated by one approach. Reducing comparisons sped up string-matching, saving time and space. Another study used vertical and horizontal projection, convex hull, and SVM to distinguish cursive English handwriting. The computers correctly identified CCC cursive handwriting patterns. Researchers discovered a way to identify handwritten number characters for document processing and digit recognition.

Gestalt-based perimetric complexity and affine transformation assessed Odia numeral shape context for point-to-point comparison. Logos on leaflets identify bank cheque issuers. Each picture segmentation, character recognition, and signature identification investigation performed well. Lack of bank cheque validation research and no automated cheque handling system. Scannable check images are formatted and amounts and signatures verified. Next, we demonstrate model implementation and success.

The task was done utilizing several approaches. Researchers used OCR in MATLAB to obtain 97.7% accuracy for machine-printed scripts. SIFT feature extraction and SVM signature verification classifiers achieved 98.10 percent accuracy. These two methods dramatically enhanced automated check processing and reconciliation. This paper's extensive research advances these fields.

III. Methodology

The bank cheque authentication process necessitates the completion of several crucial elements in a particular sequence. The system authenticates the IFSC code on the cheque and then verifies the cheque number to ensure it is from the account holder's designated set of cheque booklets.

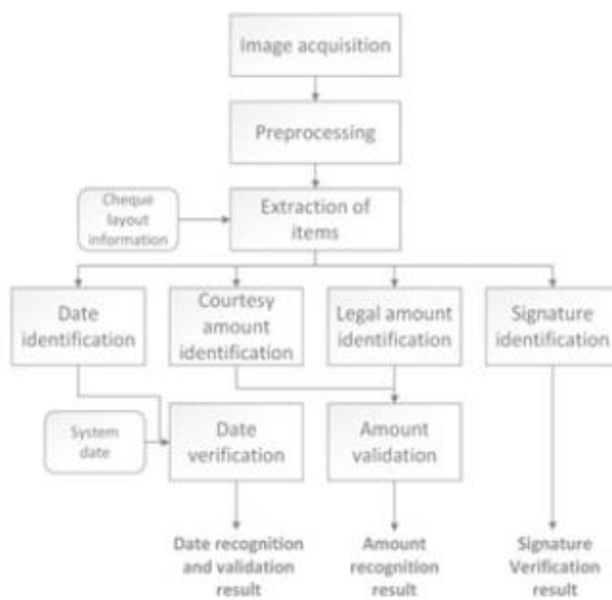

Fig. 5. Schematic depicting the intended work process

The next stage is verifying the check's issuer's signature(s) and amount against the customer's account balance. All check-clearing steps, including withdrawals and transfers, depend on your fast and accurate completion of these essential verifications. The check clearance process has interdependent and crucial steps. Several accurate and effective methods retrieve data from the check leaflet. Because of its precision, optical character recognition (OCR) extracts information from machine-printed text. Meanwhile, deep learning-based CNN handles numerical data and handwritten text.

SIFT extracts features to authenticate signatures, while SVM classifies them for better performance. These approaches require a noise-free image, so image segmentation extracts only relevant information.

A. *Image acquisition*

Pre-processing procedures were necessary for the scanned image(s) before they could be used directly in image-processing activities. The operations were conducted to ready the image(s) for additional processing.

B. *Image pre-processing*

For our research, we used a scanned copy of the cheque. The scanned images couldn't be utilized without first undergoing preprocessing. To implement this method, you must first rotate the image and then remove any extraneous background data. The "Date Box," a standard component of all bank checks, was employed to rotate the scanned picture. We then removed noise and irrelevant data, which greatly improved the accuracy of parameter identification. These preprocessing steps are critical to the reliability and validity of the verification as a whole.

