DOI: 10.2478/ias-2024-0011 • JIAS • 19(4) • 2024 • 146-161

A Two-stage CNN Based Computer Vision Framework for Automated Validation of Indian Bank Cheques

Debjani Chakraborty[1], Projjal Sahoo[1], Argha Biswas[1], Sujaan Maitra[1], Sourav Saha[1], and Biswajit Halder[2]

1 Narula Institute of Technology, Kolkata, W.B., India; debjani.cse@gmail.com (D.C.); projjalsahoo2017@gmail.com (P.S.); argha.biswas2811@gmail.com (A.B.); sjm28401@gmail.com (S.M.); souravsaha1977@gmail.com (S.S)

2 Hooghly Engineering and Technology College, Hooghly, W.B., India; biswajithalder88@gmail.com

Abstract: Automated bank cheque processing is still considered a challenging task for computer vision researchers. This article proposes a two-stage deep learning-based end-to-end computer vision framework to validate an Indian bank cheque with respect to a few mistakes commonly found in manually entered handwritten fields. The proposed framework primarily works in two stages involving two separate Mask RCNN models to detect two common mistakes due to the absence of any key handwritten field or the presence of any overwritten/strike-through handwritten character in the bank cheque image. The first stage of the Mask RCNN model aims to segment all the key handwritten fields in a bank cheque image, leading to the detection of any missing handwritten field. The second stage, the Mask RCNN model, attempts to detect the presence of any overwritten/strike-through handwritten character in a bank cheque image that may lead to the invalidation of the cheque. Due to the unavailability of any standard dataset for validation purposes, a bank cheque image data repository has been prepared exclusively for developing the proposed framework. Extensive experimentation with the prepared dataset reveals that the proposed framework can outperform some of the popular frameworks by achieving a promising accuracy (98%) in terms of reporting validation errors owing to the aforementioned mistakes in the bank cheque.

Keywords: Automated Cheque Processing, Mask RCNN, Computer Vision.

1 Introduction

Bank cheques are still considered one of the essential entities in the banking sector. The processing of bank cheques plays a critical role in ensuring valid financial transactions without risking any scope of untoward systematic failure [1]. Surprisingly, despite the advancement of technologies, the current mode of processing cheques in many developing countries still primarily involves manual efforts, even for the validation of cheques. Given the high volume of cheques that are deposited daily at banks, implementing an automated cheque validation system can significantly reduce manual workload while enhancing customer service at a reduced cost. Presently, many developed countries worldwide have adopted reasonably accurate as well as fast image-based Cheque clearing Systems (ICS) using advanced computer vision technologies with minimal human intervention to streamline cheque processing in an automated manner. In today's artificial intelligence-dominated world, the automated ICS as a modern document image processing system uses machine learning models in conjunction with digital image processing algorithms to extract required information based on processing a few important fields in cheque images [2].

Typically, the automation of cheque clearing systems may include a series of automated tasks, viz., cheque validation, extraction of important data from the cheque, signature verification, identification, and processing of the relevant data in the cheque for authentic financial transactions. Among many critical data fields that are relevant for bank cheque processing, the most crucial as well as important fields that are to be filled manually on an Indian bank cheque generally consist of the legal amount, courtesy amount, date, payee's name, and signature of the account holder, as shown in Fig.1. The validation of these fields is one of the prerequisites for processing a bank cheque. During the deposition

sciendo © 2024 Debjani Chakraborty et al., published by Sciendo. BY-NC-SA This work is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 4.0

of a bank cheque, even a simple automation module for cheque validation can expedite the entire ICS by rejecting invalid cheques at a very early stage.

In general, there may be multiple criteria that are to be examined for validating a bank cheque. Most of the common mistakes are found in the manually entered fields of the bank cheques. These fields are usually scrutinized minutely during the deposition of a bank cheque to avoid any failure while processing them at a later stage. In connection with these manually entered fields, some of the basic criteria for validating a bank cheque are listed below.

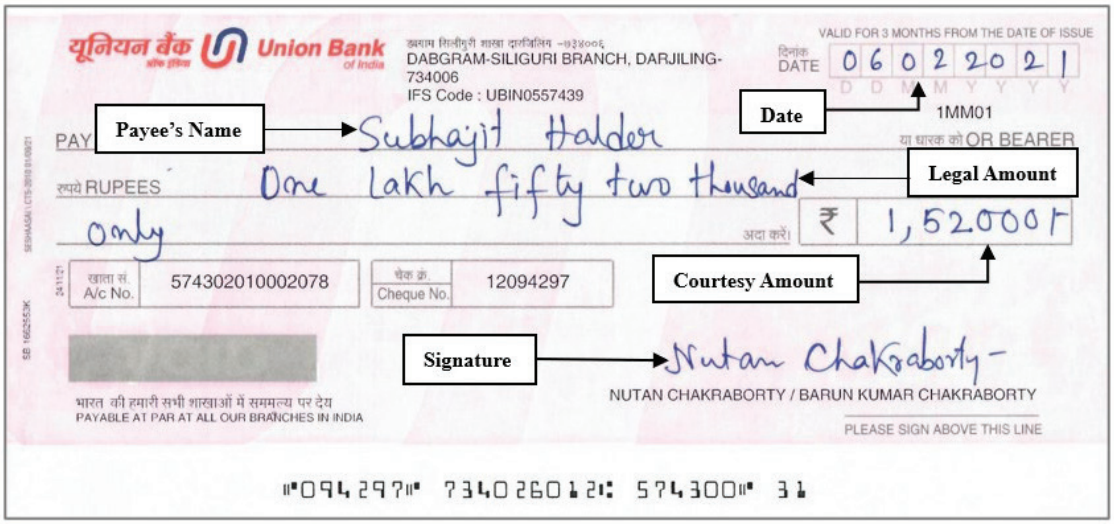


Figure 1. A Valid Indian Bank Cheque Image

1. A bank cheque is considered invalid if any of its key fields remains empty.
2. A bank cheque is considered invalid if any of its key fields contain any overwritten/strike-through handwritten characters.

Over the past few years, many research works have been reported in context with automating bank cheque processing. The majority of the researchers have worked on designing automated tools for key-field segmentation, character recognition, and signature authentication in a bank cheque image with reasonably good accuracy [3]. However, we have rarely come across any prominent research proposals that solely focus on designing an automated validation tool for Indian bank cheque images using deep learning techniques. In view of the dearth of such efforts, we have attempted to develop a simple two-stage computer vision framework with the objective of automating the validation of an Indian bank cheque. The challenges of developing such a validation framework that concern us initially while establishing its efficacy primarily stem from the unavailability of any standard dataset exclusively meant for cheque validation purposes, including other widely faced issues involving the presence of noise and spurious handwritten characters, a varied range of bank cheque layouts and handwriting, the illegibility of handwritten characters in bank cheque images, and many other generic document image processing issues.

Objectives and Contributory Scope of the Proposed Framework:

This article intends to present a two-stage computer vision framework based on deep learning tools that have been proposed to validate an Indian bank cheque with respect to the above-mentioned validation criteria regarding manually entered handwritten key fields, as illustrated in Fig. 2.

During the deployment phase, the initial stage of the framework employs a Mask RCNN-based deep learning tool to detect any manually entered key handwritten fields that are missing in a bank cheque while attempting to segment them in its scanned image format. The second stage of the framework tries to locate the presence of any overwritten/strike-through handwritten character in a bank cheque image using another Mask RCNN-based segmentation model. The performance of the framework is found to be promising in terms of validating a bank cheque with respect to the two aforementioned mistakes. The performance of our Mask RCNN-based model is also compared with two other deep learning models, namely U-Net and YOLO, as all of these models, including Mask RCNN, are widely used as segmentation models in various computer vision applications [4] [5] [6] [7] [8]. The contributory scope of the proposed work is outlined below to highlight the significance of our research efforts.

- During the initial stage, a Mask RCNN model has been developed that aims to segment the key handwritten fields in a bank cheque image. The segmented fields are analyzed to detect any missing handwritten field, and a validation error is generated if the model fails to segment any key handwritten field, i.e., legal amount, courtesy amount, date, payee's name, and signature of the account holder.
- For the second stage, another Mask RCNN segmentation model has been developed to locate the presence of any overwritten/strike-through handwritten character in a bank cheque image that may lead to the invalidation of the cheque.
- Due to the unavailability of any standard dataset for validation purposes, a bank cheque image data repository has been prepared exclusively for developing the proposed framework. This repository has also been made publicly available for future reference.

