S	
E1	
E2	
For Office Use Only	

Masters Project Final Report

(MCS)

2019

Project Title	Methodology to authorize bank cheque automatically
Student Name	G. H. M. C. Senanayaka
Registration No. & Index No.	2017/MCS/074 17440747
Supervisor's Name	Dr. M. G. N. A. S. Fernando

For Office Use ONLY

Methodology to Authorize Bank Cheque Automatically

A dissertation submitted for the Degree of Master of Computer Science

G. H. M. C. Senanayaka
University of Colombo School of Computing
2019

UCSC

1. ABSTRACT

This study shows the automation of a simple cheque captured by a camera or a scanner.

The bank cheque is a declining bank instrument. The cheque is still one of the payment methods in Sri Lanka. It is mostly used by the corporate community. While I am studying the research, I have contacted several people to inquire about cheques and its realization process. Unfortunately, most of them are not sentient of the check realization process or details printed on the cheque.

In-order to possess cheque issuing facility, a person should have a cheque account where the bank is offering a cheque book. There are several areas in a cheque for its realization process. Those areas are date, pay, bearer, amounts in words and amounts in figures, signature, and magnetic ink area. Each of these areas plays a salient role in cheque realization process.

In the date area, cheque issued date is mentioned. Generally a cheque is valid for 6 months of period. Therefore if a cheque issued date is more than 6 months, it is considered as an expired cheque. Amounts in words and amounts in figures should match with each other, otherwise it is an invalid cheque. Also, the issuer signature should be there on the cheque signature area. The signature should validate with the signature stored in the bank's database. If there is a miss-match, the cheque is consider as an invalid cheque. Bank systems can identify unique chque number, branch of the cheque account, cheque account number by using magnetic ink are details. From the pay area, bank officer can identify the payee details (Cash cheque).

Currently, above-mentioned processes have been done manually by bank officers [23]. According to the cheque clearance type, it can change the time taken to clear the cheque.

There are three cheque clearance types. Inward cheques, outward cheques and transfers. Inward cheques are cheques a bank received from a 3rd party cheque clearance organization. Transfers are a bank received its cheques by customers. Outward cheques are cheques that a bank received other bank cheques from its customers. Inward cheques and transfers are done by the bank while transferring outward cheques to a 3rd party clearance organization.

In this study, when identifying the areas, first the cheque area will be identified and extracted from the given input image; then, the system will remove noise and resize the image for a defined size and crop key areas to process the optical character recognition and comparison. For noise removal, it is using Gaussian filter while the Tesseract Optical Character Recognition Engine is used to read English characters in the cheque. Also, it is using Scale Invariant Feature Transform algorithm to identify signature on the cheque. By applying optical character recognition and comparison techniques, the realization process is automated.

The accuracy of the system for printed character identification is 90%, handwritten character identification is 60%. Signature identification accuracy by using image comparison techniques is 70%.

2. DECLARATION

The thesis is my original work and has not been submitted previously for a degree at this or any other university/institute.

To the best of my knowledge it does not contain any material published or written by another person, except as acknowledged in the text.

Student Name: G. H. M. C. Senanayaka

Registration Number: 2017/MCS/074

Index Number: 17440747

Signature: Date:

This is to certify that this thesis is based on the work of

Mr./Ms.

Under my supervision. The thesis has been prepared according to the format stipulated and is of acceptable standard.

Certified by:

Supervisor Name:

Signature: Date:

3. ACKNOWLEDGEMENT

I am grateful to my supervisor Dr. Noel Fernando of the University of Colombo, School of Computing for his valuable support to make my dissertation and the research project a success. The door to Dr. Fernando's office was always open whenever I was in a trouble or had a question about my research or writing. He consistently allowed this paper to be my own work, but directed me in the right the direction whenever he thought I needed it. His great guidance, well-organized meetings, excellent dissertation ideas and confidence pointed me to come up with a successful research project.

Finally, I must direct my deep gratitude towards my parents for providing me with consistent support and continuous encouragement throughout my years of study and through the completion of researching and writing this thesis. This achievement would not have been conceivable without them. Also, I thank all those whose names, though not mentioned for their support and encouragement in completing this project. Thank you.

4. TABLE OF CONTENTS

Contents

5. LIST OF FIGURES

6. LIST OF ABBREVIATION

1. INTRODUCTION

1.1.MOTIVATION

Recently, it is noticed that financial institutions are rapidly moving to the digital era by using modern technology. Also, as a matter of fact, nowadays most software companies invest by moving to the bank and other financial functions. With the development of technology, banks use digital currencies, BlockchainATechnology, Automated Financial Services, Mobile and Digital Banking, Upgraded ATMs and Wearable.

Nowadays in Sri Lanka, in some technological leading banks, the concept of digital banking/digital branches have come into the picture. In-order to move all banking functions to a digital platform. Banks in Sri Lanka are doing different research areas such as Robotic Process Automation, Automatic Teller Machines, Cash Deposit kiosk, Virtual Teller Machines, Chat Bots.

One of the good examples of digital banks can survive in any kind of circumstances is that the most people all over the world are browsing for online banking applications to satisfy their basic needs with the effect of COVID-19 pandemic. Other than the technological era, this outbreak also influenced the essence of the digital platform for any bank in Sri Lanka as well as other countries in the world.

For a period, customers are suffering from cheque clearance, because it takes several days to clear a cheque after a complex manual process. Therefore, there is a need for an automated and faster cheque clearance process for banks in Sri Lanka.

1.2.STATEMENT OF THE PROBLEM

Cheque details, cheque holder details, date and amounts validation of a bank cheque is a manual process in financial institutions in Sri Lanka.

A bank draft is also a type of cheque that is issued by a financial institution on behalf of a person. When a person wants to receive money from another individual in a formal and a trusted way, that person can request the sender to provide it via a bank draft. The person who is sending the bank draft can visit the nearest bank, pay the required amount to a cashier and can request a bank draft for the amount. The bank will issue a cheque/draft mentioning the amount while placing an authorized signature on it. This type of cheque is called a bank draft [7].

After receiving the draft, the person can submit this draft to any nearest bank and change it to the cash. When changing a bank draft to money, bank officers will authorize the signature manually by comparing signatures with the help of a signature authorization system provided by other financial institutions.
In order to realize a cheque, sufficient balance should be there in the current account holder's account. Therefore, currently, the amount validation is done manually by bank officers [23].

Since this cheque details, cheque holder details, date and amounts in both the figure and words validation on a bank draft or a cheque is a manual process and plenty of drafts and cheques are being received by financial institutions per day, it is an essential requirement to have an automated signature and amount validation system for drafts/cheques.

1.3.AIMS AND OBJECTIVES

- Capturing the image of a cheque as a JPG format and recognize the related information (handwritten and typed) using image processing techniques
- Validating the content of the cheque (Cheque details, cheque holder details, date, amounts in figures and words)
- Providing a solution to automate the cheque realization process

1.4.PROJECT SCOPE

- Only the cheque realization is considered
- An image of the bank cheque with good resolution (1500*500px) is needed for the process (Scan image of a cheque is preferred since banks are focusing more on that)
- If the input image is a captured image by using a camera, it should be capture by keeping the cheque on a white A4 sheet or with white background.
- Only English language will be considered in the study (Sinhala and Tamil will not be considered)
- Only simple good hand writing will be considered (combined hand writing not considered)
- Payable party's signature is out of this scope and will not be considered
- Only non-crossed inward cheques will be considered (Crossed cheques will not be considered)
- Only outward cheque clearance and transfer clearance will be considered
- Image processing, natural language processing and machine learning techniques will be used in this study
- Only a single signature validation is considered

1.5.MACHINE LEARNING

This is a sub domain of Artificial Intelligence where a system self-learn itself without programming or guiding explicitly. By self-learning it improves the accuracy of the system's output for given inputs [11]. Machine learning algorithms build mathematical models according to the given sample data which is also known as sample data. This sample data will train the algorithms to do better predictions or decisions.

Machine learning is generally divided into three main categories considering the learning style,

I. Supervised Learning [12]
II. Unsupervised Learning [12]
III. Semi-supervised Learning [12]

Other than these three approaches, there is another approach called reinforcement learning approach, which interacts with a dynamic environment to learn the algorithm with those environment inputs.

1.6. OPENCV

OpenCV stands for Open Source Computer Vision Library. As for the definition, this is an open source project that has more than 2500 optimized algorithms, which contains a comprehensive set of both classic and state-of-the-art computer vision and machine learning algorithms. These algorithms can be used to detect and distinguish faces, recognize objects, classify human actions in videos, trail camera movements, trail moving objects, extract 3D models of objects, generate 3D point clouds from stereo cameras, stitch images together to generate a high resolution image of an whole scene, find related images from an image database, eliminate red eyes from images taken using flash, track eye movements, identify scenery and establish markers to overlay it with augmented reality [13].

1.7. TESSERACT

Tesseract is an open-source Optical Character Recognition (OCR) engine, which can recognize more than 100 languages. Also, it supports Unicode (UTF-8). It can be trained to recognize other languages as well. This engine is used machine learning algorithms and OpenCV techniques to recognize characters using the training dataset. Tesseract is developed by using C++ and built to different language libraries such as Python, Java. Currently, Tesseract 4 is released, where there is a separate repository for Tesseract [14]. It is using Tesseract 4.2 JAR file with non-free algorithms.

1.8. GAUSSIAN FILTER

In image processing applications, the Gaussian distribution needs to be estimated by a convolution kernel. Therefore, values from this distribution are used to form a convolution matrix then applied to the original image. Each pixel's novel value is a weighted average of that pixel's neighborhood. Thus, the original pixel's value obtains the heaviest weight (having the highest Gaussian value) and adjacent pixels receive smaller weights as their distance to the original pixel increases [15].

1.9. EDGE DETECTION

There are different algorithms to detect edges on an image such as Sobel Edge Detection algorithm, Canny Edge Detection Algorithm. In Canny edge detection algorithm, there are three main criteria taken into account [16],

I. Detect all the significant edges in the source image
II. Edge points to be detected as close as probable to the true edge
III. Not to have over one response to a single edge

1.10. HOUGH TRANSFORM

There are plenty of Hough transform application areas such as, traffic and transport applications, biometrics and man-machine interactions, 3D applications, object tracking, under water applications and medical applications [16].

2. LITERATURE REVIEW

Several studies had been conducted on identifying handwritten digits [2], [3], [4], characters [5], [9] and image similarity comparison [6], [7] using image processing techniques such as machine learning, artificial neural networks and image comparison in domains of shark fish classification, tilted handwritten character recognition and improve the accuracy of handwritten character recognition.