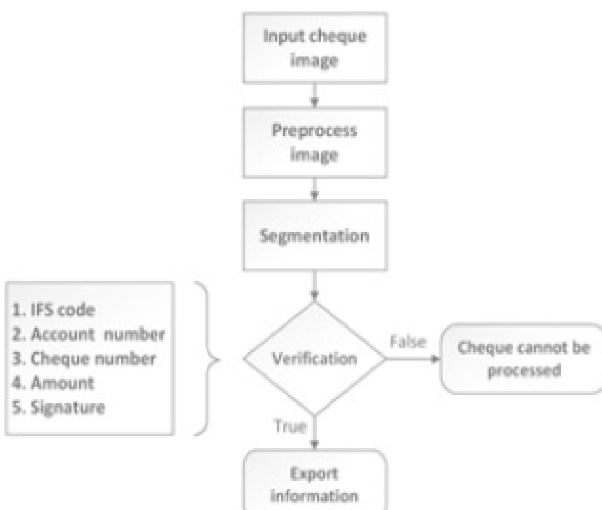

Fig. 6. A schematic showing the steps to verify a bank check

- The scanned photographs were rotated using the date field typically found on standard check booklets. Contour extraction was used to identify the position of the date box, serving as a reference point for rotation.
- The rotation point was set at the center of the image, and the rotation angle was based on the quadrant where the date box was located.
- Scanned photo noise was removed by evaluating the date box dimensions and applying length mapping to determine the range of the check.

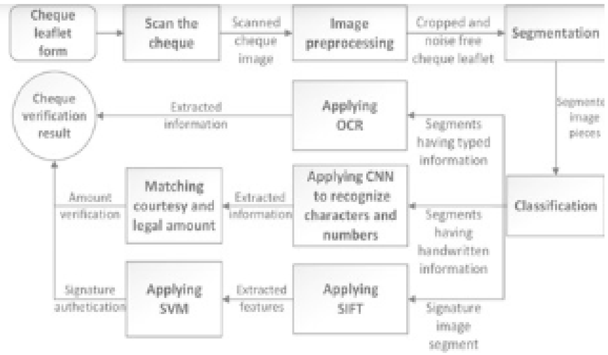

Fig. 7. Block diagram of bank cheque verification process

- The RGB image was transformed into a monochromatic grayscale image using a designated formula.
- A 2D Gaussian function with a specified standard deviation was utilized for Gaussian filtering to remove noise from the grayscale image. Following morphological procedures, erosion, and dilation were performed to improve the picture.

The formula for 2-D Gaussian function is

$$G(x, y) = \frac{1}{2\pi\sigma^2} \times e^{-\frac{x^2+y^2}{2\sigma^2}} \tag{1}$$

In this case, the typical arrangement of the values is denoted by $G(x, y)$. In contrast, the arrangement's standard deviation is denoted by σ.

To accurately and precisely extract contours from the grayscale image, we utilized the binary image format. Put simply, we used the algorithm stated below to find the region's border points after converting the picture to binary representation :

$$\begin{cases} f(i, j-1) = 0 \\ f(i, j) = 1 \\ f(i, j) \geq 1 \\ f(i, j+1) = 0 \end{cases} \tag{2}$$

The binary image function $f(i, j)$ gives the pixel value of the needed point in this equation, whereas the functions $f(i, j-1)$ and $f(i, j+1)$ give the pixel values of the nearby points for$f(i, j)$.

C. *Segmentation*

Image segmentation separated critical data and identified cheque document sections of interest. This allowed access to the full procedure while processing only necessary data. Contour extraction and standard templates are segmented according to RBI proportions. Transfer learning and OCR identified handwritten digits and typographic letters. SIFT retrieved features and SVM classified them for signature verification. Operator intervention was needed to check size or national or bank standard deviations during semi-automatic segmentation.

Fig. 8. Financing image segmentation sample

D. CNN legal and courtesy amount identification model

We employed a CNN to read the courtesy amount's handwritten numbers and convert them to strings for our study. MATLAB's Deep Learning Toolbox created the CNN with two convolution layers and max pooling.

MNIST was utilized for digit identification training and evaluation, while EMNIST was employed for English letter recognition. 80% of these datasets were utilized for training, and 20% for testing.

We also sorted numerical digits using the Indian Place Value (IPV) system after a CNN detected letters and integers. We used transfer learning to handle handwritten number sequence variations. We converted the number to text so you could compare it to the allowed amount.

Both string values were verified by a dictionary function. Same-string verification was done. Comparing the legal number to the textual amount ensured the check was accurate and legitimate. Remember that this comprehensive technique follows CTS-2010 check dimensions for Indian banks. It may also process international checks with unique dimensions required by different countries or banks. It uses standardized check measurements and meets international check regulations, making it semi-automatic.