During experimentation, the performance of our Mask RCNN-based framework is compared with exploring other popular segmentation models like U-Net [5] and YOLO [6]. The viability of the proposed framework has been established through extensive experimentation with scanned images of Indian bank cheques that include various kinds of mistakes, such as missing handwritten fields and containing overwritten/strike-through handwritten characters.

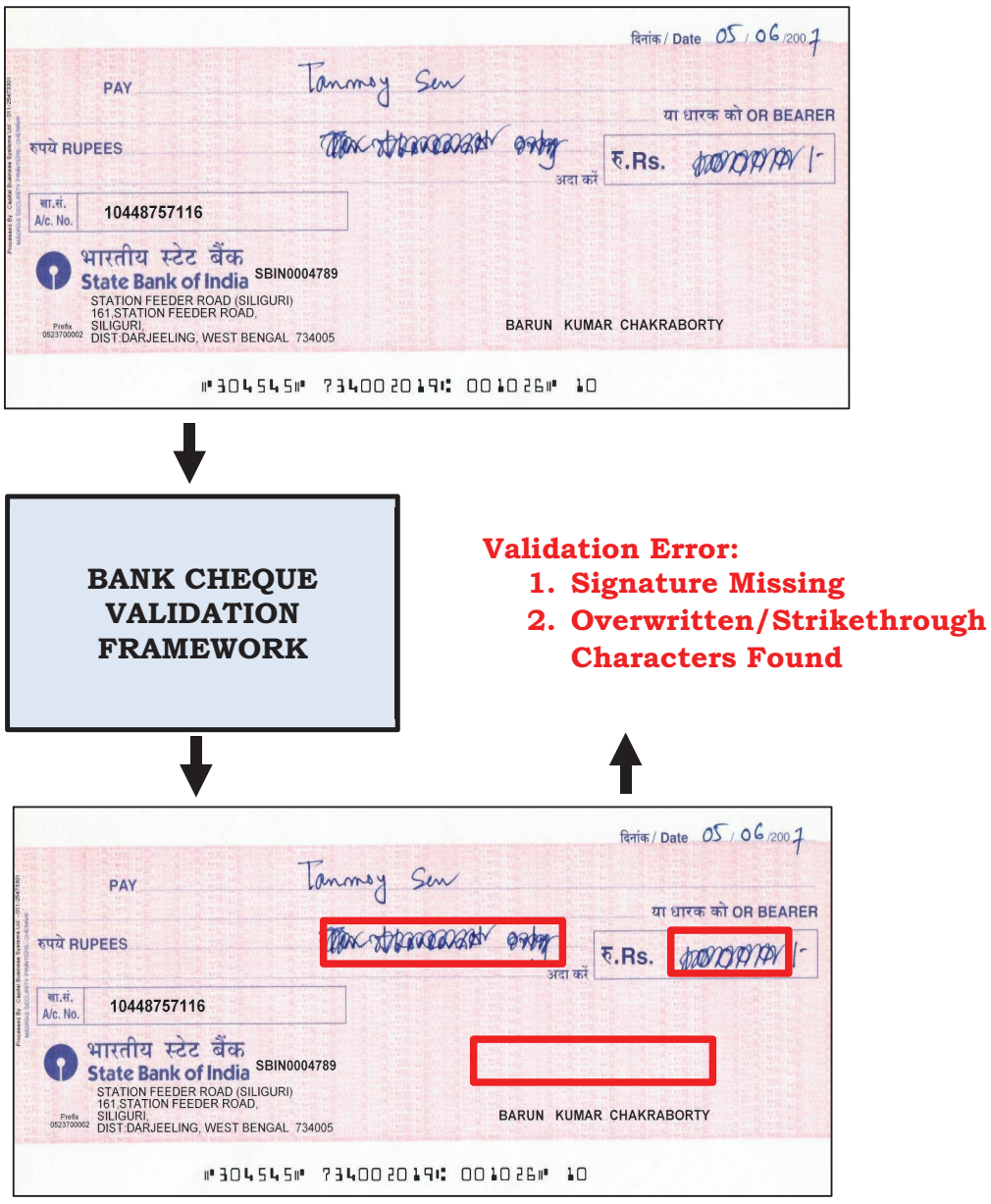


Figure 2. Automated Detection of an invalid Indian Bank Cheque Image

The rest of the article begins with Section 2, which highlights the related research work carried out so far for automated bank cheque processing. Section 3 describes the proposed methodology with a detailed illustration of the deep learning models that have been employed to validate a scanned Indian bank cheque. Section 4 reports the results of our experimentation along with an analysis of the proposed framework's performance. Finally, Section 5 summarizes the proposed work while indicating the probable future work as an extension of the proposed work.

2 Related Work

The literature study reveals many notable works attempted by the researchers to automate bank cheque processing [3] [1]. The majority of works have been dedicated to the extraction and recognition of various key fields in the bank cheque [9]. This section summarizes some of the notable works reported so far in the context of bank cheque image processing.

A machine learning-based system for identifying and detecting the signature and amount fields in the digital format of a cheque image was proposed by Miah M.B. *et al.* [10]. Their suggested framework locates the bank cheque's courtesy field region using layout-specific data. After extracting characteristics from the segmented digits inside the bank cheque's courtesy field, they used an ANN model to recognize the courtesy amount. Vamsi *et al.* [11] applied sliding window techniques in a bank cheque image to locate the position of the signature. The cheque image is examined block-wise by a movable window of varying height and breadth, and the entropy of the pixels inside the window is calculated based on pixel density. The signature field on the bank cheque is localized by comparing the entropy of each window to the pre-determined reference range.

In [12], Dansena *et al.* proposed a unique method for extracting regions of interest (ROI) from a cheque image. It analyses the dimensional and spatial properties of various horizontal lines in a bank cheque to isolate the ROIs.

In order to segment the courtesy field from a bank cheque, Chandra *et al.* [13] used cheque layout characteristics to locate a bounding box around the courtesy amount in a bank cheque. Palacios *et al.* [9] reported a hybrid architecture to recognize courtesy amounts from US and Brazilian cheques written in numerical format. Their proposed framework extracts individual digits from the segmented courtesy field and employs a neural network architecture to recognize the extracted digits.

Hakim A. Abdo *et al.* explored the potential of the Faster RCNN deep learning model [14] to segment crucial bank cheque fields by training it with annotated bank cheque images. Their proposed model automatically finds probable locations of bank cheque fields. They used IDRBT Cheque Image Data to develop the model, and the results showed that the model had a promising ability of segmenting the crucial bank cheque fields with an accuracy of 97.4%.

Agarwal *et al.* [15] developed a framework that used SIFT (Scale Invariant Feature Transform) features to locate various key fields of a bank cheque which are essential for cheque clearance. They employed a deep learning-based CNN method for recognizing handwritten numeric characters on bank cheques with an accuracy of 99.14%. They also used SIFT features in fusion with SVM classifiers to verify signatures with 98.10% accuracy.

Rajib Ghosh *et al.* developed an automated bank cheque processing system for Indian bank cheques that attempts to recognize handwritten characters present in the 'payee name', 'courtesy amount (both in words and figures)' and 'date' fields. They employed SVM classifiers on handwritten field-specific features, which are derived by using the Histogram of Oriented Gradients (HOG) method on the Grey Level Co-occurrence Matrix (GLCM) texture of handwritten fields [16].

Mukesh Jha *et al.* trained a convolutional neural network with an IAM dataset to extract handwritten fields from cheque leaves and used OCR tools to recognize those fields [17]. They also explored the principal component analysis technique for verifying the signatures on a bank cheque.

Alirezaee *et al.* [18] developed an algorithm based on the concept morphological and connected component analysis for segmenting the components of Iranian bank cheques. The segmented fields are further processed by the OCR application for recognition. Their proposed algorithm achieves 96.58% accuracy in recognizing five handwritten fields of the Iranian bank cheques.

Madasu *et al.* [19] used the idea of connected component analysis for the extraction of various information fields in bank cheques and used entropy, energy, aspect ratio, and the average fuzzy membership value of the extracted fields as the basis of feature formation for developing a fuzzy neural network model to recognize the information fields.