Literature reveals that the Mixed National Institute of Standards and Technology (MNIST) handwritten digits data set [1] is the most popular and the widely used dataset used by researchers to research on this domain. MNIST dataset has used in research conducted to analyze character recognition patterns. There are two different types of character recognition identified by previous studies, which are online character recognition and offline character recognition. In the online character recognition, characters are recognizing while writing; and in the offline character recognition, it identifies the characters from a written paper or an article or a paper [2]. Within the two categories aforementioned, it noticed that there are plenty of techniques to recognize handwritten numbers where each of them has its advantages and disadvantages. Some algorithms used for character recognition such as K-nearest neighbor (K-NN), random forest and decision tree have higher accuracy while its processing speed is relatively low. In addition to this, most researchers conducted their researches by using machine learning and artificial neural network techniques. It identified that the classifier accuracy increased with the increment of the training dataset. K-NN has the accuracy improvement of 96.7% with compared to the Neural Network, which is having the 96.8% improvement. K-NN is ten times faster than the Neural Networks, where its efficiency is higher [2].

In the research conducted by Zhao, K., three main parts were identified in Optical Character Recognition process and they are image pre-processing, feature extraction and classification. When it comes to recognition of handwritten digits, slant correction or elastic distortion have a considerable effect on the feature selection of the sample. Slant correction is considered as effective because different people's handwriting is tilted/skew. In order to reduce the dimension of data and to extract the relevant information, some researchers used Principal Component Analysis and Histogram of Oriented Gradients were used. Pre-processing and feature extraction were performed properly in the research to reach the highest accuracy in recognizing handwritten digits [3].

The heart of the problem in recognizing handwritten digits lies within the capability to develop an efficient and effective algorithm that can distinguish handwritten digits, and which is submitted by users with the use of a scanner, tablets, and other digital devices. According to the study [4], the number of machine learning algorithms such as Multilayer Perceptron, Support Vector Machine, Naïve Bayes, Bayes Net, Random Forest, J48 and Random Tree has used for the recognition of digits. The result shows that the highest accuracy of 90.37% has obtained for Multilayer Perceptron [4]. One of the challenges in handwritten character recognition wholly lies in the distinction and distortion of handwritten character set because different communities may use various style of handwriting and control to draw the similar pattern of the characters of their familiar writing. Handwritten character's lines are not straight as the printed characters and curves are not smooth like printed characters. Therefore, suitable feature extraction techniques should be used to address the problem [4].

Image comparison is one of the main processes in the field of image processing. At the same time, it is essential to compare two images to approximate the similarities and dissimilarities among them. In image comparison, there are several methods and tools to compare images by using image processing techniques. Pixel by pixel comparison, comparing images using the Hausdorff distance, Distance-based functions for image comparison, Image Change Detection Algorithms [5] are some of the methods and ImageMagick, Perceptual Diff, Image Comparer, Image Diff, ImageJ, OpenCV are some of the tools which were discussed in the research done by Katuam, R. [5].

It was mentioned in the paper [6], the newest handwritten character recognition algorithm with newly designed network structure proposed. According to the study, even though handwritten character recognition based on deep networks which significantly increases time consumption in parameter training, it has higher accuracy than the traditional methods. Currently, there are two main methods to recognize handwritten characters [6]. They are, handwritten character recognition methods based on conventional feature extraction and techniques that originate from deep learning. Traditional methods consist of three processes which are image pre-processing, feature extraction of characters and character classification [6].

Template matching-based, k-nearest neighbour (K-NN) algorithm, shallow artificial neural network algorithm, and support vector machine are some commonly used character classification methods. Results cannot meet the needs of practical applications due to low processing capacity and non-clear model in traditional classification methods [6]. The results obtained from the convolutional neural network and classification methods based on the residual network are superior with compared to conventional methods. However, current deep-learning methods also present some problems, such as time complexity issues that have not yet been resolved [6].

One of the most important steps in feature extraction is image segmentation. Image segmentation means dividing the image into parts where it is helpful for further processing by reducing the information for analysis. According to research conducted by R.Yogamangalam and B.Karthikeyan, there are several image segmentation techniques in image processing field with their advantages and disadvantages. Some of them are suitable for noisy images where some consume pre-processed images with noise reduction [7].

Segmentation can do either by using a binary image or RGB colour image. Edge-based segmentation is one of the widely used image segmentation technique. Edges are detected to identify the discontinuities in the image [7]. Edges on the region traced by identifying the pixel value and it compared with the neighbouring pixels. Famous algorithm for edge-based image segmentation is the Canny Edge Detector Algorithm. Its process is as follows [7],

I. Gaussian filter is used to reduce the effect of noise.
II. Sobel Operator is applied to the image to detect the strength and the edge directions.
III. The edge directions are taken into account for non-maximal suppression.
IV. Removing broken edges.

In edge detection, it is necessary to point out the true edges to get the best results from the matching process. Classical edge detection techniques have the primary advantage of simplicity compared to other techniques. Another advantage is that classical operators are detecting edges and their orientations [8]. The main disadvantage of this is that it is very much dependent on the noise of the image. Advantage of Canny edge detection technique is improving the signal with respect to the noise ratio and this is established by Non-maxima suppression method as it results in one-pixel wide ridges as the output. It has better detection of edges especially in noise state with the help of the thresholding method [8].

Efficiency of the Canny Edge Detection is higher, it is well fit with simple images and complex images as well and higher signal to noise ratio than the other edge detection [16].

Parameters	Sobel	Canny
Computation[5]	Simple and time efficient	Complex and time consuming
Signal to Noise ratio	Low	High
Texture based image Fig.2	Less efficient	More efficient
No of objects in image Fig.3, Fig.4	Suitable for simple images	Suitable for simple as well as complex images
Application area /domain [5] [6]	Massive data communication and data transfer	Medical field for X-ray diagnosis and object recognition

Figure 1: Comparison between Sobel Edge Detection and Canny Edge Detection [16]

A survey conducted by Monica Patel and Shital P. Thakkar has identified that there are four methods of cursive (joined characters) handwritten word recognition [9].

I. Holistic Approach
II. Segmentation Based Approach
III. Recognition Based Segmentation Approach
IV. Mixed Approach

According to the survey, in the image preprocessing step, there are several operations such as noise removal, binarization, skew correction performed to the scanned image. In the segmentation step, several segmentation methods like line segmentation, word segmentation and character segmentation are performed [9]. In feature extraction, statistical, structural and global transformation feature extractions performed. Supervised and unsupervised classifiers used to classify an image. Even though they have done immense work and research to identify unconnected handwritten characters, 100% accuracy is not achieved [9].

According to literature, it can identify that artificial neural network has more accuracy than the algorithm K-NN, so, therefore, it is better to use artificial neural networks to authorize bank cheque since it considers accuracy rather than the speed. In addition to this, current neural networks and deep learning algorithms have more precision than traditional methods for image classification. Methods such as slant correction are useful because most people's handwriting skewed. According to the research [7], it can be improving the accuracy when the image is pre-processed by removing noise. To do that the Gaussian filter can apply to the high-quality image. Edge detection algorithms also can be applied to increase the accuracy of the handwritten character identification. There are plenty of image comparison techniques available to compare the human signature on the cheque. Some algorithms are Pixel by pixel comparison, comparing images using the Hausdorff distance, Distance-based functions for image comparison, Image Change Detection Algorithms.

When it comes to OCR engines, Tesseract is the better and free open source OCR engine compared to GOCR and Transym OCR [21, 22].

Accuracy of identifying different types of fonts, different resolution levels, different brightness levels and performance speed is better in Tesseract with compared to other open source OCR engines [21, 22]

Comparison between Tesseract and GOCR is as follows,

IMAGE TYPE	TOTAL CHARACTERS EXTRACTED BY TESSERACT	CHARACTERS CORRECTLY EXTRACTED BY TESSERACT	TOTAL CHARACTERS EXTRACTED BY GOCR	CHARACTERS CORRECTLY EXTRACTED BY GOCR	TESSERACT ACCURACY (IN %)	GOCR ACCURACY (IN %)	TESSERACT PRECISION (IN %)	GOCR PRECISION (IN %)
Color	39	38	28	25	97.4	64.1	97.4	89.2
Gray scale	39	38	24	18	97.4	46.1	97.4	75.0
Black and White	39	38	27	19	97.4	48.7	97.4	70.3

Figure 2: Comparison between Tools using different Image types [21]

BRIGHTNESS	TOTAL CHARACTERS EXTRACTED BY TESSERACT	CHARACTERS CORRECTLY EXTRACTED BY TESSERACT	TOTAL CHARACTERS EXTRACTED BY GOCR	CHARACTERS CORRECTLY EXTRACTED BY GOCR	TESSERACT ACCURACY (IN %)	GOCR ACCURACY (IN %)	TESSERACT PRECISION (IN %)	GOCR PRECISION (IN %)
25	39	37	28	23	94.8	58.9	94.8	82.1
50	39	38	27	26	97.4	66.6	97.4	96.2
100	39	37	1	1	94.8	02.5	94.8	100

Figure 3: Comparison between Tools using different Brightness values [21]

FONT	TOTAL CHARACTERS EXTRACTED BY TESSERACT	CHARACTERS CORRECTLY EXTRACTED BY TESSERACT	TOTAL CHARACTERS EXTRACTED BY GOCR	CHARACTERS CORRECTLY EXTRACTED BY GOCR	TESSERACT ACCURACY (IN %)	GOCR ACCURACY (IN %)	TESSERACT PRECISION (IN %)	GOCR PRECISION (IN %)
Arial	39	37	7	0	94.8	0	94.8	0
Roman	39	38	6	2	97.4	05.1	97.4	33.3
Tahoma	39	37	3	0	94.8	0	94.8	0

Figure 4: Comparison between Tools using different Font Types [21]

IMAGE TYPE	RESOLUTION	TOTAL CHARACTERS EXTRACTED BY TESSERACT	CHARACTERS CORRECTLY EXTRACTED BY TESSERACT	TOTAL CHARACTERS EXTRACTED BY GOCR	CHARACTERS CORRECTLY EXTRACTED BY GOCR	TESSERACT ACCURACY (IN %)	GOCR ACCURACY (IN %)	TESSERACT PRECISION (IN %)	GOCR PRECISION (IN %)
Color	75	39	38	39	38	97.4	97.4	97.4	97.4
Color	300	39	38	–	–	97.4	–	97.4	–
Color	1200	0	0	–	–	–	–	–	–
Gray scale	75	–	–	39	38	–	97.4	–	97.4
Gray scale	300	–	–	–	–	–	–	–	–
Gray scale	600	39	38	4	2	97.4	05.1	97.4	50.0
Black and White	75	39	38	39	38	97.4	97.4	97.4	97.4
Black and White	300	39	38	–	–	97.4	–	97.4	–
Black and White	1200	–	–	–	–	–	–	–	–

