Extraction of Bank Cheque Fields Based on Faster R-CNN

Hakim A. Abdo[1,2], Ahmed Abdu[3], Ramesh Manza[1], and Shobha Bawiskar[4](✉)

[1] Dr. Babasaheb Ambedkar Marathwada University, Aurangabad, India
[2] Hodeidah University, Al-Hudaydah, Yemen
[3] Northwestern Polytechnical University, Xi'an, China
[4] Government Institute of Forensic Science, Aurangabad, India
shobha_bawiskar@yahoo.co.in

Abstract. The cheque field extraction is a critical step in automating bank cheque processing and is the first step in implementing a cheque recognition system. Many approaches for extracting the bank cheques components have been suggested. However, the complexity of the backdrop, the design variety of bank cheques, the variety of font sizes, and different patterns of writing remain a difficulty that necessitates the employment of precise algorithms. In this paper, we present a novel approach to extract the bank cheque components, in presented approach we used an innovative model called Faster R-CNN. This model represents the pinnacle of object recognition since it eliminates the need to manually extract image features and instead segments images to provide candidate region suggestions automatically. The IDRBT Cheque Image Dataset is used to train and test the Faster R-CNN model. The findings demonstrate that the model is capable of properly detecting the bank cheque fields. The extraction of bank cheque fields using Faster R-CNN achieves an accuracy of 97.4%, which outperforms other techniques.

Keywords: cheque fields · Object detection · Faster R-CNN

1 Introduction

A cheque is a document that instructs a bank to transfer a specified amount of money from a person's account to the person named on the cheque. The cheque includes printed identifying information and handwritten information. The printed fields in magnetic ink on the cheques are the number of account and bank code. While, the handwritten fields include the payee, date, courtesy amount (numerical format), the amount to be paid (legal amount), and the signature of the cheque' writer. Extract the bank cheque fields is to locate the bank cheque fields and segmenting this fields by using computer vision technologies.

The extraction of cheque fields is the critical step of automation of cheque processing. Only after completing the extraction step, the process of recognizing the payee, courtesy amount, and signature can be started. Developing an effective cheque fields extraction system, on the other hand, is a challenging undertaking, especially when the cheques

R. Manza et al. (Eds.): ACVAIT 2022, AISR 176, pp. 130–139, 2023.
https://doi.org/10.2991/978-94-6463-196-8_12

Fig. 1. Fields in a bank Cheque

have intricate and colorful backgrounds. In real applications, bank cheques may contain a variety of colorful backgrounds. In such cases, it is very difficult to find a thresholding method that will produce a satisfactory binarized image. Bank cheques do not conform to a single global standard. Even within the same country, cheques come in a variety of sizes and background colors. Cheques are issued in several sizes, and each bank offers its own collection of background drawings. On an international basis, the cheque fields are not located in the same place on the cheque, and cannot be located in terms of coordinates alone; no special box or icon is provided to help locate the cheque fields. However, a more general method is needed to extract cheque fields or for a system intended to handle international cheques, one that cannot benefit from any particular document structure.

In this paper, we present a novel approach, in this approach we extend the use of deep learning to extract bank cheque content. We use an advanced method called Faster R-CNN. Faster R-CNN is an improved version of R-CNN, which combines region proposals with CNNs. This represents the highest level in the field of object detection by now. The proposed approach is a critical step in automating bank cheque processing, it is an automatic segmentation step of bank cheque regions (Fig. 1).

2 Related Work and Overview

2.1 Related Work

Several methods for extract bank cheque fields are proposed, some of them extract handwritten fields, in [1] proposed a system to detect the handwritten items from bank cheque images; the system adjusted and evaluated for extraction of numeral format amounts in US cheques and Brazilian cheques. This method consists of three stages: arranging data into linked blocks finding possible string candidates and deciding which string best depicts the courtesy amount in [2] Proposed an approach for detecting and recognizing the fields of the signature and amount in the digital format in cheque image, in the proposed approach, the processing is divided into five parts, the courtesy amount and signature are manually selected from a preprocessed cheque image.

In [3] proposed approach for Extraction of Signatures from Cheques image, the sliding window method is used to determine the approximate area in which the signature is located. In this method, a window with adjustable height and width is dragged over the picture one pixel at a time, and the density of pixels within the window is determined.

This density is then used to calculate the entropy, which aids in fitting the box that can segment the signature.