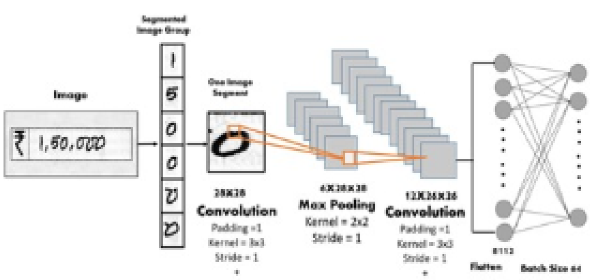

Fig. 9. CNN architecture proposed for recognizing handwritten digits in courteous amounts

E. Signature verification

Our study focused on crucial signature authentication. We examined the critical procedure of signature authentication. We collected 110 offline handwritten signatures from scanned cheques from 50 diverse individuals. There were 80 real and 30 phony.

We extracted attributes from each signature image portion and standardized them. Normalization helped remove scaling, thinning, rotating, and cropping features. We extracted features and outlines from normalized photos using Scale-Invariant Feature Transform (SIFT).

SIFT's usefulness depends on accurately recreating visual areas at key points while keeping their size and orientation. SSE creates scale-space extrema, KPL localizes key points, OA assigns orientation, and KDE extracts key points. Four phases comprise the method.

SVM classifiers with SIFT-derived features authenticated signatures. Genuine and fraudulent signature data were uniformly divided 80:20 throughout training and testing datasets. An SVM classifies data into two portions using a separating hyperplane in two dimensions. This method creates optimal hyperplane-grouped data points for categorization.

This thorough procedure ensures check authenticity and precision, producing a robust check verification and processing system. It follows CTS-2010 criteria for Indian banks and unique dimensions set by individual countries or banks for cheques from other countries, making this technique semi-automatic. This technique ensures honest and empathetic verification.

Fig. 10. Classification of account holder signatures for bank cheque verification

F. Findings and examination

MATLAB was used to examine performance and accuracy for a specific issue. 114 bank cheque images were scanned, 112 from the IDRBT dataset, and 2 independently. Our system was trained and tested utilizing the bank cheque booklet's major parameter sections.

A Convolutional Neural Network identified handwritten digits 99.14 percent of the time. Over 850 iterations, we trained character recognition networks with 0.0077 mini-batch loss and 99.94% accuracy. We converted English-to-Indian place-value (IPV) using courtesy photos. After verifying the signatory, we compared it to the approved sum. We used a Support Vector Machine (SVM) classifier to discover and categorize patterns after standardizing the segmented picture and extracting features using SIFT. The Support Vector Machine (SVM) gives us 98.1 percent accuracy. Despite advances in forensics, handwritten signatures are still a reliable tool for authenticity verification due to their unique writing style. Online and offline signature verification have perks and cons. We examined system performance per module accuracy with advanced methods and achieved success.