Figure 5: Comparison between Tools using different Resolution values [21]

Comparison between Tesseract and Transym is as follows,

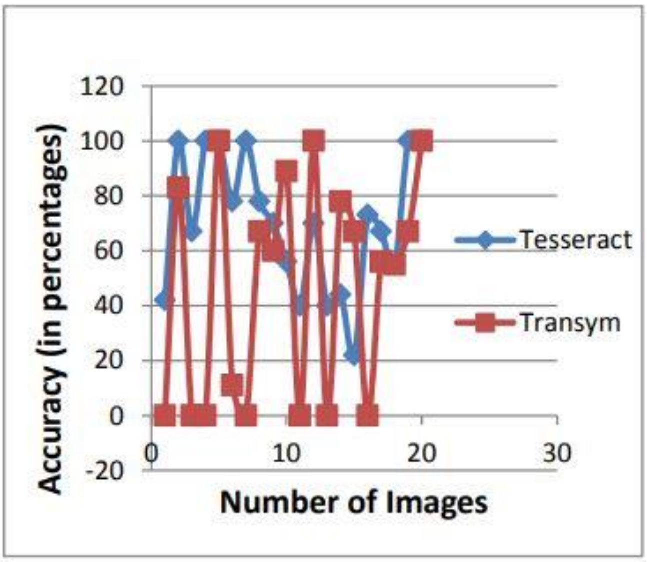


Figure 6: Comparison of Accuracy between Tesseract and Transym OCR [22]

Feature	Tesseract OCR	Transym OCR
Free	Yes	No (Trial Version is available)
Open Source	Yes	No
License	Apache	Proprietary
Online	No	No
Operating System	Window, MaC, Linux	Windows
Latest Stable version	3.01	3
Release Year	2010	2008
DLL Available	Yes	No
Accuracy (For extracting character from vehicle number plate)	61% (color images) 70% (gray scale images)	47%
Average Time	1 second (color images) 0.82 Seconds gray scale images)	6.75 seconds
σ_{AC}	34.21(For color Images) 24.64(For Gray Scale Images)	40
σ_T	0.63 (For color images) 0.61(For gray scale images)	12

Figure 7: Feature Comparison of Tesseract and Transym OCR [22]

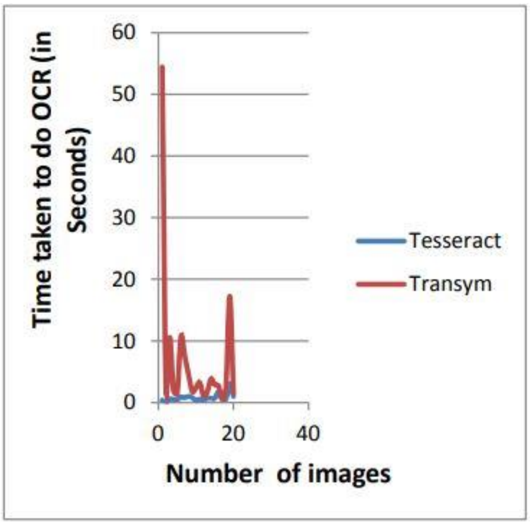


Figure 8: Comparison of Performance between Tesseract and Transym OCR [22]

There are plenty of algorithms and OCR engines to identify printed English characters, while human signature is far more different from the OCR. Therefore, image feature extraction and feature mapping techniques are more suitable to identify human signatures and to identify forged signatures [19].

Among the two feature extraction algorithms SURF and SIFT, SURF algorithm is used in this study. Because SURF algorithm's accuracy and processing speed is higher than the SIFT algorithm [20].

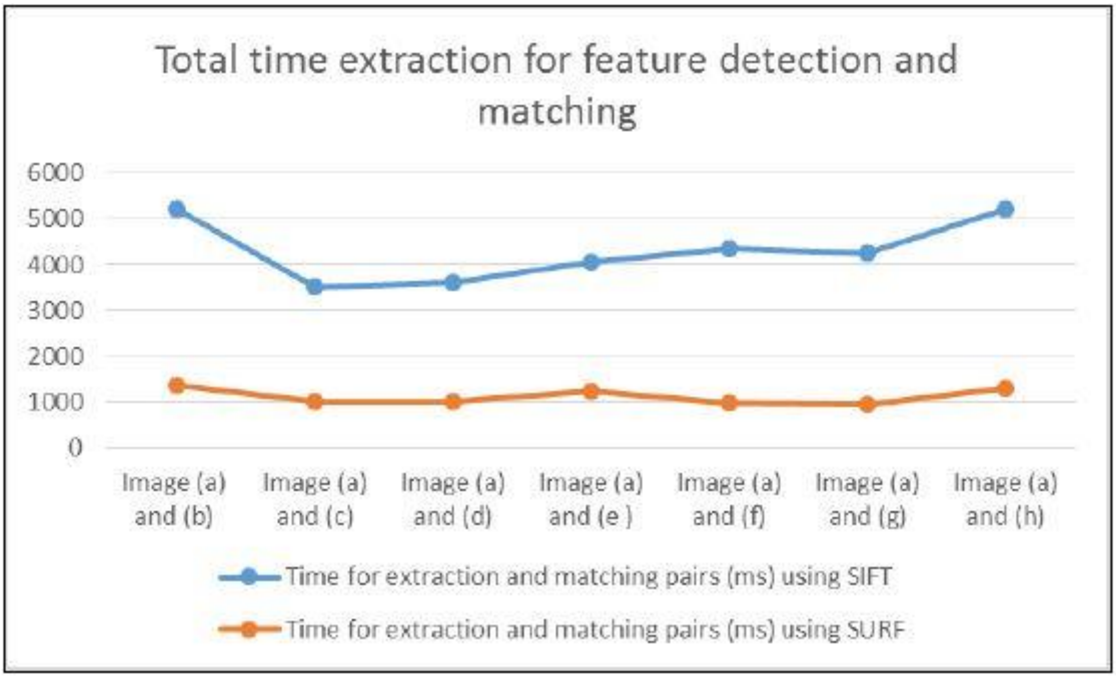


Figure 9: SURF and SIFT Time taken to Extract and Match Features of Images [20]

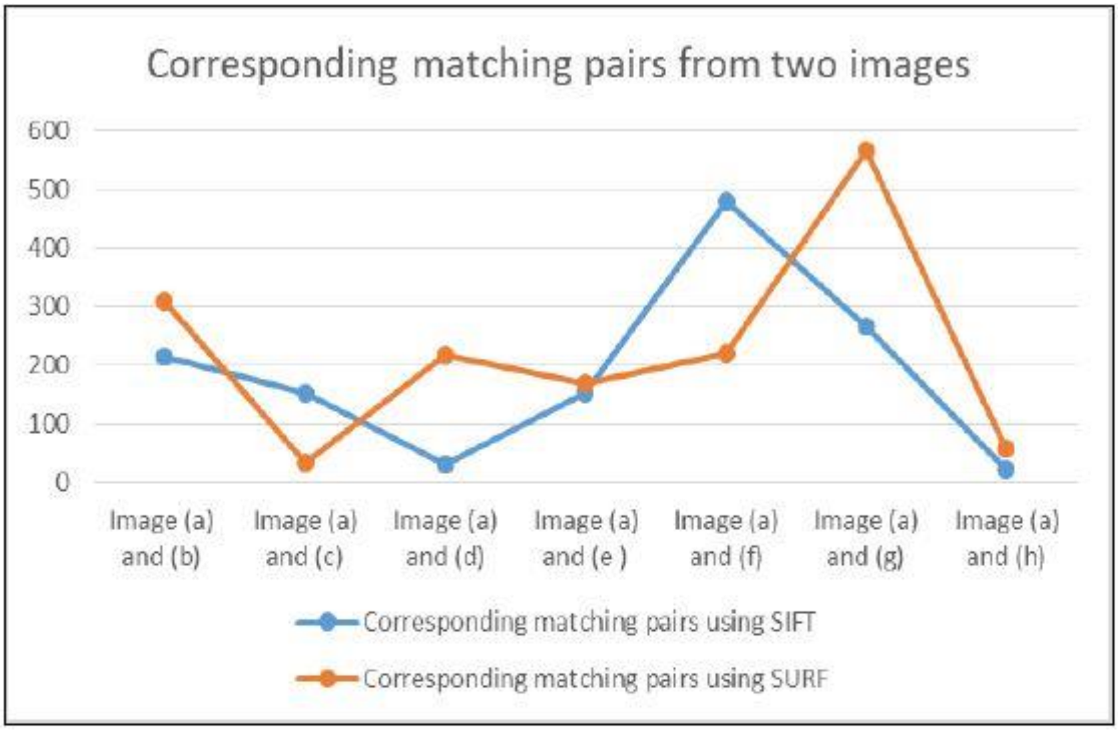


Figure 10: Corresponding Matching Pairs Using SIFT and SURF Algorithms [20]

Algorithm	*Rotation*	*Scale*	*Blur*	*Illumination*	*Warp*	*RGB noise*	*Time cost*
SIFT	*Good*	*Better*	*Good*	*Good*	*Good*	*Good*	*Good*
SURF	*Better*	*Good*	*Better*	*Good*	*Better*	*Better*	*Better*

Figure 11: Comparison of SIFT and SURF [20]

3. METHODOLOGY

In this section, the process which will be followed to create an application to automatically authorize a bank cheque is explained.