Mehta *et al.* [20] proposed a unique technique of combining HMM and sum graph approaches for forming feature vectors to authenticate signatures in bank cheques. The proposed model's character segmentation accuracy is found to be 95%, while the accuracy of non-numeric character recognition and digit recognition is reported at 83%, and 91% respectively. The model can locate the signature in a bank cheque image with 91% accuracy but scores only 80% in terms of authenticate accuracy.

In [10], Miah *et al.* employed a neural network-based model for recognizing handwritten numeric characters and signatures in bank cheque images. The proposed system extracts twenty-six geometric features from manually segmented courtesy and signature fields for training the neural network model. The proposed model achieved 93.4% and 96% recognition rates for the courtesy amount field and signatures field, respectively.

So far, several studies have been conducted on the automation of bank cheque processing. The majority of studies have focused on developing automated methods with a high level of accuracy for segmenting critical fields, recognizing characters, and authenticating signatures in bank cheque images. The challenges that the researchers have faced while attempting data field extraction and recognition in bank cheque images are listed in Table 1.

Over the last few years, many researchers have developed automated tools while leveraging the potential of deep learning techniques for handwritten text recognition in document images with high levels of accuracy [21] [22]. Interestingly, there are hardly any notable research proposals that specifically concentrate on developing an automated validation tool for Indian bank cheque images using deep learning techniques [23]. Given the lack of such initiatives, we have endeavored to develop a straightforward two-step computer vision system with the aim of automating the verification process for an Indian bank Cheque.

Table 1. Challenges of extracting and recognizing data fields from bank cheque images

Challenges	Difficulties
Data deterioration	• The post-binarization nose, including stamps, dots and lines are observed. • The decrease of quality resulting from the removal of noise, lines, and backgrounds.
The problem of Skewness	• Misalignment of the cheque during the scanning process. • The segmentation and recognition of bank cheques pose significant challenges.
Distinct handwriting	• The distinctive handwriting style complicates the process of data recognition. • Various persons use several font sizes, directions, thicknesses, and angles.
Data superposition	• Data overlap caused by adjacent words resulting in insufficient differentiation of the data.
Perplexity	• "/" occasionally contacts the adjacent digits. • The "/" symbol poses a challenge in the process of segmenting and recognizing numbers.
Document torn and folded	• The corners are predominantly folded or ragged.
Variation in image contrast	• The image is either excessively bright or excessively dark, which hinders the extraction of data. • Issues with image camera calibration, subpar printing, and incorrect thresholding of the background.
Cheque Streaks	• The presence of dark or bright streaks can be attributed to various factors, including scratches in the scan window, dirt on the camera calibration target, and malfunctions in the camera electronics.
Image compression	• Difficulties arise when the image is compressed below the minimal requirement or when the image size exceeds the maximum range.
Integrated algorithm	• There is currently no standardized method available that can extract all the information from a bank cheque simultaneously.

3 Proposed Methodology

The proposed framework for validating a bank cheque works on the scanned image of a bank cheque. Fig. 3 outlines the process flow of the proposed two-stage framework to illustrate the functionalities of various modules within the framework, which are designed to report any validation errors.

The scanned bank cheque image is initially converted into a gray image, which is later binarized using OTSU's thresholding method [24]. Next, the stage-1 Mask RCNN model is applied to the binary image to segment all the handwritten fields in the bank cheque. The stage-1 error reporting module analyzes the segmentation results of the stage-1 Mask RCNN model to detect whether any key handwritten fields are missing. Subsequently, the binary image is masked based on the segmentation results of the stage-1 Mask RCNN model. The stage-2 Mask RCNN model works on the masked image to segment overwritten/strike-through handwritten characters. The stage-2 error reporting module examines the segmentation outcome of the stage-2 Mask RCNN model to detect any presence of an overwritten/strike-through handwritten character in the bank cheque image.

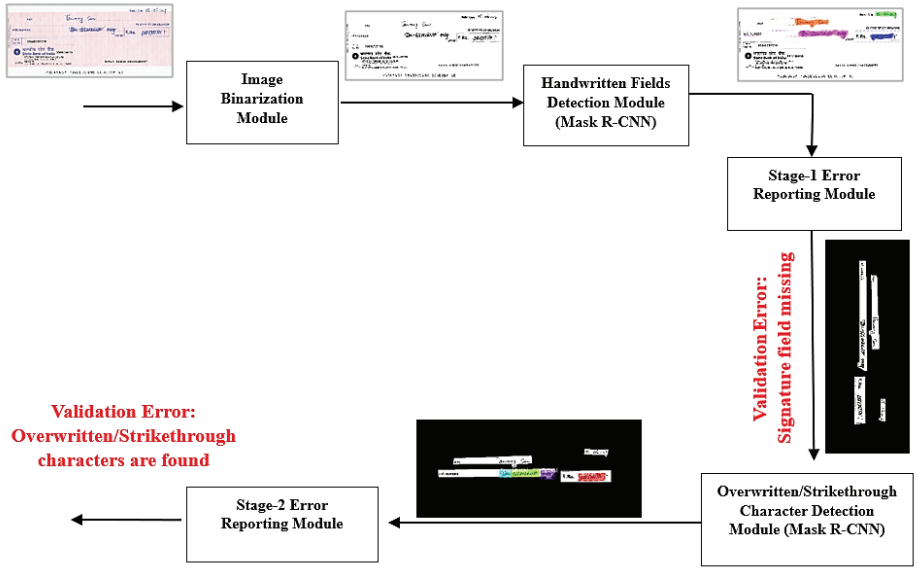


Figure 3. Overview of the process flow of the proposed framework

General Overview of the Mask RCNN Model used in the Proposed Framework:

The proposed framework employs Mask Regional Convolutional Neural Network (Mask RCNN) [25] [26] wherein ResNet101 acts as the backbone. Mask RCNN is a deep neural network model that is widely used for segmenting the instance of an object located in the given image (Fig.4).

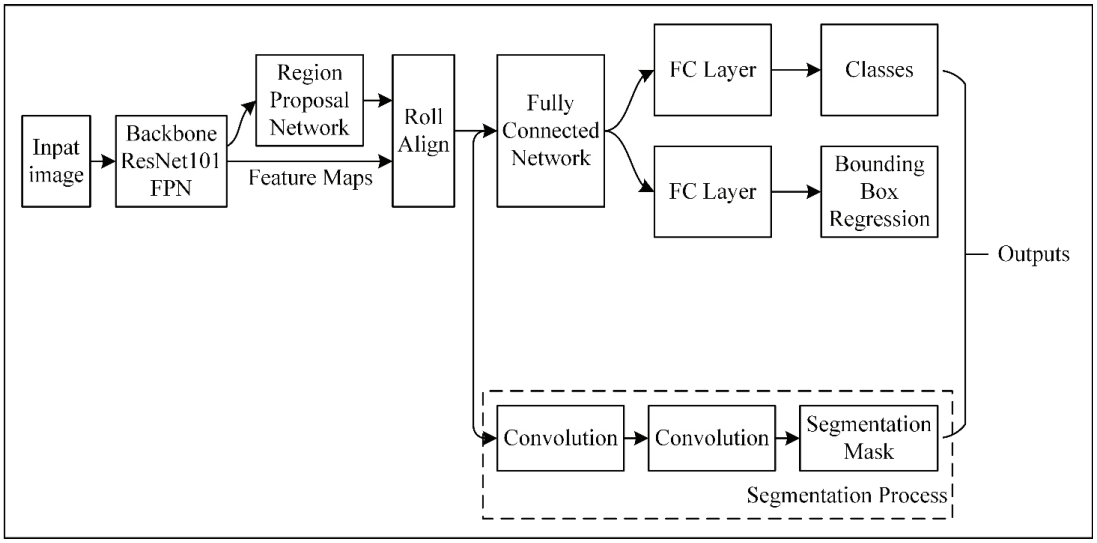


Figure 4. Overview of the Mask RCNN architecture

The multi-layer ResNet-101 architecture, which serves as a feature extractor in our proposed work, generates feature maps. The three stages of the architecture include the bottom-up pathway, top-down pathway, and lateral connections (Fig. 5).