In [4] The proposed approach used the detect lines to the extraction of significant areas from a cheque image..

A collection of heuristic rules encoded in histograms are applied to extract courtesy amount region. Approaches to extraction that make use of linked component analysis and the surrounding bounding box have also been highlighted in research works.[5].

In [6] Proposed technique uses Cartesian coordinate space to partition the cheque picture into interesting areas, at first the converting cheque image into grayscale for the purpose of decrease size, and the vertical and horizontal scanning of grayscale image for locating the AROI, finally, segmentation of the AROI regions.

Some method extract printed fields from cheque image, in [7], the proposed approach utilized cheque template structures to segment the printed items from cheque images. Template structures are determined by extracting the MICR code from the input cheque image. Important fields region is segmented, and the printed data is recognized.

2.2 Faster RCNN

Faster R-CNN [8] is a DCN model, it is using to object detection, this model is an extension of Fast R-CNN [9]. As its name implies, Faster R-CNN outperformed in speed on Fast R-CNN due to RPN.

The Faster RCNN structure consists of ConvNet layers, region proposal generation Network, region of interest pooling [11], and classifier. The structure of Faster R-CNN is shown in Fig. 2.

The Convolutional neural network is used as a feature maps extractor [10], region proposal network aim to generate the region proposal from extracted features map by sharing it with ROI pooling.

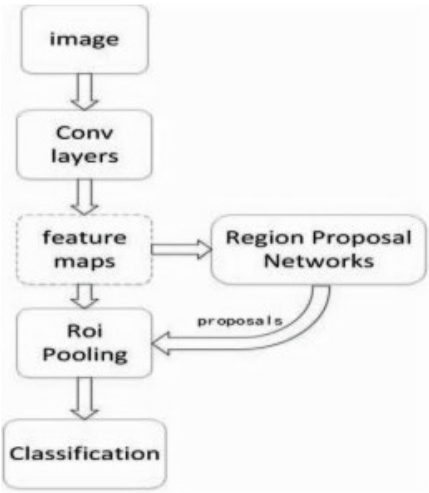

Fig. 2. Structure of Faster R-CNN

3 Methodology

The cheque fields extraction system architecture is described in this part. Figure 3 depicts the structure of cheque fields extraction system, first, the cheque image is fed to backbone convNet to extract the feature map, then the feature map extracted is fed to the PRN network for generates the region proposal, then the Roi Pooling layer uses to extract the proposal feature from feature Maps extracted, finally, the proposal feature is fed to two fully connected layers for predicting the cheque field locations and a class of filed.

3.1 ConvNet Layers

A Convolutional neural network can be used as a feature maps extractor. In the study, we use the VGG16 network as the CNN model for extraction of features map. This serves as a backbone for both the region Proposal network and convNet. The VGG-16 has 13 convolutional layers divided into five network phases by 4 max-pooling layers with relu activation function, which are used to extract various levels and scales of feature maps, we modified it to get the feature map (Feature size: 60, 40, 512) and skip the last 2 fully connected layers.

3.2 Region Proposal Networks

The RPN module [8] is in charge of creating region proposals. It makes use of the idea of attention in neural networks to direct the Fast R-CNN detection module where to seek for items in the image. To generate the region proposals, a 3x3 sliding window from the convNet layer output feature map is passed through the PRN. For each spatial point, the anchor box is sent to the regression and classification layers with varied aspect ratios and varying scales centered, as shown in Fig. 4, let k represent the number of anchors at a location. We estimate three aspect ratios and four scales, yielding $k = 12$ different anchors at a particular location. The classification layer output is a probability that the anchor box information is object or background, while the regression layer provided anchor boxes offsets.

3.3 Region of Interest Pooling Layer

Region of interest pooling layer purpose is to perform max pooling on inputs of non-uniform sizes to obtain fixed-size feature maps. The proposal suggested by RPN and