There are several important areas in a cheque to identify before understand the cheque realization process. Those areas are,

I. Bank Logo
Indicates the bank of the cheque
II. Date
Indicates the issued date of the cheque
III. Pay Area
Indicates to whom the cheque is paid or the cash cheque
IV. Amounts in Words Area
Indicates amounts in words
V. Amounts in Figure Area
Indicates amounts in figures
VI. Current Acc Number and Name Area
Indicates the cheque owner current account number and the name
VII. Signature Area
Indicates the cheque owner signature
VIII. Magnetic Ink Area
Indicates several details such as cheque number, current account number of the cheque

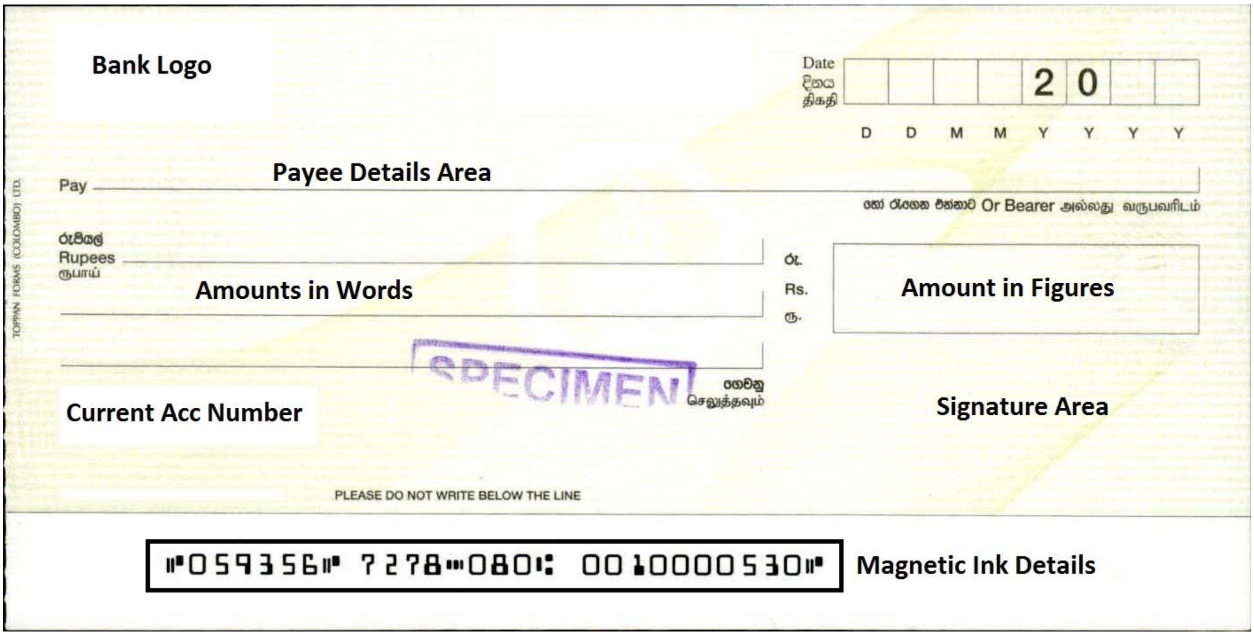


Figure 12: Sample Image of a Cheque with Areas

This study uses Java 1.8 as the programming language for the implementation. Further, it uses OpenCV library version 4.2 which is built for Java with the non-free part. In non-free part it consists of patented feature detection algorithms such as SURF and SIFT.

3.1.PREPROCESSING

Initially, the state-of-the-art in capturing the cheque area from the camera image will be explained.

As for the first step in this activity, the system reads the image by converting the input image (Figure 13) to a grayscale image. Extract the cheque area from the input image, in the preprocess step, first it detects all edges in the input image (Figure 14). To perform this, Canny's edge detection method is being used [16]. After that, edge detected image again converts to a RGB image and perform the Standard Hough transform to detect the lines in the image. Then probabilistic Hough line transform used to identify the real margins (Figure 15) with the end points of the cheque area [16]. After identifying all the lines in the image, it will again identify the outlines of the cheque and then crop (Figure 16).

Figure 13: Captured Original Input Image

Figure 14: Standard Hough Transformed Image

Figure 15: Probabilistic Hough Transformed Image

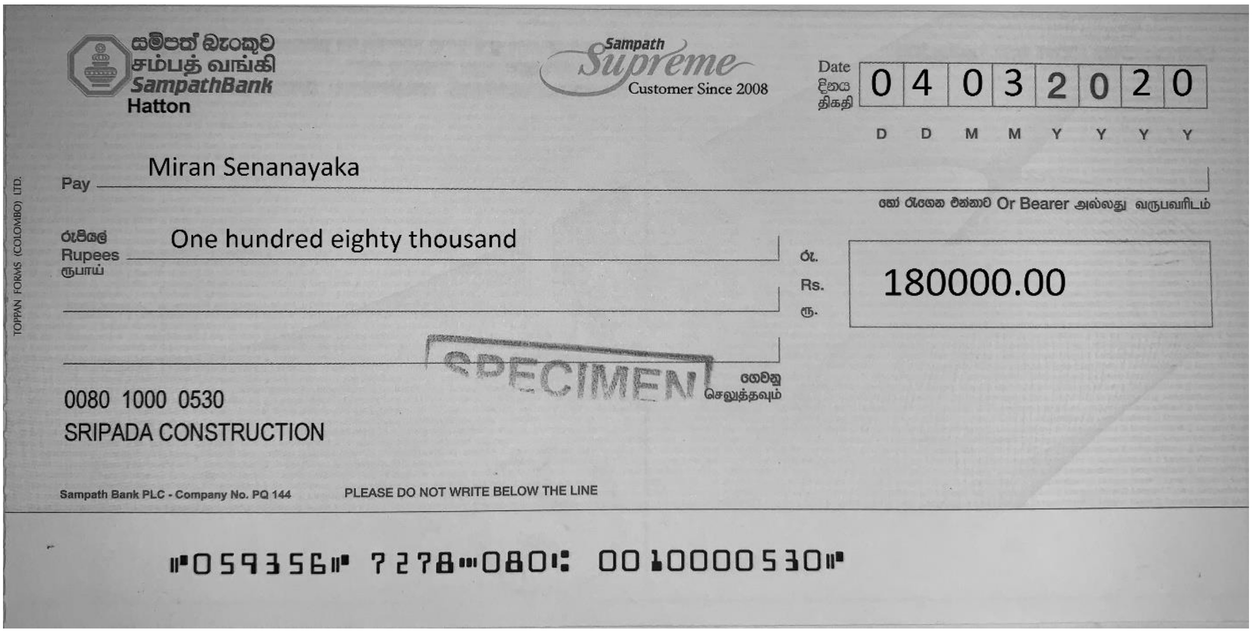


Figure 16: Final Image after Extracting the Image

After cropping, the image is resized to pre-defined width and height. Then the image again crops to 3 parts,

I. Top part of the cheque
This part contains, bank logo and the issued date of the cheque (Figure 17)

Figure 17: Top Part of the Cropped Image

II. Middle part of the cheque
This part contains, payee details, amount in words, amounts in figure and the signature (Figure 18)

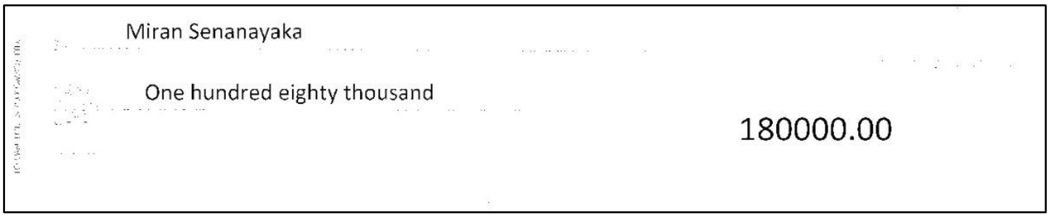


Figure 18: Middle Part of the Cropped Image

III. Bottom part of the cheque

This part contains, magnetic ink details (Figure 19)

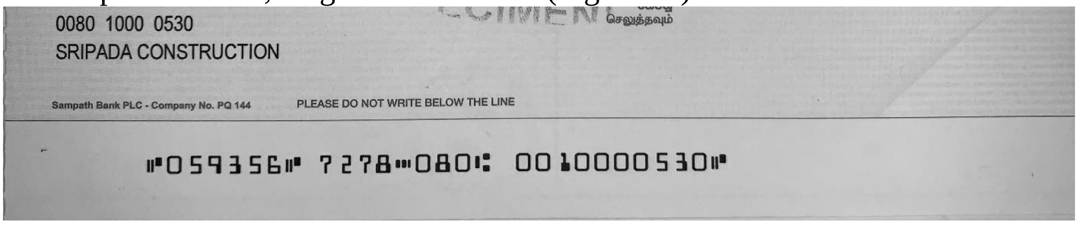


Figure 19: Bottom Part of the Cropped Image

Each of these parts will enhance according to pre-defined threshold values. After enhancing each part, these three parts combine vertically from top to bottom. Finally, this combined image (Figure 20) converts to a grayscale image and remove the noise in the image. Later removing the noise, it converts to a binary image. Finally the image is ready to process area extraction and for the OCR.

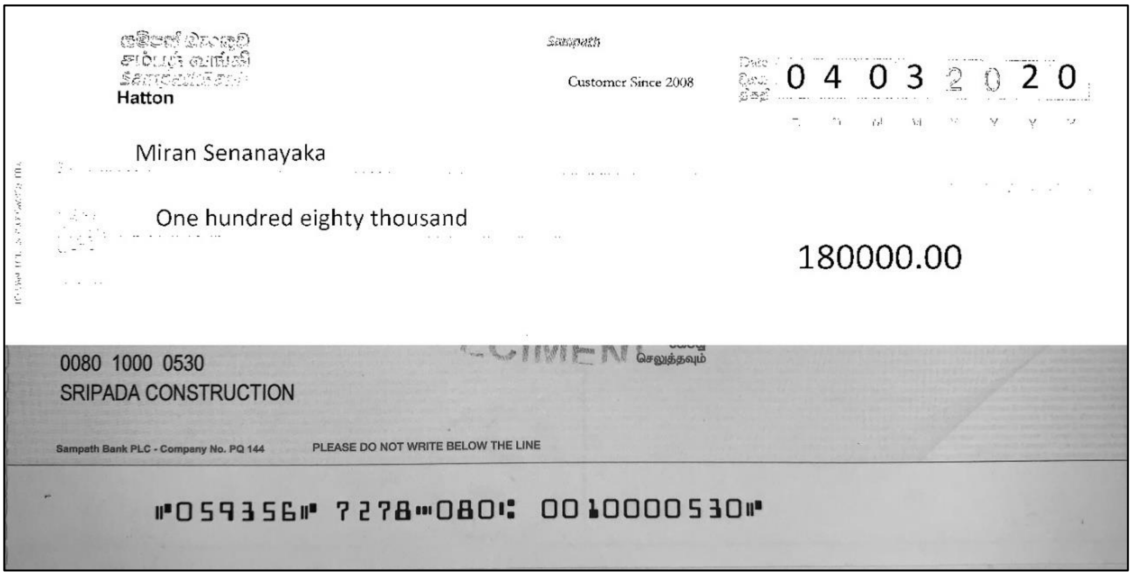


Figure 20: Enhanced Final Image (Captured Input)

If the image is a scan image of the cheque, system converts the image to a gray image, then remove noise by applying a Gaussian Filter with a pre-defined Gaussian kernel size, and then, convert to the binary image. While doing this process, image will sharpen several times with pre-defined binary threshold. After generating the binary image it is ready to process area extraction and for the OCR.