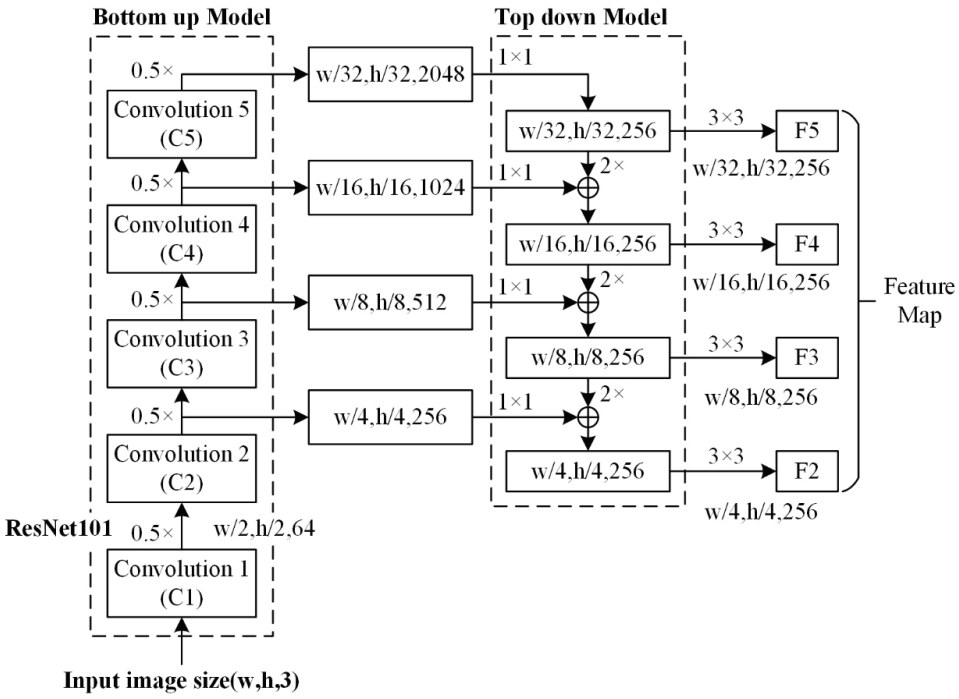


Figure 5. Overview of the Resnet 101 architecture of the Mask RCNN Model

Each layer of the bottom-up pathway reduces the feature map size by half while doubling the number of feature maps. During the top-down pathway stage, the upscale operations begin with a top feature map of a smaller dimension and gradually continue to produce larger feature maps. Before sampling, the 1x1 convolution is applied to reduce the number of channels (256). As a result, element-wise addition merges the feature maps of the two pathways. After element-wise addition, 3 x 3 convolution layers generate the four feature maps (F2, F3, F4, and F5). Finally, we apply the max pooling operation to F5 to generate an additional feature map (F6).

In order to determine the segmentation mask for the target object, these feature maps, along with region proposals, are passed through a series of convolution layers.

A. Stage-1 Mask RCNN Model Development for detection of missing handwritten key fields:

During the initial stage, a Mask RCNN model has been developed that aims to segment the key handwritten fields in a bank cheque image, as shown in Fig. 6. The segmented fields are analyzed to detect any missing handwritten field, and a validation error is generated if the model fails to segment any key handwritten field, i.e. legal amount, courtesy amount, date, payee's name, and signature of the account holder. An Indian bank cheque along with annotations of its various handwritten fields, is shown in Fig. 7.

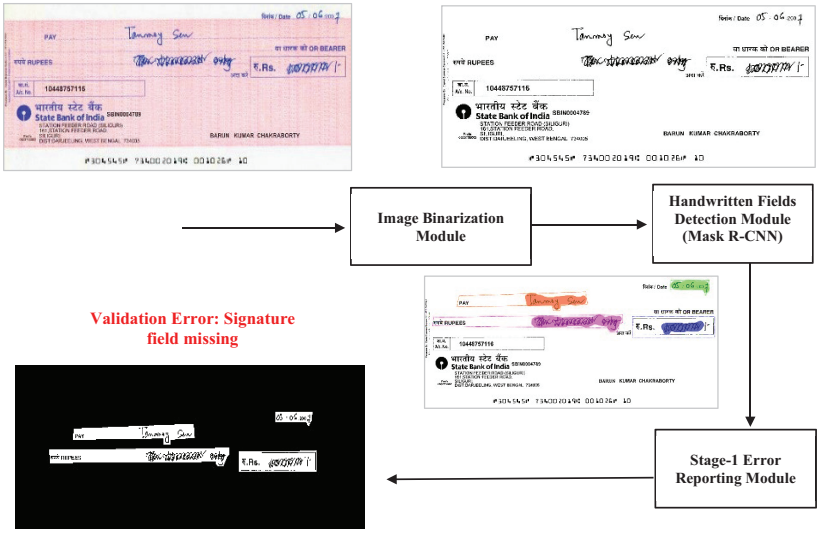


Figure 6. Flow-diagram of First Stage validation error reporting module

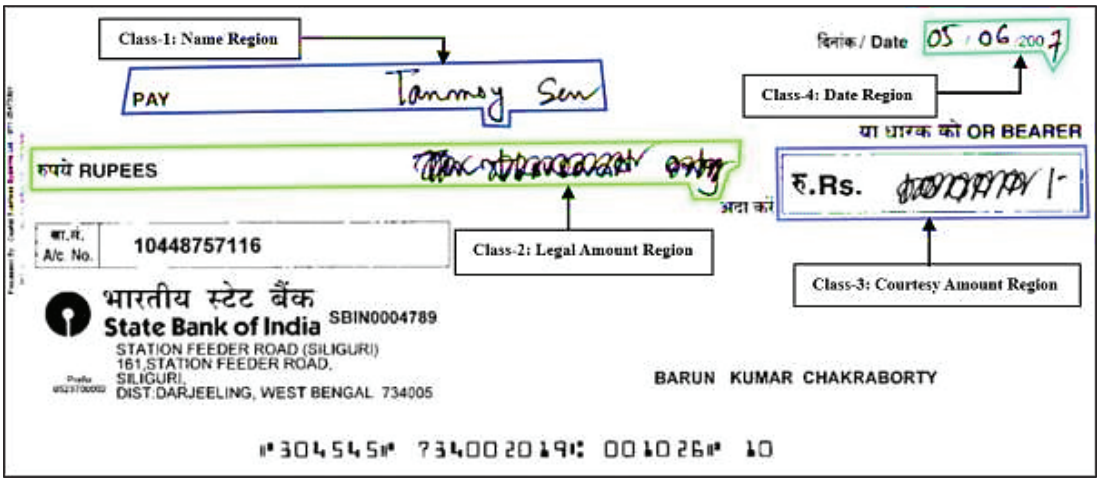


Figure 7. Annotated Bank cheque image for stage-1 Mask RCNN model

B. Stage-2 Mask RCNN Model Development for detection of overwritten/strike-through handwritten characters:

At the second stage (Fig. 8), another Mask RCNN segmentation model has been developed to locate the presence of any overwritten/strike-through handwritten character in a bank cheque image that may lead to invalidation of the cheque. As stated earlier, the binary image of the bank cheque is masked based on the segmentation results of the stage-1 Mask RCNN model. The annotation is performed on the masked binary image of the bank cheque image for stage-2 of the Mask RCNN model, as shown in Fig. 9.