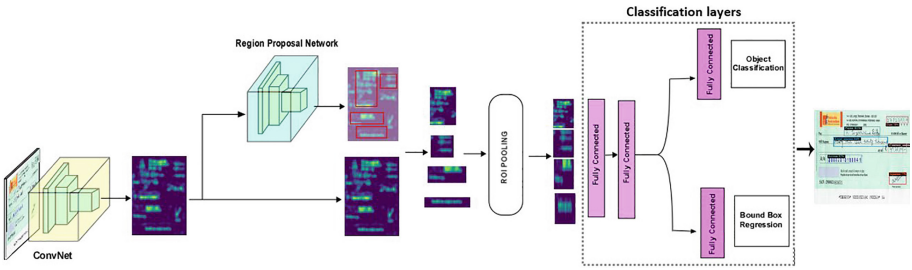

Fig. 3. Cheque fields extraction system architecture

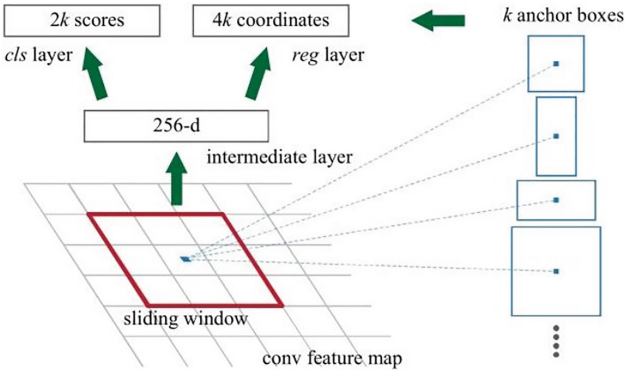

Fig. 4. Region proposal

A fixed-size feature map obtained from backbone convNet are passed to Roi pooling layer, The ROI layer divides the region corresponding to a proposal from the backbone feature map into a specified number of sub-windows, Then, max-pooling is applied to these sub-windows to get a fixed-size output.

3.4 Classification Layers

Classification layers have two fully connected layers, and two parallel sister branches, respectively, for classifying the area proposal and exact bounding box regression. The classification branch is a binary classification layer with a softmax activation function to distinguish respectively between the cheque fields and the background. Once the region proposal is expected to encapsulate cheque fields, the regression branch outputs a tetrad (x,y,w,h), indicating the exact bounding box of cheque fields in the input image based on the region proposal.

4 Experiment and Results

4.1 Dataset

The IDRBT Cheque Image dataset [15] is used in this work. We created an annotation on the dataset to suitable the faster RCNN model, partially presented in Fig. 4, the dataset consists of 112 cheque images collected from Indian banks, the size of original images is 2372×1093, the cheque images written by 9 volunteers in English language, in writing cheques the 14 pens with blue and black ink are used.

To train the Faster R-CNN on IDRBT dataset, we first resized the cheque images into 416×416 for purpose of reducing the model training time, and then we manually annotate them, as the Faster R-CNN demands. To accomplish the annotation task, we use the graphical image annotation application LabelImg designed by [16]. LabelImg is a Python program with a Qt graphical user interface; in LabelImg, the first step is pre-defining the classes of cheque bank fields, and then annotations the image of cheques by drawing a rectangular surrounding the cheque' fields with labeling them as shown in Fig. 4, finally, saving the annotation file as a JSON file.