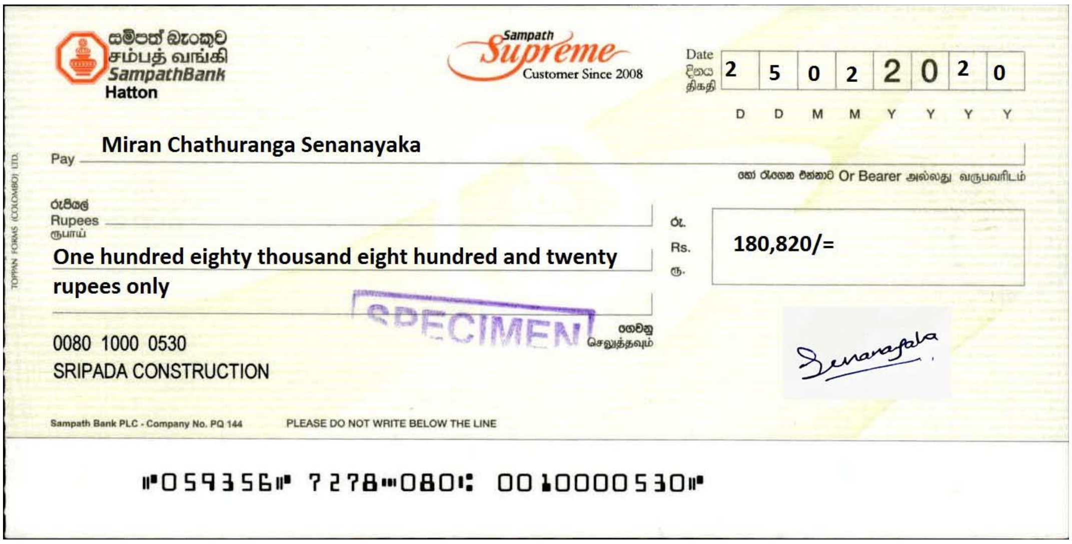


Figure 21: Scanned Input Image

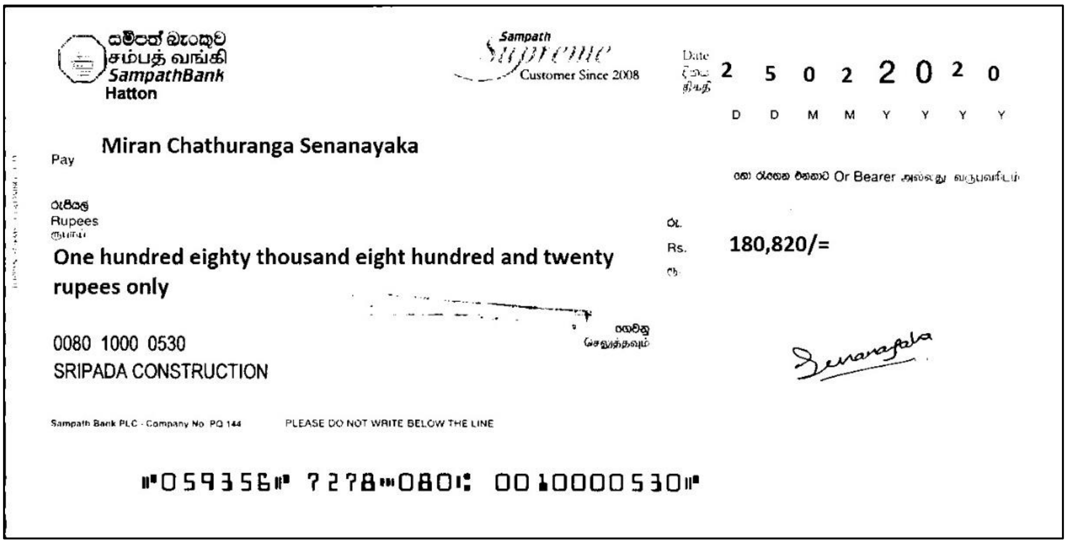


Figure 22: Enhanced Final Image (Scanned Input)

In the next steps, the system will segment separate areas of a cheque for extract details. Those areas are as follows,

I. Date Area
II. Payee Details Area
III. Amounts in Words Area
IV. Amounts in Figures Area
V. Signature Area
VI. Magnetic Ink Area

3.2. IDENTIFY CHEQUE HOLDER DETAILS

After converting the input image to a binary image, the system performs to identify the system reads the magnetic ink details at the bottom of the cheque. System can identify the Current account number related to the cheque, cheque number, branch by using this magnetic ink details.

Figure 23: Identified Magnetic Ink Details Area

The system used a customized Tesseract dataset to OCR these characters in the Magnetic Ink details area. There is a difference in the font of the magnetic ink area and the normal standard numbers. Therefore, there is a separate trained dataset to identify the magnetic ink area numbers. Tesseract MICR OCR is the sample project used to demonstrate the project with the trained dataset [18].

After identifying the magnetic ink area digits, it can generate the current account number. According to this cheque, it is a Sampath Bank cheque. Therefore, the current account number has 12 digits. Current account number generation differ from bank to bank, when it comes to Sampath Bank, they have 12 digits for a Current Account. To generate particular current account number from the details extracted,

- First 4 digits:
  To generate first 4 digits, bank implemented a method where they use 0 as the first digit and branch code as the other 3 digits.

Figure 24: 3 Digit Branch Code in the Image

- Last 8 digits:

Figure 25: Last 8 Digit of the Current Account

3.3. IDENTIFY CHEQUE DETAILS

The system performs to identify the amounts in figures and words by using the image processing techniques. In order to recognize handwritten amounts, MNIST benchmark dataset used as the training dataset [1], [6]. Tesseract Java library used as the natural language processor library [10]. In this process, amounts in words identified and then generated the respective figures. Then the result will be compared with the amounts in figures on the cheque. This increases the accuracy of identifying and validating amounts in figures and amounts in words. Handwritten character and digit recognition methods used to perform this task by using Tesseract library. OpenCV will be used as image comparison tool. Tesseract OCR engine internally use machine learning techniques to recognize characters by using its training dataset.

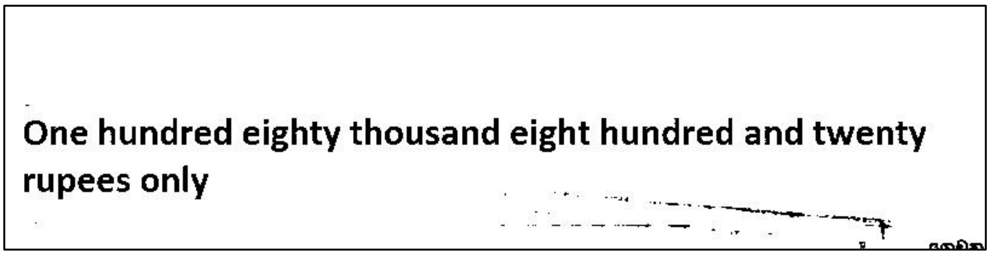


Figure 26: Identified Amounts in Words Area

Figure 27: Identified Amounts in Figure Area

Then, the system detects the date on the cheque while validating the date. Tesseract library is used to extract the date from the cheque. As for the basic validation, in-order to valid a cheque, the cheque date should not be greater than 6 months. If the difference between today's date and cheque date is more than 6 months, the cheque is considered as an expired cheque.

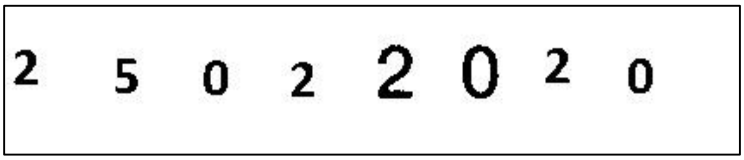


Figure 28: Identified Date Area

After all of these, as for the next step, the system detects the details of the payee. The system is using Tesseract library to extract the payee details.

Figure 29: Identified Payee Details Area

3.4. IDENTIFY SIGNATURE

Finally, the system detects the Signature on the cheque and validate it with the signature which is stored in the database.

Figure 30: Identified Signature Area

To identify signature in this system, it is first used histogram matching algorithm. Later it is used SURF algorithm which is a patented algorithm.

In the histogram matching it was able to identify the exact same signature. But when it comes to a signature, each signature is different with each other. There is at least a single mismatch between two signatures. Therefore, this approach is good only when there is a signature pasted by using a rubber seal. There are some organizations where they use a rubber seal as the signature. This approach is not suitable to identify signatures signed by a person’s hand, so the SURF feature detection algorithm implementation came to address the issue.

In histogram approach, it creates two histograms, one for the horizontal side for the given input image and other one for the vertical side for the input image. Before performing the histogram generation, input image converts to binary image, then there are only black and white pixels where it equals to 1 and 0. To generate binary image from the grayscale image, all pixel values converts to 1 and 0 by the threshold value. In this study, 110 is used as the binary threshold. By doing that all pixel values which are less than 110 converts to 0 and greater than 110 are converted to 1.