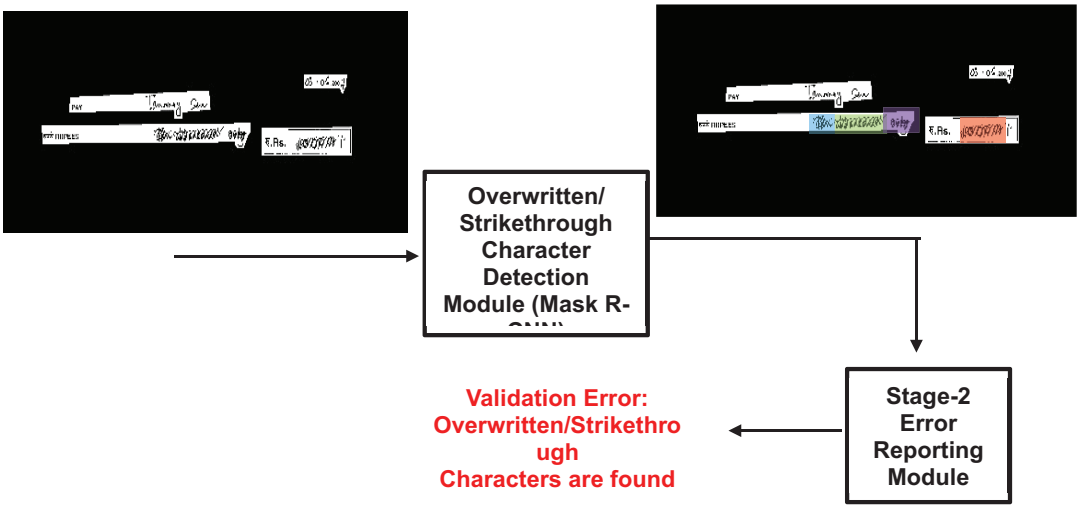


Figure 8. Flow-diagram of Second Stage validation error reporting module

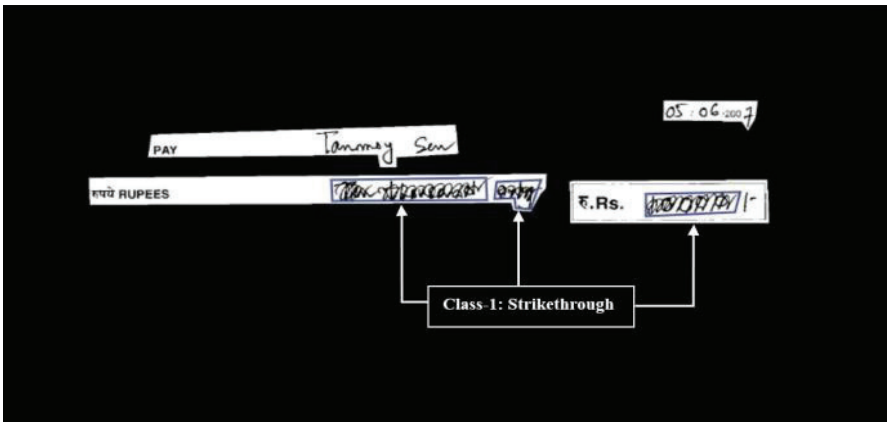


Figure 9. Annotated Bank cheque image for stage-2 Mask RCNN model

The stage-2 Mask RCNN model is applied to the masked image for segmenting overwritten/strike-through handwritten characters. The second-stage Mask RCNN model is developed using various types of overwritten/strike-through handwritten characters, as shown in Fig. 10.

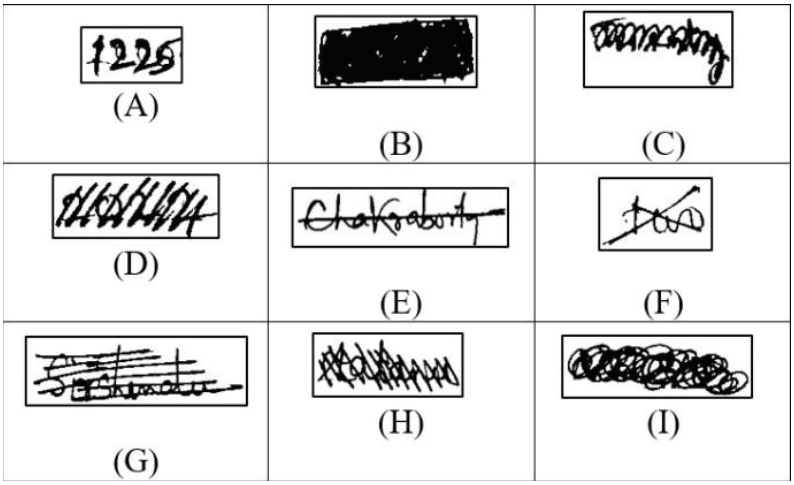


Figure 10. Different types of unacceptable overwritten/strike-through handwritten characters

4 Experimentation, Results & Analysis

This section provides a detailed overview of our experimental setup and reports on the results of our proposed framework. The preparation of the dataset that has been used in our experimentation is described here in detail. We deep dived into the experimental findings for both stages of our Mask RCNN-based two-stage framework and specified the individual performance for each stage. Additionally, a comparative analysis is also conducted by contrasting the performance of our framework with that of YOLOv8 and U-Net in both stages.

A. Data set Preparation

Due to the unavailability of any standard dataset for validation purposes, a bank cheque image data repository has been prepared exclusively for developing the proposed framework. This repository has also been made publicly available for future references [27]. For our experimental investigation, we have meticulously compiled a distinctive dataset of 120 bank cheque images with various dimensions from five renowned Indian banks: the State Bank of India, Axis Bank, Canara Bank, Central Bank, and Union Bank. The dataset comprises 50 bank cheques from the State Bank of India, 30 from Axis Bank, 20 from Canara Bank, 10 from the Central Bank, and 10 from Union Bank. All the important handwritten information fields in the bank cheques have been manually filled out by 15 volunteers in English using 12 pens with blue and black ink.

The dataset contains 30 valid bank cheques and 90 invalid bank cheques. For the first stage analysis of the proposed framework, a small set of cheques was intentionally kept devoid of one or two key handwritten information fields, such as the date or signature, as depicted in Fig 11. For the second stage of investigation, overwritten and strike-through handwritten characters were deliberately added to the bank cheques to create authentic complications, as depicted in Fig.

12. Table 2 summarizes the nature of variations in our dataset exclusively prepared for validating bank cheques with respect to the mistakes commonly found in manually entered handwritten fields of Indian bank cheques. A few binary images of bank cheques that are used in our experimentation for developing stage-1 and stage-2 Mask RCNN models are shown in Fig. 11 and Fig. 12 respectively.

Table 2. Summary of the Dataset

Cheques with Missing Handwritten Field	Cheques with No Missing Handwritten Field	Cheques with Overwritten/ Strike-through Characters	Cheques with no Overwritten/ Strike-through Characters
90	30	78	12

Figure 11. A few Bank Cheque sample images used for developing Stage-1 Mask RCNN Model

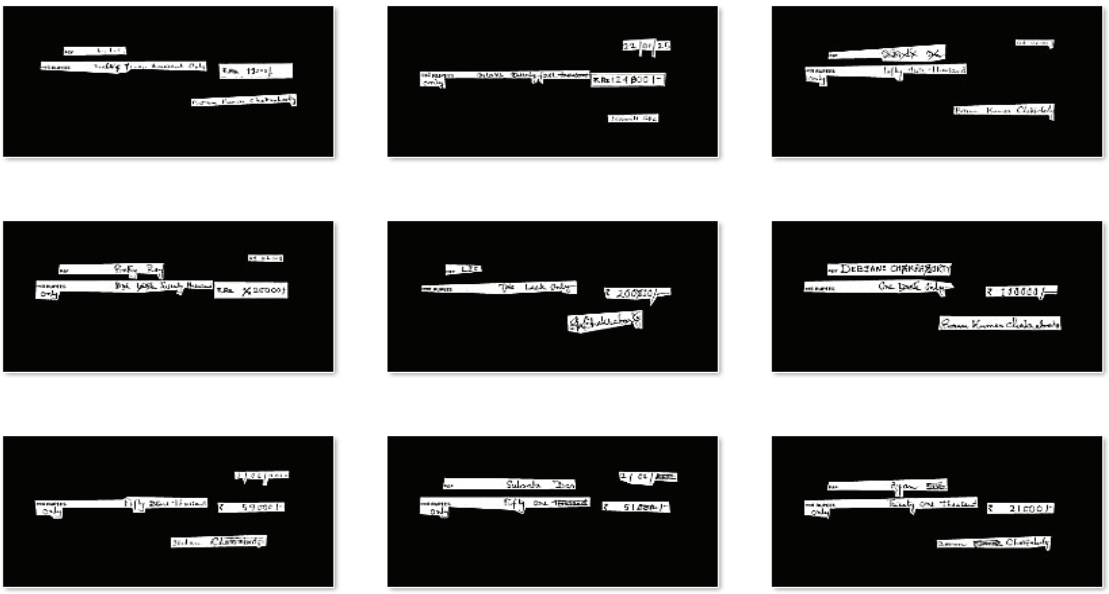
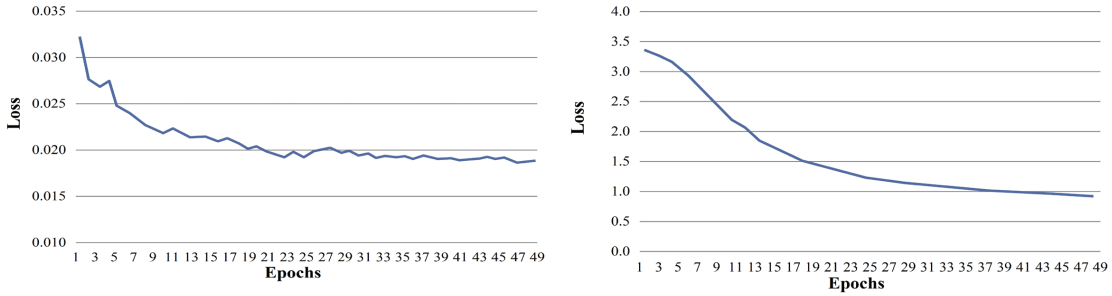