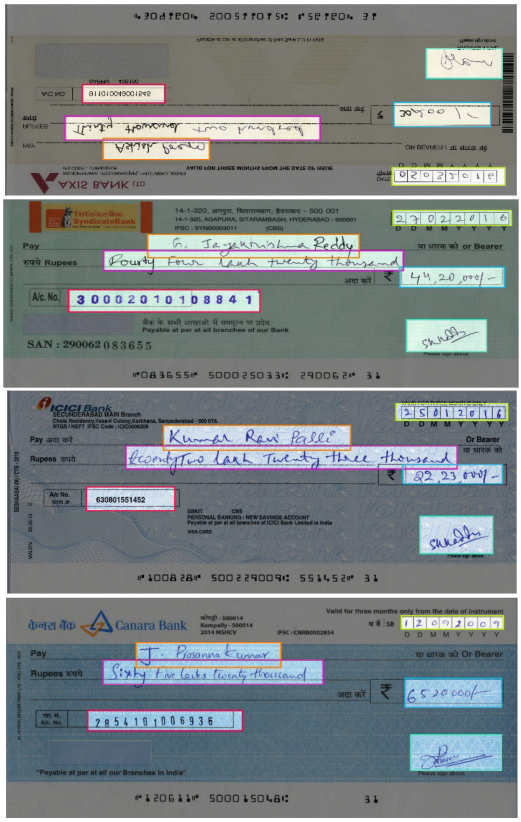

Fig. 5. Samples from modified dataset

4.2 Experiment Results

In this part, the below extraction results have been achieved using Faster R-CNN. Some samples of the model's testing results for extracting bank cheque fields are shown in Fig. 5. It can be observed that the Faster R-CNN is capable of properly detecting the Indian bank cheque fields. Table 1 shows the outcomes of utilizing Faster R-CNN retrained on our dataset, which consists of six classes of cheque fields. The dataset consists of 112 cheque images; 101 images for training and 11 images for testing; all cheques are written in English. The various font sizes and patterns of writing the cheques in the dataset increase the model's ability to extract ROI (fields) from Indian cheques that are written in English. In order to increase the ability of the proposed model to extract the ROI (fields) from Indian cheques that are written in other Indian regional languages, the dataset has to be extended to include various regional languages and retrain the model. The Faster R-CNN model performance is assessed in 2 way: classification precision and bounding box prediction precision, also known as regression accuracy. We can observe that the classification's mean Average Precision (mAP) is reached 97.4 percent. Comparison of different cheque fields extraction methods' Performance is identified in Table 2. As evidenced by the experimental findings, the Faster R-CNN surpasses

the other techniques. Figure 6 is shown the outcomes of RPN network loss. The model realizes minimal loss through the training of the model, as can be observed. In the RPN network, the minimal classification and regression losses are 0.0194 and 0.490, respectively (Fig. 7).

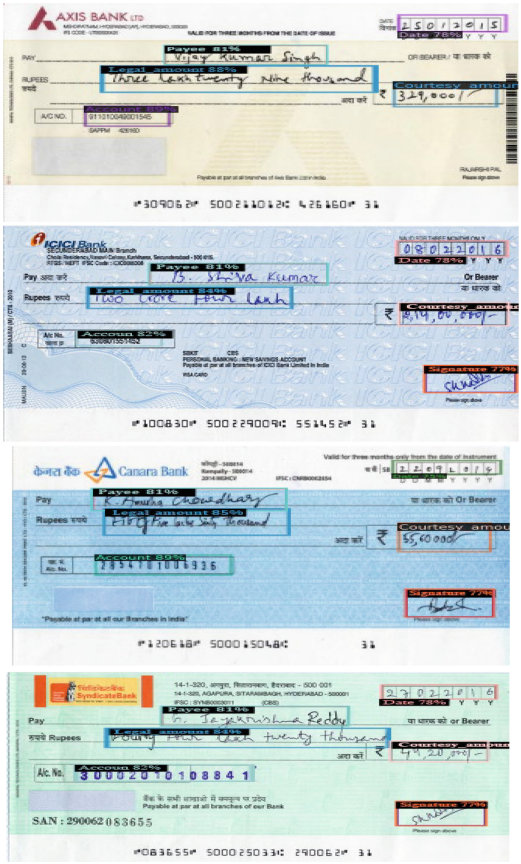

Fig. 6. Samples of cheque fields' extraction result

Table 1. Accuracy rate of classification and regression

Classes	Date	Courtesy amount	Legal amount	Account number	Signature	Time of testing(s)
Classification accuracy	1	0.99	0.91	0.98	0.99	25
Regression accuracy	1	0.99	0.93	.98	0.99	

Table 2. Comparison of different cheque fields extraction approaches' Performance

Cheque field	Method/Accuracy							
	[3]	[12]	[13]	[2]	[14]	[4]	[7]	our
Date	–	Prior knowledge position/NA	–	–	–	based on identification of important lines/NA accuracy	Template Matching method/NA accuracy	Faster R-CNN/97.4%
Courtesy amount	–		ROI based segmentation/95%	Manually selected/NA	recursive thresholding algorithm/90%			
Legal amount	–			–	–			
Payee	–			–	–			
Account No	–	Prior knowledge position/NA	–	–	–			
Signature	Sliding window method /99.26%		ROI based segmentation/91%	Manually selected/NA	–			