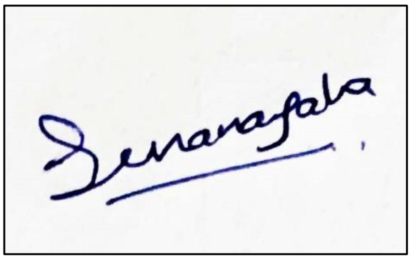


Figure 31: Input Image for Histogram Generation

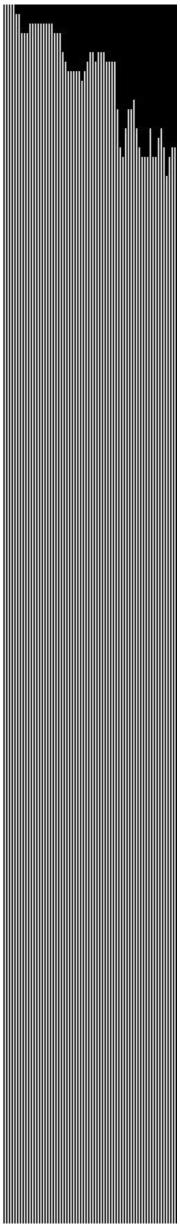


Figure 32: Horizontal histogram

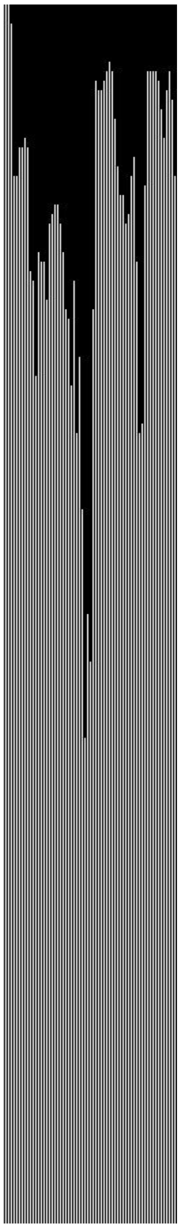


Figure 33: Vertical Histogram

Figure 32 is the horizontal histogram generated for the *Figure 31* signature image where *Figure 33* is the vertical histogram. This same histogram values can only be generated for the given Figure 24 signature image, because humans cannot sign the identical signature every time [26, 27]. After generating histograms for the signature on the cheque, these two histograms compare with the histograms generated for the persons sample signature images provided when the person register for the Bank Current Account. Performing this task, the system can only identifies the 100% equal image which matches the signature which is on the cheque. In other words, by histogram matching it can identify only the same image.

When it comes to the SURF matching, it has good performance with higher processing speed than the SIFT algorithm where it identifies two images by its features [25]. According to the algorithm implementation it is identifying similar features in two images and later match all features with one another. Since this is a patented algorithm, it is not built for default OpenCV jar file which is used in Java programming. Therefore, there was a need to re build the jar file with the non-free part of the original C# source code. After rebuilding with the non-free part, the jar file is ready to consume the SURF and SIFT algorithms. The both algorithms are scale and rotation invariant algorithms which are suitable to match two images like human signatures which has not scaled and which has not defined angle for the signature.

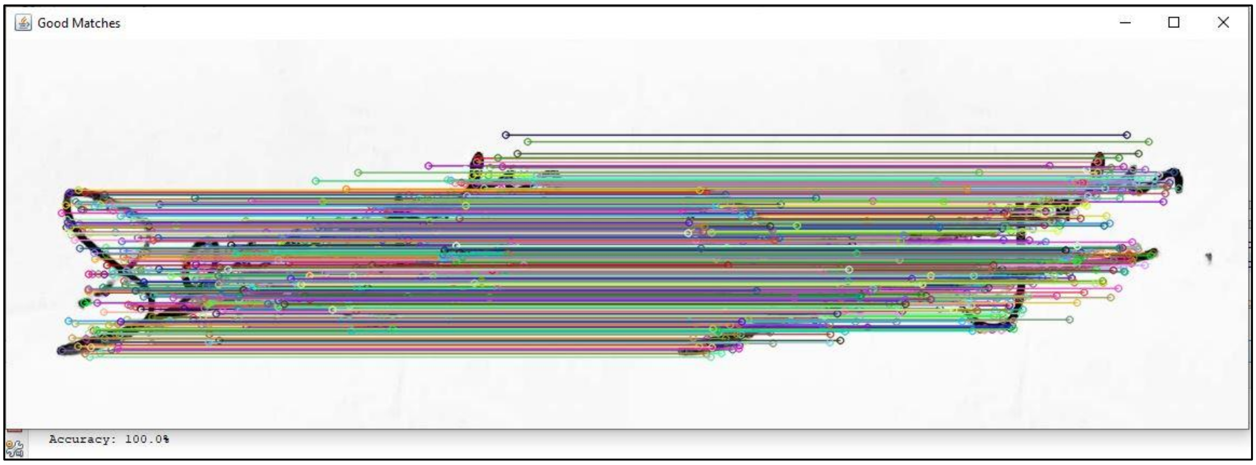


Figure 34: Same Signature Image Identification

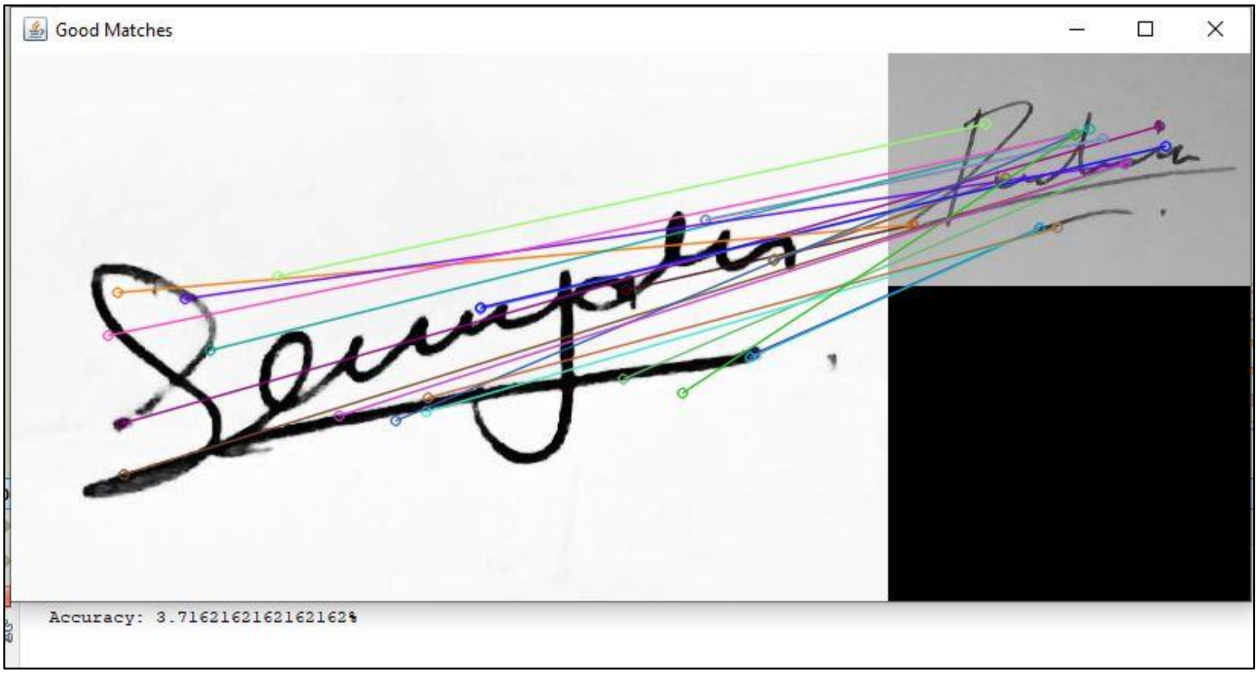


Figure 35: Different Signature Images Identification

Before performing the SURF algorithm, signature images cropped according to its ROI, by doing that unwanted white space is eliminated where it helps to identify only the signature features from the given image.

Figure 36: Cropped Original Image

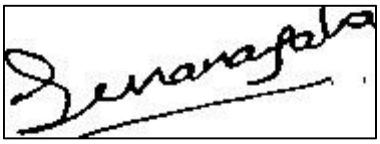


Figure 37: ROI Image

There are three cheque clearance types,

I. Outward Cheque
Outward cheques are cheques receiving from other financial institutions to a financial institution. For example, ABC bank receives a cheque from BCD bank's customer is an outward cheque. In this situation, ABC bank scans the cheque image and send that image to a 3rd party clearing organization (Lanka Clear in Sri Lanka) with a UI. UI contains some information to identify the cheque uniquely such as unique ID, branch, cheque number.

II. Inward Cheque
The 3rd party cheque clearing organization (Lanka Clear in Sri Lanka) send the received cheques from other banks to a respective bank. ABC

III. Transfers
When a customer visits to a branch with the same financial institution cheque, the branch officer will scrutinize the cheque and validate its details and deposit the respective account number instantly if there are not any issues. However, if there are some issues with the account balance, the cheque will transfer to transfer zone where at the end of the day branch manager will take the decision.

According to financial institutions, currently they validate the signatures manually with the help of plethora of employees [23, 24]. It is a time-consuming task where sometimes they have to contact the superior officers as well for further clarifications [24]. The automated cheque authorization system will address these issues and smoothen the process of authorization.