Figure 12. A few Masked Bank Cheque sample images used for developing Stage-2 Mask RCNN Model

B. Experimental Setup:

For our experimental environment, the Google Collaboratory platform is used with a NVIDIA SMI GPU and 32 GB of RAM. Our Mask RCNN model employs stochastic gradient descent with momentum, while U-Net uses the Adam opti- mizer and YOLOv8 uses the AdamW optimizer. A total of 85% of the dataset was allocated for the training phase during the development of the framework, with the remaining 15% being set aside for testing purposes. During the prepossessing phase, all the differently sized bank cheque images were standardized to a uniform format. The crucial handwritten information fields on bank cheques have been polygonally annotated for the first stage, while in the second stage, we have performed polygonal annotation of the overwritten and strike-through handwritten characters on the bank cheques. The loss graphs during the training phases of both stage-1 and stage-2 Mask RCNN models are shown in Fig. 13a and Fig. 13b, respectively.



(a) Loss Graph: Stage-1 Mask RCNN

(b) Loss Graph: Stage- 2 Mask RCNN

Figure 13. Loss Graph During Training Phase of Mask RCNN Models

C. Performance Evaluation Criteria:

In our proposed framework, we primarily determine if a bank cheque is valid or invalid by detecting the (I) presence of all the manually entered handwritten fields and (II) presence of overwritten or strike-through characters in the segmented key information fields in bank cheques. The stage1 Mask RCNN model detects the presence of handwritten fields, while the stage-2 Mask RCNN model detects the presence of overwritten or strike-through handwritten characters in the bank cheque images. The performance of the proposed framework is assessed in terms of the detection accuracy of these two Mask RCNN models as they are deployed to segment the above-mentioned regions in an Indian bank cheque image. The accuracy of detection is measured using the IOU metric, which primarily focuses on the percentage of the number of pixels common between the target and prediction masks. In our framework, the IOU threshold is set above 50% for validating the detection of a target region.

The accuracy of any handwritten field detection is measured in terms of the IOU metric, which primarily focuses on the percentage of the number of pixels common between the target and prediction masks. IOU is computed using Eqn. 1, where $Predict_{BBox}$ indicates the rectangular bounding box enclosing the courtesy field area predicted by a framework and GT_{BBox} denotes the rectangular bounding box enclosing the original target area detected as ground truth.

$$IOU = \frac{Predict_{BBox} \cap GT_{BBox}}{Predict_{BBox} \cup GT_{BBox}} \tag{1}$$

The accuracy of the proposed framework is computed based on the True Positive (TP), True Negative (TN), False Positive (FP), False Negative (FN) as determined based on the following criteria:

- True Positive (TP): If our model can detect any manually entered handwritten field or Overwritten/Stricken-through characters properly in the scanned bank cheque images, it will be treated as a True Positive case.
- True Negative (TN): If our model correctly identifies a valid bank cheque in the absence of any mistakes in manually handwritten fields, it will be treated as True Negative case.
- False Positive (FP): If our model falsely generates any validation error, even in the absence of any mistakes, it will be treated as False Positive.
- False Negative (FN): False Negative cases refer to situations when our model fails to identify any mistakes in the handwritten fields of the bank cheque images.

$$Detection\ Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{2}$$

D Results of Stage-1 Validation Module:

The results of our stage-1 validation module are reported here, providing both qualitative and quantitative measures. Table 3 displays the qualitative performance of the proposed framework in comparison to U-Net and YOLOv8. It has been observed that our proposed Mask RCNN based framework outperforms both U-Net [28] and YOLOv8 in terms of accurate detection of manually entered handwritten fields through segmentation. The quantitative performance measures are provided in Table 4 wherein the detection accuracy is compared with other frameworks. Notably, the detection accuracy of various handwritten fields in bank cheque images using the proposed framework indicates its supremacy over other frameworks.

Table 3. Qualitative Observations of Stage-1 Validation Module's Performances

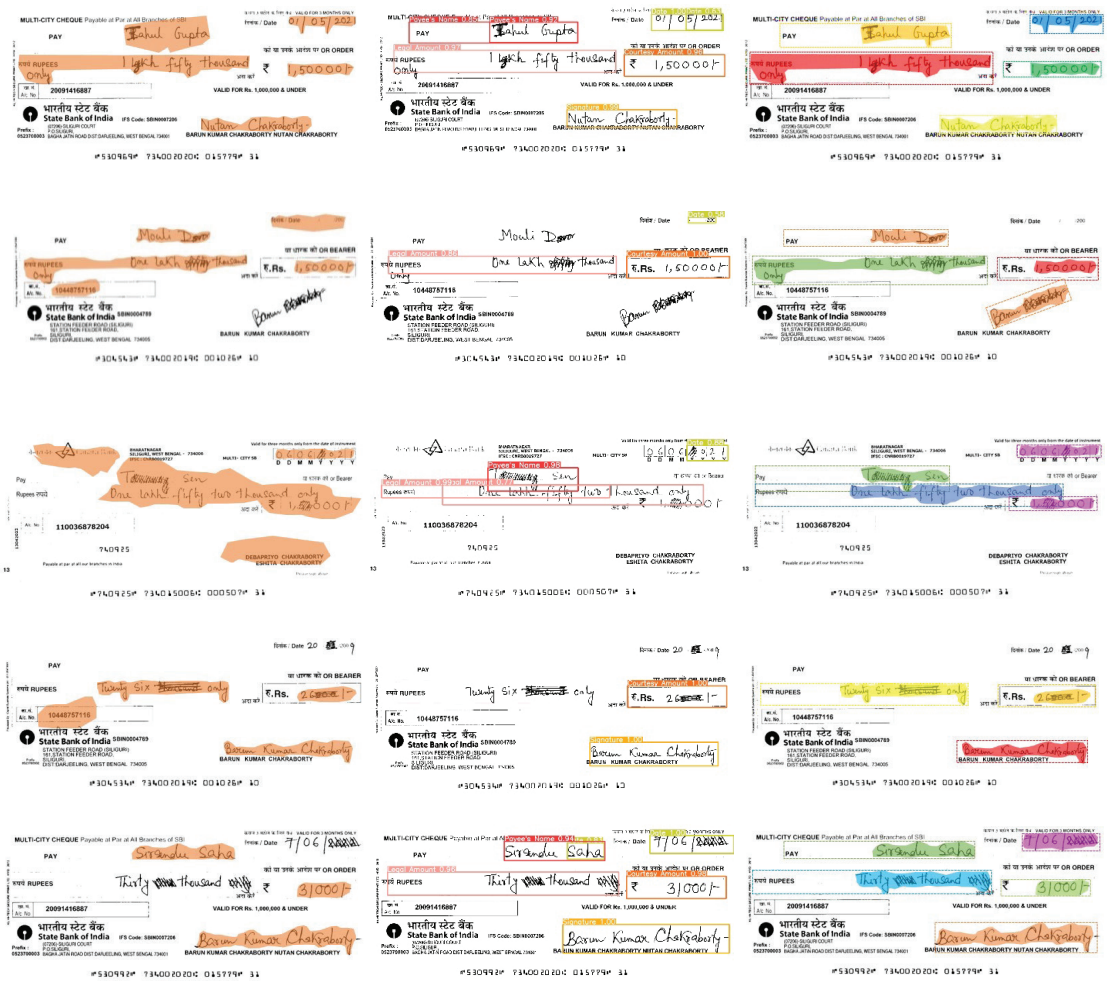
U-Net	YOLO v8	Proposed Mask RCNN based model
		

Table 4. Quantitative Results: Detection Accuracy of Stage- Validation Module

Model	Payee's Name (%)	Legal Amount (%)	Courtesy Amount (%)	Signature (%)	Date (%)
Madasu [19]	90.0	91.6	85.2	90.3	87.5
Hakim [14]	97.5	72.0	75.4	69.0	75.8
Alirezaee [18]	94.4	93.2	94.15	90.8	92.1
U-Net [29]	95.3	94.9	95.1	95.2	91.8
YOLO v8 [30]	96.2	94.8	95.1	91.0	93.3
Proposed Model	98.1	98.0	97.4	97.2	98.2

B. Results of Stage-2 Validation Module:

As described before, our stage-2 validation module attempts to detect the presence of any overwritten or strike-through of handwritten characters in bank cheque images. The qualitative as well as the quantitative aspects of the performance of our stage-2 validation module are reported in Table 5 and Table 6 respectively. The detection accuracy of our proposed Mask RCNN based model in the second stage is found to be better as compared to YOLOv8 and U-Net. Notably, a few FP cases were observed during our experimentation, wherein some valid handwritten characters (e.g., t, g, y, f, etc.) were misidentified as strike-through characters.