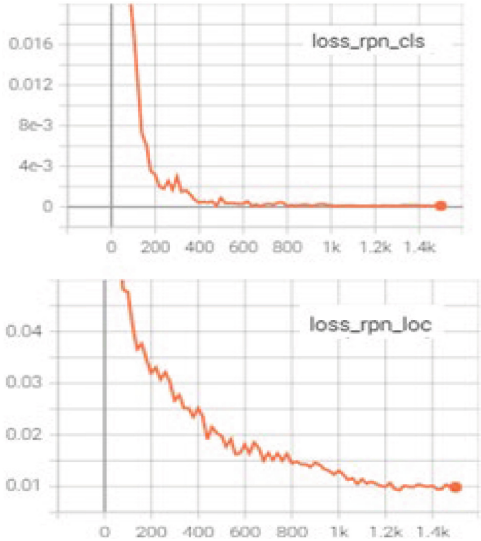

Fig. 7. The RPN network loss outcomes

5 Conclusion

In this work, we used Faster R-CNN to extract bank cheque fields, the Faster R-CNN is a novel object detection method. We modified The IDRBT Cheque Image dataset by creating an annotations to suitable the faster RCNN model. The Faster R-CNN model trained and tested on 112 bank cheque images. The model achieves a mAP value that is 97.4%.The outcomes we got shows that the Faster R-CNN model is capable of properly extracting the bank cheques fields.However, as compared to previous techniques, deep learning improves the effectiveness of the detection method.

References

1. R. Palacios and A. Gupta, "A system for processing handwritten bank checks automatically," *Image Vis. Comput.*, vol. 26, no. 10, pp. 1297–1313, 2008, https://doi.org/10.1016/j.imavis.2006.04.012.
2. M. B. Alam and S. Mia, "Handwritten Courtesy Amount and Signature Recognition on Bank Cheque using Neural Network," vol. 118, no. 5, pp. 21–26, 2015.
3. V. K. Madasu, M. Hafizuddin, M. Yusof, and M. Hanmandlu, "Automatic Extraction of Signatures from Bank Cheques and other Documents," pp. 10–12, 2003.
4. P. Dansena, K. Pramod Kumar, and R. Pal, "Line based extraction of important regions from a cheque image," in *2015 Eighth International Conference on Contemporary Computing (IC3)*, 2015, pp. 183–189, https://doi.org/10.1109/IC3.2015.7346676.
5. L. Chandra, N. Software, T. Limited, S. V. Marg, and Q. I. Area, "Automatic Courtesy Amount Recognition for Indian Banks ' Checks," pp. 6–10, 2008.
6. M. Sankari and M. Benazir, "Verification of Bank Cheque Images using Hamming Measures," no. December, pp. 7–10, 2010.
7. P. M.P, "Automatic Extraction of Attributes from Printed Indian Cheque Images by Template Matching Technique," vol. 5, no. 1, pp. 30–37, 2019.
8. S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 6, pp. 1137–1149, 2017, https://doi.org/10.1109/TPAMI.2016.2577031.
9. R. Girshick, "Fast R-CNN," *Proc. IEEE Int. Conf. Comput. Vis.*, vol. 2015 Inter, pp. 1440–1448, 2015, https://doi.org/10.1109/ICCV.2015.169.
10. S. Bell, C. L. Zitnick, K. Bala, and R. Girshick, "Inside-Outside Net: Detecting Objects in Context with Skip Pooling and Recurrent Neural Networks," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 2016-Decem, pp. 2874–2883, 2016, https://doi.org/10.1109/CVPR.2016.314.
11. C. Szegedy *et al.*, "Going deeper with convolutions," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 07–12-June, pp. 1–9, 2015, https://doi.org/10.1109/CVPR.2015.7298594.
12. P. Agrawal *et al.*, "Automated bank cheque verification using image processing and deep learning methods," *Multimed. Tools Appl.*, vol. 80, no. 4, pp. 5319–5350, 2021, https://doi.org/10.1007/s11042-020-09818-1.
13. A. K. Talele, "Automatic Extraction of Legal and Courtesy amount , Payee Name and signature in Bank Cheque Processing System," vol. 3, no. 5, pp. 4417–4425, 2011.
14. C. Y. Zhang, Q. Chen, Z. Lou, and J. Y. Yang, "Extraction of courtesy amount item from Chinese check," *2004 8th Int. Conf. Control. Autom. Robot. Vis.*, vol. 1, no. December, pp. 109–114, 2004, https://doi.org/10.1109/icarcv.2004.1468807.
15. Dansena, P., S. Bag, & R. Pal. (2017). Differentiating Pen Inks in Hand-written Bank Cheques Using Multi-Layer Perceptron. *Proc. of 7th International Conference on Pattern recognition and Machine Intelligence, Kolkata, India.*
16. Tzutalin. (2015). labelImg:A graphical image annotation tool. Retrieved from https://github.com/tzutalin/labelImg