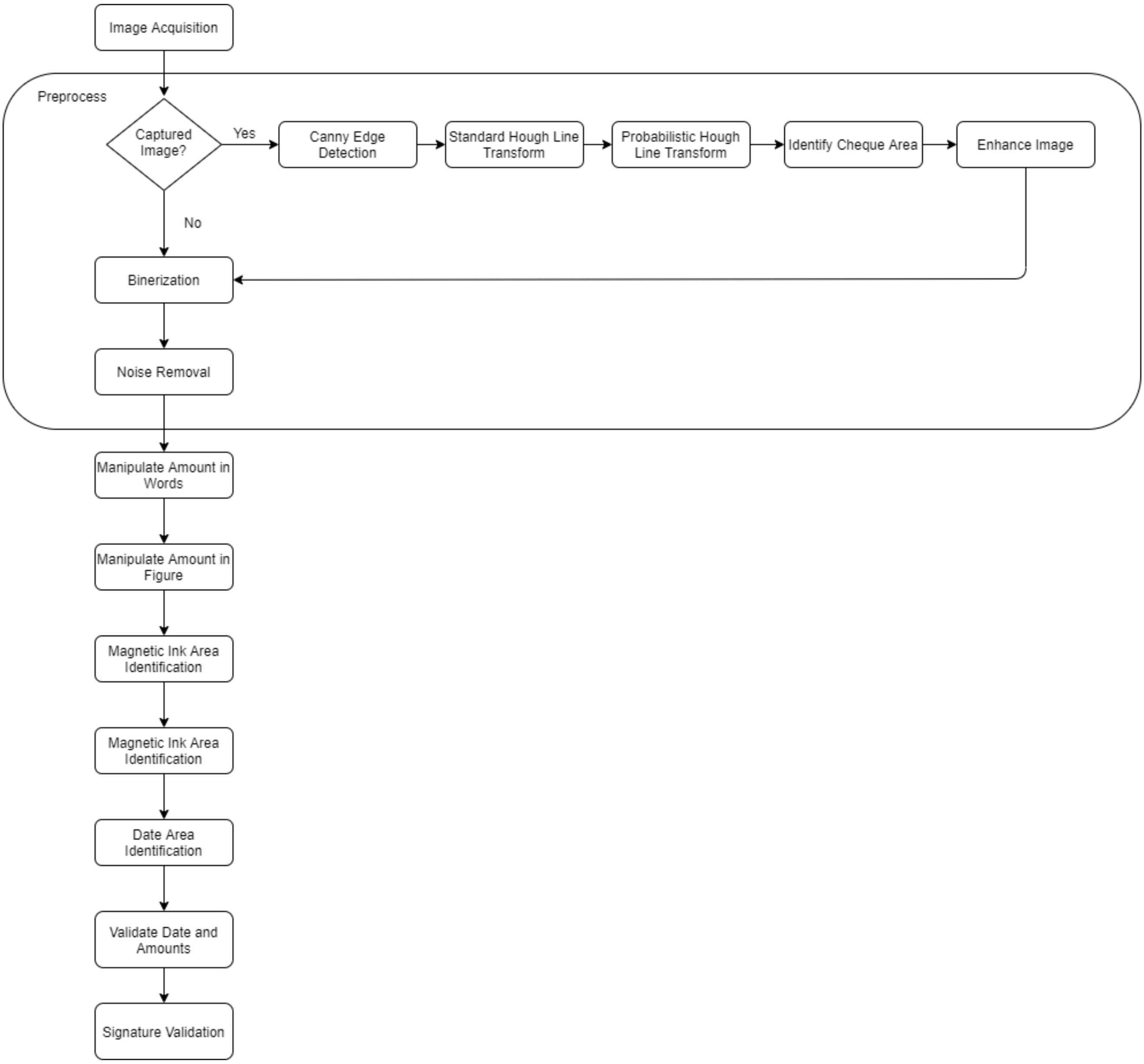


Figure 38: Basic Process

4. EVALUATION

When planning performance evaluation of an image processing application, there could be situation that an algorithm processes an image with rare or unusual features. Such situations, the performance depends on the performance indicators. Typical performance indicators include [1]

1. *Accuracy:* algorithm's performance with respect to some reference
2. *Robustness:* algorithm's ability for enduring numerous conditions
3. *Sensitivity:* algorithm's responsive to small changes in structures
4. *Adaptability:* algorithm's behavior with variation in images
5. *Reliability:* output of the algorithm when repeatedly using the same stable data
6. *Efficiency:* the practicability of an algorithm (space and time)

In this study, it is considering image processing domain area with image comparison, optical character recognition of each printed characters and handwritten characters. It is using experiment and mathematical proof based approach to evaluate these modules separately.

Generally minimum image resolution is decided to 1500*500px by the *Table 1* experimental results.

Resolution	Problem
800*600	Printed fonts, date and signature is blurred
1000*800	Handwritten fonts, signature and date is blurred
1200*1000	Handwritten fonts are blurred

Table 1: Experimented Results for Resolution

There are 4 main modules in this study,

4.1.PREPROCESSING

Preprocessing consists of the dynamic identification of the border of bank cheque from the given image and straight the image if it is skewed. Then later in this phase, it has to identify the respective areas such as date area, payee details area, bearer details area, amounts in words area, amounts in figure area, signature area and cheque details area as well.

In-order to manually evaluate these steps, there are separate tabs in the GUI.

4.2.IDENTIFY CHARACTERS USING THE DEFAULT TESSERACT DATASET

This study is using famous open source optical character recognition (OCR) tool which is the Tesseract. According to a survey it is identified that Tesseract tool is much better in accuracy than the Transym which is not an open source OCR tool, further it mentioned that the performance of the Tesseract is comparatively higher than the Transym tool [22].

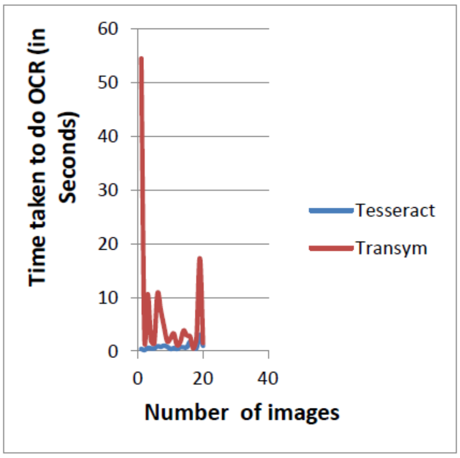


Figure 39: *Tesseract and Transym Accuracy [22]*

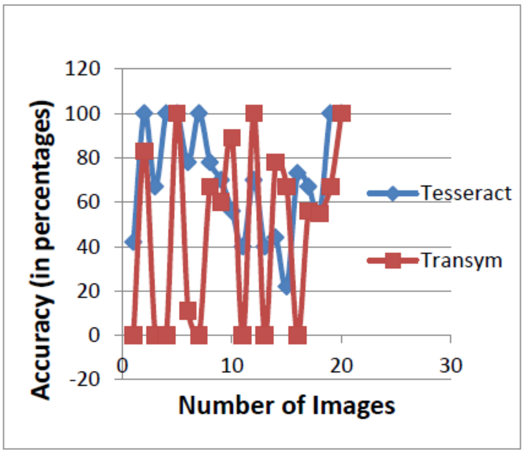


Figure 40: Tesseract and Transym Performance [22]

Therefore to identify characters in date, payee details, bearer details, amount in words, amount in figures, cheque details areas it is used the Tesseract tool and the dataset [10].

4.3.IDENTIFY CHEQUE DETAILS BY USING THE TRAINED TESSERACT DATASET

Font of the cheque details differ from the default Tesseract dataset. Therefore when the default Tesseract dataset is used to identify characters of the cheque details area, it is giving less accuracy with compared to others. Therefore there is a need to train a dataset in-order to get a higher accuracy for cheque area character identification.

To train this dataset it is planning to use supervised learning algorithm by comparing the accuracy and the efficiency of selected several algorithms. Among those algorithms, best algorithm will be selected and use to train the dataset.

After using this trained dataset, it can retrieve the output from the application to the given cheque details. By comparing it manually with the physical cheque it can be determined the accuracy level. Extracted sample cheque details area is as follows.

Figure 41: Extracted cheque details area

4.4.SIGNATURE COMPARISON

Signatures should also be evaluated by manually by checking the cheque's signature and stored signature in the database. By doing it for a number of cheques with different signatures it can come to an accuracy level.

Previous literatures show that there are several algorithms to compare signatures by using its features and pixels. There are 4 given algorithms which are, Binary Robust Independent Elementary Feature Algorithm (BRIEF), Oriented Fast and Rotated BRIEF Algorithm (ORB), Scale Invariant Feature Transform Algorithm (SIFT), Speeded Up Robust Feature Algorithm (SURF). Among these algorithms, best 2 algorithms to compare signatures are SIFT and SURF algorithms [28]. Therefore those two algorithms will be used to compare signatures.

Feature detection method	Feature descriptor	Accuracy	AUC	F1-measure
SIFT	SIFT	0.8333	0.924	0.8462
SURF	SURF	0.875	0.935	0.8800
FAST	BRIEF	0.75	0.854	0.7000
ORB	ORB	0.8333	0.915	0.8333
BRISK	BRISK	0.7916	0.894	0.7826
FAST	FREAK	0.7916	0.889	0.8000

Figure 42: Performance Evaluation with Chinese Dataset [28]

Feature detection method	Feature descriptor	Accuracy	AUC	F1-measure
SIFT	SIFT	0.7917	0.911	0.8000
SURF	SURF	0.8333	0.928	0.8462
FAST	BRIEF	0.6250	0.823	0.5714
ORB	ORB	0.8263	0.882	0.8182
BRISK	BRISK	0.6667	0.86	0.6923
FAST	FREAK	0.7083	0.876	0.6957

Figure 43: Performance Evaluation with Dutch Dataset [28]

Other than above mentioned evaluation methodologies, it is planning to input the same image in different resolutions, different sizes, and different brightness levels and check whether the output is the same. It can guarantee the consistency and robustness of the application.

To check the efficiency of the application and algorithms, it can calculate the time taken for the whole process and particular modules and compare it with the manual process. If the manual process takes less time than the application it is not a viable to use the application. Not only the time but it should consider resource utilization of the computer from the application. If the application use more resources than the computer has, it is not an effective application.

5. RESULTS

5.1.PREPROCESSING

In preprocessing using above mentioned process, accuracy of the correctly extracted cheque from the camera captured image is more than 85%. In-order to calculate the percentage, it is used 200 images.

If the image is skewed, then the system tend to identify the incorrect area from the captured image.

Figure 44: Sample Skewed Input Image

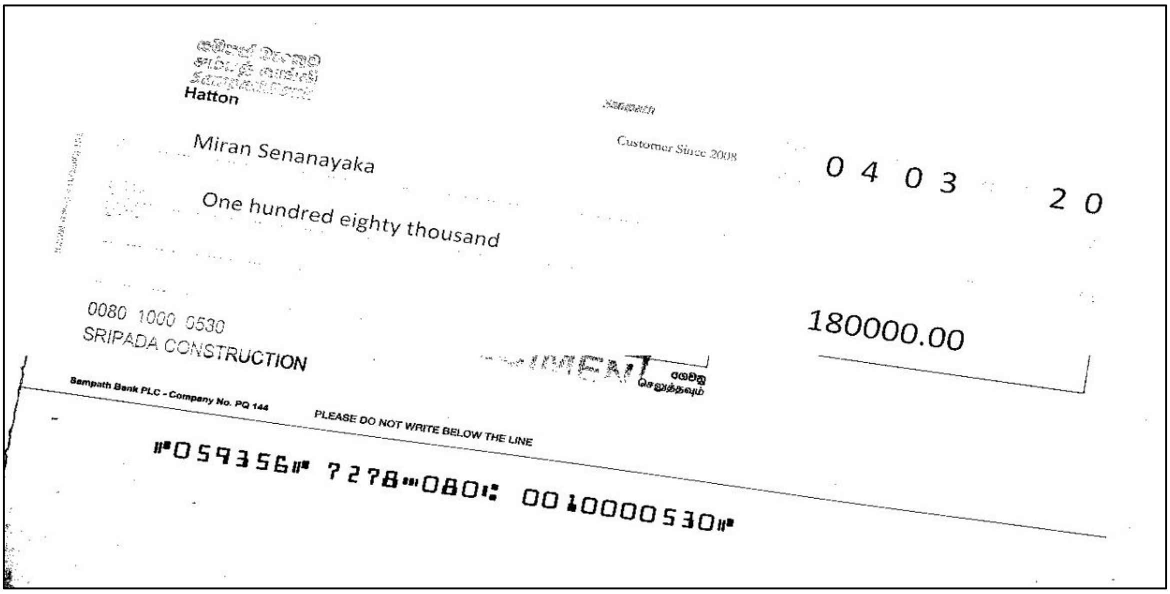


Figure 45: Incorrectly Identified Image (Skewed)

Preprocessing is failing because of the brightness issues of the input image.