Table 5. Qualitative Observations of Stage-2 Validation Module's Performances

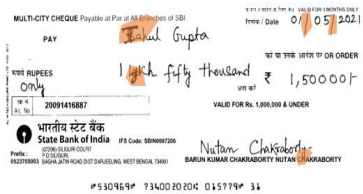
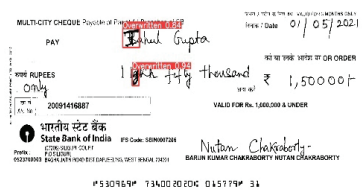
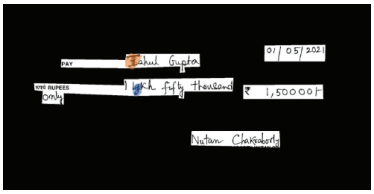
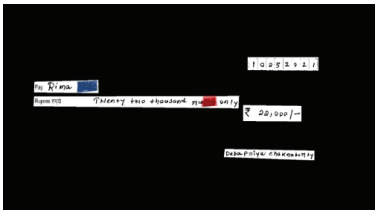
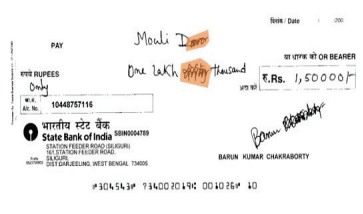
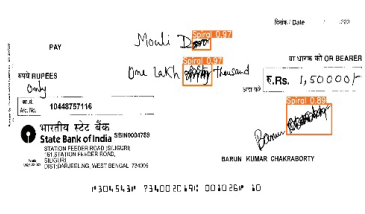
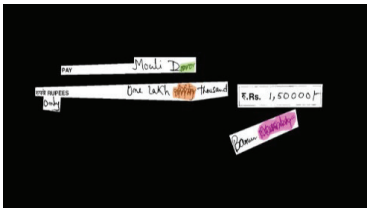
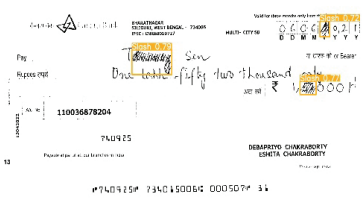
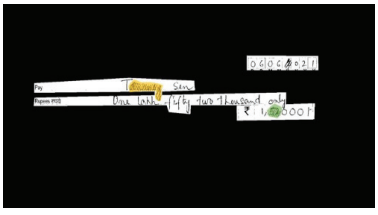
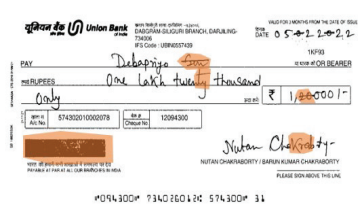
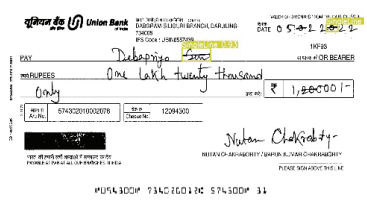
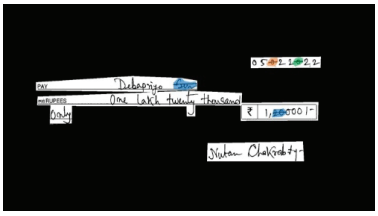
U-Net	YOLO v8	Proposed Mask RCNN based model
		
		
		
		
		

Table 6. Quantitative Results: Detection Accuracy of Stage-2 Validation Module

Overwritten/ Strikethrough class types	U-Net based model (%)	YOLO v8 based model (%)	Proposed Mask R-CNN based model (%)
Type - A	94.30	95.40	98.20
Type – B	93.90	94.56	98.03
Type – C	94.75	95.2	97.57
Type – D	95.39	95.42	98.50
Type – E	92.67	93.71	98.90
Type – F	93.25	95.83	97.82
Type – G	95.28	95.07	98.23
Type – H	92.33	94.89	97.90
Type – I	93.80	94.77	97.88

5 Conclusion

This article presents a computer vision framework that utilizes deep learning tools in fusion with some refinement strategies to validate an Indian bank cheque with respect to the mistakes commonly found in manually entered handwritten fields. The proposed framework primarily works in two stages, involving two separate Mask RCNN models to detect mistakes due to the absence of any key handwritten field or the presence of any overwritten/strike-through handwritten characters in the bank cheque image. The first stage Mask RCNN model aims to segment all the key handwritten fields in a bank cheque image, leading to the detection of any missing handwritten field. The second-stage Mask RCNN model attempts to locate the presence of any overwritten/strike-through handwritten character in a bank cheque image that may lead to invalidation of the cheque. The proposed framework achieves promising accuracy (98%) in terms of reporting validation errors owing to the aforementioned mistakes in the bank cheque.

The framework may further be extended to validate many other issues related to the handwritten fields, like verifying the date field, matching the legal amount with the courtesy amount, and authenticating the signature field.

Acknowledgments

The authors acknowledge the support from the Narula Institute of Technology, which provided the necessary infrastructural facilities to carry out the research activities.

References

[1] Jayadevan, R., Kolhe, S., Patil, P., & Pal, U. (2012). 'Automatic processing of handwritten bank cheque images: a survey'. *International Journal On Document Analysis And Recognition (IJDAR)*, 15, 267–296.

[2] NIGAM, S., Verma, S., & Nagabhushan, P. (2023). 'Document Analysis and Recognition: A survey'. *Authorea Preprints*.

[3] Thakur, N., Ghai, D., & Kumar, S. (2023). 'Automatic imagery Bank Cheque data extraction based on machine learning approaches: a comprehensive survey'. *Multimedia Tools And Applications*, 1–56.

[4] Maity, S., Biswas, S., Manna, S., Banerjee, A., Llado´s, J., Bhattacharya, S., & Pal, U. (2023). 'Selfdocseg: A selfsupervised vision-based approach towards document segmentation'. In *International Conference On Document Analysis And Recognition* (pp. 342–360).

[5] Fizaine, F., Bard, P., Paindavoine, M., Robin, C., Bouye´, E., Lefe`vre, R., & Vinter, A. (2024). 'Historical Text Line Segmentation Using Deep Learning Algorithms: Mask-RCNN against U-Net Networks'. *Journal Of Imaging*, 10, 65.

[6] Kumar, A., & Lehal, G. (2024). 'Layout Detection of Punjabi Newspapers using the YOLOv8 Model'. *International Journal Of Performability Engineering*, 20.

[7] Affes, N., Ktari, J., Ben Amor, N., Frikha, T., & Hamam, H. (2023). 'Comparison of YOLOV5, YOLOV6, YOLOV7 and YOLOV8 for Intelligent Video Surveillance'. *Journal Of Information Assurance & Security*, 18.

[8] Shaikh, A., Mishra, K., Kharade, P., & Kanojia, M. (2023). 'Comprehensive Study on Emotion Detection with Facial Expression Images Using YOLO Models'. *Journal Of Information Assurance & Security*, 18.

[9] Palacios, R., & Gupta, A. (2008). 'A system for processing handwritten bank cheques automatically'. *Image And Vision Computing*, 26, 1297–1313.