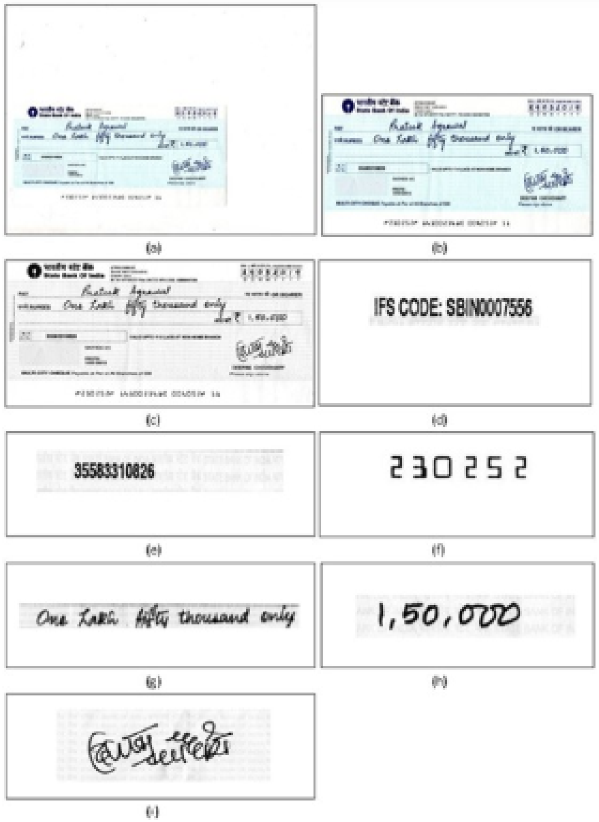

Fig. 11. An actual scanned copy of the check, a cropped copy of the cheque booklet, a grayscale image of the cheque, an IFS code, an account number, a cheque number, a legal amount, a courtesy amount, and a signature are the components that make up an SBI bank cheque.

SUMMARY AND PROSPECTIVE RESEARCH

Complete bank cheque verification model incorporating OCR, CNN, SIFT, and SVM. Cheque clearance automation improves efficiency and accuracy. Machine-printed checks were OCR'd. The OCR matched typographic characters 97.7%. Machine-printed check data was useful.CNN model with different databases found handwritten numbers. Intensive training and testing improved accuracy to 99.14 percent. CNNs improved digit recognition from 99.05 percent. Handwriting digit detection is required for checks. Our CNN character recognition was 99.94%. CNN correctly spotted cheque characters, boosting verification.SIFT and SVM found signatures. 98.1% of cheque-clearing signatures were right. SIFT uncovered key signature picture elements that SVM detected for accurate signature verification. Banks gain from our method. Automating cheque clearance saves time. Central check transaction tracking simplifies and verifies. Simply learning English checks. Future research could use our method for more languages, expanding its utility. Our revolutionary bank cheque verification uses OCR, CNN, SIFT, and SVM. To speed bank cheque clearance, the system recognizes machine-printed and handwritten data and confirms signatures.

REFERENCES

[1] Dhanva K, Harikrishnan M, Babu PU (2018) Cheque image security enhancement in online banking. In: Second international conference on inventive communication and computational technologies (ICICCT), pp 1256–1260

[2] Singh HK, Tigga AE (2012) Impact of information technology on indian banking services. In: 1st International conference on recent advances in information technology (RAIT), pp 662–665

[3] Wang P, Shiau CR (1972) Machine recognition of printed Chinese characters via transformation algorithms. Pattern Recognit 5(4):303–321

[4] Zramdini A, Ingold R (1998) Optical font recognition using typographical features. IEEE Trans Pattern Anal Mach Intell 20(8):877–882

[5] Zhu Y, Tan T, Wang Y (2002) Font recognition based on global texture analysis. IEEE Trans Pattern Anal Mach Intell 23(20):1192–1200

[6] Saenthon A, Sukkhadamrongrak N (2014) Comparison the training methods of neural network for English and Thai character recognition. In: Signal and Information Processing Association Annual Summit and conference (APSIPA), pp 1–4

[7] Ramanathan TT, Sharma D (2017) Multiple classifications using svm based multi knowledge-based system. In: International conference on advances in computing & communications, (ICACC'17), pp 307– 311.

[8] Stewart S, Pinto L, Barrett B (2018) Segmentation and stitching improves handwriting recognition on datasets with few samples. In: 16th International Conference on Frontiers in Handwriting Recognition (ICFHR), pp 465–470

[9] Kajale R, Das S, Medhekar P (2017) Supervised machine learning in intelligent character recognition of handwritten and printed nameplate. In: International conference on advances in computing, communication and control (ICAC3), pp 1–5

[10] Feng W, Gao S (2010) A vehicle license plate recognition algorithm in night based on hsv. In: 3rd International Conference on Advanced Computer Theory and Engineering (ICACTE), vol 4, pp 53–56

[11] Singh S, Kariveda T, Gupta JD, Bhattacharya K (2015) Handwritten words recognition for legal amounts of bank cheques in English script. In: Eighth IEEE International Conference on advances in Pattern Recognition (ICAPR), pp 1–5

[12] Miah MBA, Yousuf MA, Mia MS, Miya MP (2015) Handwritten courtesy amount and signature recognition on bank cheque using neural network. Int J Comput Appl 118(5):21–26