[10] Miah, M., Yousuf, M., Mia, M., & Miya, M. (2015). 'Handwritten courtesy amount and signature recognition on bank cheque using neural network'. *International Journal Of Computer Applications*, 118.

[11] Kubik, K., Hanmandlu, M., Yusof, M., & Krishna Madasu, V. (2003). 'Automatic extraction of signatures from bank cheques and other documents'. In *Proceedings Of The Seventh International Conference On Digital Image Computing: Techniques And Applications, DICTA 2003*.

[12] Dansena, P., Kumar, K., & Pal, R. (2015). 'Line based extraction of important regions from a cheque image'. In *2015 Eighth International Conference On Contemporary Computing (IC3)* (pp. 183–189).

[13] Chandra, L., Gupta, R., Kumar, P., & Ganotra, D. (2008). 'Automatic courtesy amount recognition for Indian banks' cheques'. In *TENCON 2008-2008 IEEE Region 10 Conference* (pp. 1–5).

[14] Abdo, H., Abdu, A., Manza, R., & Bawiskar, S. (2023). 'Extraction of Bank Cheque Fields Based on Faster RCNN'. In *First International Conference On Advances In Computer Vision And Artificial Intelligence Technologies (ACVAIT 2022)* (pp. 130–139).

[15] Agrawal, P., Chaudhary, D., Madaan, V., Zabrovskiy, A., Prodan, R., Kimovski, D., & Timmerer, C. (2021). 'Automated bank cheque verification using image processing and deep learning methods'. *Multimedia Tools And Applications*, 80, 5319-5350.

[16] Ghosh, R., Panda, C., & Kumar, P. (2018). 'Handwritten text recognition in bank cheques'. In *2018 Conference On Information And Communication Technology (CICT)* (pp. 1–6).

[17] Jha, M., Kabra, M., Jobanputra, S., & Sawant, R. (2019). 'Automation of cheque transaction using deep learning and optical character recognition'. In *2019 International Conference On Smart Systems And Inventive Technology (ICSSIT)* (pp. 309–312).

[18] Alirezaee, S., Allahiari, M., Askari, F., & Alijamat, A. (2008). 'Segmentation of Iranian bank cheques'. In *The International Conference On Electrical Engineering* (pp. 1–7).

[19] Madasu, V., & Lovell, B. (2005). 'Automatic segmentation and recognition of bank cheque fields'. In *Digital Image Computing: Techniques And Applications (DICTA'05)* (pp. 33–33).

[20] Mehta, M., Sanchati, R., & Marchya, A. (2010). 'Automatic cheque processing system'. *International Journal Of Computer And Electrical Engineering*, 2, 761.

[21] Ambikapathy, & Singh, A. (2018). 'Fundamental Concepts of Neural Networks and Deep Learning of Different Techniques to Classify the Handwritten Digits'. In *Advances In Systems, Control And Automation: ETAEERE-2016* (pp. 287–297).

[22] Chauhan, V., Singh, S., & Sharma, A. (2024). 'HCR-Net: A deep learning based script independent handwritten character recognition network'. *Multimedia Tools And Applications*, 1–35.

[23] Chaudhary, D., Agrawal, P., & Madaan, V. (2019). 'Bank cheque validation using image processing'. In *Advanced Informatics For Computing Research: Third International Conference, ICAICR 2019, Shimla, India, June 15–16, 2019, Revised Selected Papers, Part I 3* (pp. 148–159).

[24] Poudel, U., Regmi, A., Stamenkovic, Z., & Raja, S. (2023). 'Applicability of OCR Engines for Text Recognition in Vehicle Number Plates, Receipts and Handwriting'. *Journal Of Circuits, Systems And Computers*, 32, 2350321.

[25] He, K., Gkioxari, G., Dolla´r, P., & Girshick, R. (2017). 'Mask r-cnn'. In *Proceedings Of The IEEE International Conference On Computer Vision* (pp. 2961–2969).

[26] Kiatphaisansophon, P., Wanvarie, D., & Cooharojananone, N. (2024). 'Efficient Text Bounding Box Identification Using Mask R-CNN: Case of Thai Documents'. *IEEE Access*.

[27] Bank-Cheque-Validation-Dataset: https://github.com/souravsahaGITHUB1/BankCheque-Validation-Dataset1 (2024).

[28] Wang, Y., Tian, S., Wen, M., Ruan, Y., Tao, Q., Zhou, X., Gao, F., Lu, H., & Zhang, Z. (2023). 'An end-to-end method for palm-leaf manuscript segmentation based on U-Net'. *Journal Of Cultural Heritage*, 63, 169–178.

[29] Nguyen-Trong, K. (2022). 'An End-to-End Method to Extract Information from Vietnamese ID Card Images'. *International Journal Of Advanced Computer Science And Applications*, 13.

[30] Jawale, A., Patil, K., Patil, A., Sagar, S., & Momale, A. (2023). 'Deep Learning based Money Detection System for Visually Impaired Person'. In *2023 7th International Conference On Intelligent Computing And Control Systems (ICICCS)* (pp. 173–180).

Author Biographies

Debjani Chakraborty was born in Siliguri, India, in 1982. She has received B. Tech and M. Tech in Computer Science from West Bengal University of Technology (WBUT). She is currently a research scholar in the Department of Computer Science at Maulana Abul Kalam Azad University of Technology (MAKAUT). Her research interests include image processing and deep learning.

Projjal Sahoo was born in Kolkata, India, in 2000. Currently, he is pursuing his bachelor's degree in Computer Science and Engineering from Narula Institute of Technology, West Bengal, India. During his undergraduate studies, he has been fascinated with the latest research activities in the field of Artificial Intelligence (AI). The wonders of advanced AI tools during the past decade have sparked his research interests in Computer vision, Deep learning, and Machine learning.

Argha Biswas was born in Kolkata, India, in 2001. He is currently enrolled in Narula Institute of Technology, West Bengal, India, where he is pursuing a bachelor's degree in Computer Science and Engineering. His current research interests include Computer vision, Machine learning, and Image processing. During his undergraduate studies, he has been involved in a few computer vision related research projects. The contentment of carrying out challenging research works drives him toward exploring further career opportunities in this area.

Sujaan Maitra was born in Kaity, India, in 2001. He completed his BTECH degree in Computer Science and Engineering in 2023 from Narula Institute of Technology, Agarpara. He pursued his passion for technology by starting as an intern at Amazon, followed by a role as a Senior Analyst at Ernst & Young, where he focused on using AIML algorithms to detect cyber threats. He also worked on various interface designs to host and implement Machine learning tools and worked on central government projects. Currently, Sujaan works as a SE II at Amazon. His area of research includes Machine learning, Computer vision, and Digital image processing.

Sourav Saha was born in Kolkata, India, in 1977. He is currently working as a professor in the Department of Computer Science & Engineering at Narula Institute of Technology. He worked in industry for around ten years before joining academic institutes where he has been continually gathering more than 10 years of teaching experience so far. He was awarded a Ph.D. degree in 2019 from the University of Kalyani (West Bengal, India) in Computer Science & Engineering. He obtained his Master of Engineering (M.E.) degree in Computer Science and Engineering from Bengal Engineering and Science University (Shibpur, Howrah, India) in 2002 and his graduation (B. Tech) in Computer Science & Engineering from Kalyani Govt. Engg. College (Kalyani University, West Bengal, India) in 2000. He has published many research articles in international journals and conferences. Currently, his areas of research interest include Computer vision, Soft computing, and Machine learning.

Biswajit Halder was born in Siuri, India, in 1972. He has been working in the field of security printed document analysis and digital tempering detection for the last fifteen years. He has a standing collaboration with the scientists of Computer vision and Pattern Recognition (CVPR) Unit, ISI, Kolkata, India. He was awarded Ph.D. degree from the University of Burdwan (BU). Majority of his research work revolves around the area of printed security document authentication and verification while having numerous publications in prestigious conferences like IJDAR, ICPR, ICVGIP, ICDAR, etc. His Ph.D. research work was accepted and presented on Ph D Symposium track, Springer, ICISS 2014, IDBRT (RBI Research Lab), Hyderabad, India. Presently, he is working as a professor in Hooghly Engineering and Technology College affiliated to Maulana Abul Kalam Azad University of Technology, W.B., India.