[13] Jayadevan R, Pal U, Kimura F (2010) Recognition of words from legal amounts of Indian bank cheques. In: 12th International Conference on frontiers in Handwriting recognition. IEEE, pp 166–171

[14] Raghavendra SP, Danti A (2009) A novel recognition of Indian bank cheques based on invariant geometrical features. In: International conference on trends in automation, communications and computing technology (I-TACT'15), pp 1–5

[15] Shirai K, Akita M, Okamoto M, Tanikawa K, Akiyama T, Sakaguchi T (2012) Removal of background patterns and signatures for magnetic ink character recognition of checks. In: 10th IAPR international workshop on document analysis systems, pp 190–194

[16] Lu H, Guo B, Liu J, Yan X (2017) A shadow removal method for tesseract text recognition. In: International Congress on image and Signal Processing Biomedical Engineering and Informatics (CISP-BMEI), pp 1–5

[17] Wiraatmaja C, Gunadi K, Sandjaja IN (2017) The application of deep convolutional denoising auto coder for optical character recognition preprocessing. In: International conference on soft computing intelligent system and information technology (ICSIIT), pp 72–77

[18] Aronoff J, Simske S, Sturgill M (2011) Automated optimization of void pantograph settings. HP Laboratories 690–693

[19] Shenoy SS, Nayak SU, Neelima B (2016) Swift—a performance accelerated optimized string matching algorithm for Nvidia gpus. In: International symposium on parallel and distributed computing, pp 80–87

[20] Joshi OS, Upadhvay BR, Musica S (2017) Parallelized advanced rabin-karp algorithm for string matching. In: International conference on computing, communication, control, and automation (ICCUBEA), pp 1–5

Plagrism Report

Similarity Report

PAPER NAME

Project_report _PCSE24-39.pdf

WORD COUNT	CHARACTER COUNT
7947 Words	49784 Characters
PAGE COUNT	FILE SIZE
47 Pages	2.0MB
SUBMISSION DATE	REPORT DATE
May 27, 2024 2:03 PM GMT+5:30	May 27, 2024 2:04 PM GMT+5:30

16% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.

- 11% Internet database
- 5% Publications database
- Crossref database
- Crossref Posted Content database
- 12% Submitted Works database

Excluded from Similarity Report

- Bibliographic material
- Quoted material
- Cited material
- Small Matches (Less then 8 words)

16% Overall Similarity

Top sources found in the following databases:

- 11% Internet database
- 5% Publications database
- Crossref database
- Crossref Posted Content database
- 12% Submitted Works database

TOP SOURCES

The sources with the highest number of matches within the submission. Overlapping sources will not be displayed.

1	KIET Group of Institutions, Ghaziabad on 2024-04-19 Submitted works	3%
2	KIET Group of Institutions, Ghaziabad on 2024-05-24 Submitted works	2%
3	HTM (Haridus- ja Teadusministeerium) on 2023-12-25 Submitted works	<1%
4	link.springer.com Internet	<1%
5	Prateek Agrawal, Deepak Chaudhary, Vishu Madaan, Anatoliy Zabrovsk... Crossref	<1%
6	Meerut Institute of Engineering & Technology on 2024-05-22 Submitted works	<1%
7	KIET Group of Institutions, Ghaziabad on 2024-05-15 Submitted works	<1%
8	assets.researchsquare.com Internet	<1%

9	Liverpool John Moores University on 2023-08-24 Submitted works	<1%
10	Westcliff University on 2023-12-20 Submitted works	<1%
11	University of North Texas on 2023-10-13 Submitted works	<1%
12	Asia Pacific University College of Technology and Innovation (UCTI) on... Submitted works	<1%
13	Berlin School of Business and Innovation on 2023-10-02 Submitted works	<1%
14	Signal Mountain Middle High School on 2023-12-07 Submitted works	<1%
15	"Recent Trends in Image Processing and Pattern Recognition", Springe... Crossref	<1%
16	Liverpool John Moores University on 2024-03-18 Submitted works	<1%
17	Mahidol University on 2009-09-29 Submitted works	<1%
18	University of Illinois at Urbana-Champaign on 2024-05-08 Submitted works	<1%
19	ebin.pub Internet	<1%
20	Arts, Sciences & Technology University In Lebanon on 2024-01-20 Submitted works	<1%