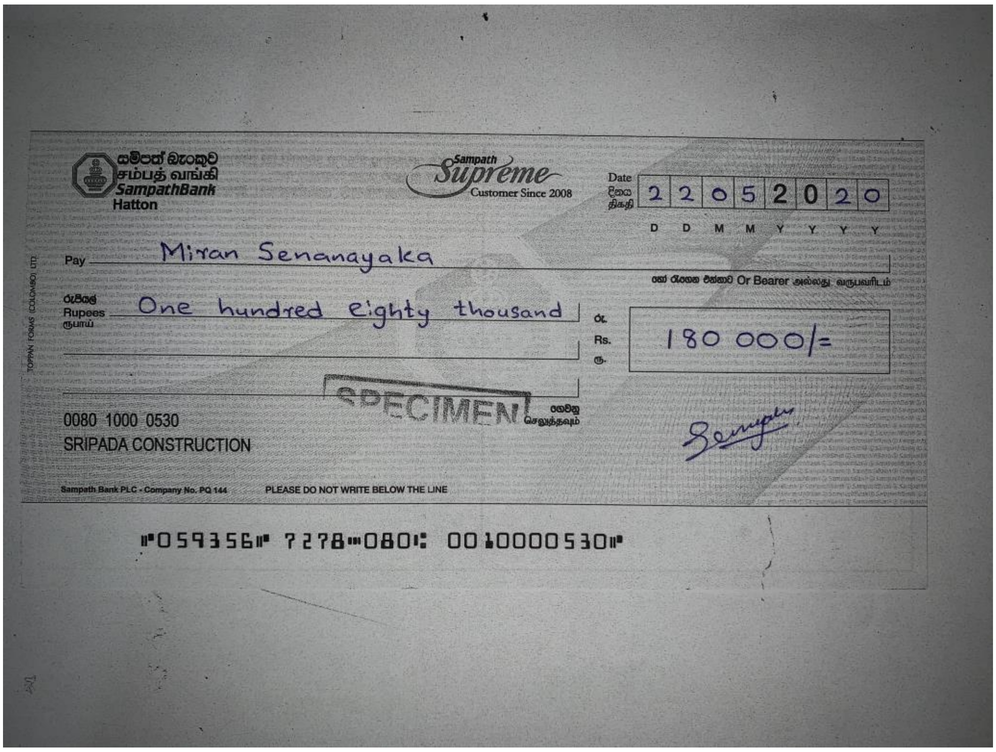


Figure 46: Less Brightened Image

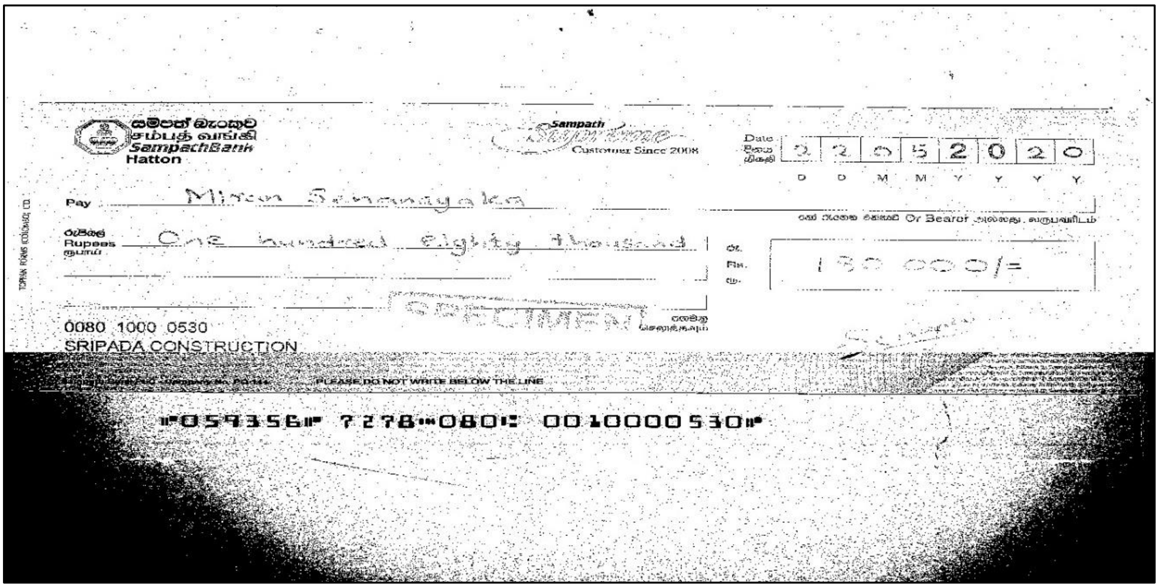


Figure 47: Failed Preprocessed Image

5.2. IDENTIFY CHEQUE HOLDER DETAILS AND IDENTIFY CHEQUE DETAILS

There are two types of identification in this step,

- Printed Characters
  On one hand, printed characters are identified with the accuracy rate of more than 90% of a preprocessed image. When it comes to a scan image of a cheque it is much higher than the captured image of a cheque. Scan image's character identification is more than 95% while captured image character recognition is around 90%. More than 100 images has being used to evaluate these results.
  When identifying Magnetic Ink Characters (MICR) it has an accuracy of more than 90% for both scanned and captured images.

No. of Images Used	**No. of Images Passed**	**No. of Images Failed**	**Accuracy**
100 (Scanned Images)	95	5	95%
100 (Camera Captured)	90	10	90%

Table 2: Printed Character Identification Accuracy

- Handwritten Characters
  On the other hand, hand written characters have huge gap in accuracy percentage which is around 60%. When it is a scanned image, the accuracy is around 65% while, captured image character recognizing accuracy is around 60%. 100 images have being used to evaluate these results.

No. of Images Used	**No. of Images Passed**	**No. of Images Failed**	**Accuracy**
100 (Scanned Images)	65	35	65%
100 (Camera Captured)	60	40	60%

Table 3: Handwritten Character Identification Accuracy

5.3. IDENTIFY SIGNATURE

When identifying signatures, it only consider to identify the particular person signature without identifying fraud signatures of the same signature. Therefore, identification of the signature on the cheque with the given samples has an accuracy of around 80%. More than 200 images are used to evaluate these results.

If two images are look-alike or image features are similar to each other, those kind of images are failed to recognize with each other.

No. of Images Used	**No. of Images Passed**	**No. of Images Failed**	**Accuracy**
200 (Scanned Images)	160	40	80%
200 (Camera Captured)	140	60	70%

Table 4: Signature Identification Accuracy

6. DISCUSSION / CONCLUSION

In this study, it is implemented a system to automate check realization process by using Java programming language and image processing. The different identified areas such as date, pay, amounts in words, and amounts in figure, signature, and magnetic ink are captured by using Tesseract OCR and feature extraction algorithms such as SURF. Obviously this will help to reduce human work involvement for realization process in a financial institution.

Due to the COVID-19 pandemic situation all together around 40 images are used to evaluate the study and the system. According to the obtained results of the study, typed cheque realization automation process has a higher success rate while handwritten cheque realization automation has a less or equal percentage if the handwritten characters are simple and in a good readable format.

When it comes to signature identification, it is used better feature extraction approach rather applying OCR technique to identify signatures. That increases the accuracy of signature identification. By using Tesseract OCR, study's accuracy and the performance is better where Tesseract internally use machine learning approaches to recognize characters with the trained dataset.

7. FUTURE WORK

As for the future work, this system can enhance to identify more than one signature on the cheque. Currently it is only identifying on signature on the image. Other than that, this can be extended to identify skewed images, currently the system not support for the skewed images where it fails the preprocessing phase. Further, the system's handwritten character recognition could be enhanced using other algorithms which are using to identify handwritten characters.

8. REFERENCES

[1] “MNIST Hand-written Digit Database, Yann LeCun, Corinna Cortes and Chris Burges.” Accessed September 18, 2019. http://yann.lecun.com/exdb/mnist/.

[2] Chen, Shengfeng, Rabia Almamlook, Yuwen Gu, and Lee Wells. “Offline Handwritten Digits Recognition Using Machine Learning,” 2018.

[3] Zhao, K. “Handwritten Digit Recognition and Classification Using Maching Learning”. M.Sc. in Computing (Data Analytics), Technological University Dublin, 2018.

[4] S M Shamim, Mohammad Badrul Alam Miah, Angona Sarker, Masud Rana and Abdullah Al Jobair. “3-Handwritten-Digit-Recognition.Pdf.” *Global Journal of Computer Science and Technology: D Neural & Artificial Intelligence Volume 18 Issue 1 Version 1.0 Year 2018.* [Online]. Available: https://globaljournals.org/GJCST_Volume18/3-Handwritten-Digit-Recognition.pdf. [Accessed September 23, 2019].

[5] Ravi Katukam. “(PDF) Image Comparison Methods & Tools” *1st National Conference on, EMERGING TRENDS IN INFORMATION TECHNOLOGY[ETIT], 28th -29th December 2015.* [Online]. Available:

https://www.academia.edu/19962797/Image_comparison_Methods_and_Tools_A_Review. [Accessed September 23, 2019].

[6] Zheheng Rao, Chunyan Zeng, Minghu Wu, Zhifeng Wang, Nan Zhao, Min Liu, Xiangkui Wan. “Research on a handwritten character recognition algorithm based on an extended nonlinear kernel residual network” KSII Transactions on Internet and Information Systems. 12. 413-435. 10.3837/tiis.2018.01.020.

[7] R.Yogamangalam, B.Karthikeyan “Segmentation Techniques Comparison in Image Processing”. School of Computing, SASTRA University, Thanjavur, TamilNadu, India, 2013.

[8] G.T. Shrivakshan and Dr.C. Chandrasekar, "A Comparison of various Edge Detection Techniques used in Image Processing" in IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 5 [Online].Available: ISSN, https://www.ijcsi.org/papers/IJCSI-9-5-1-269-276.pdf. . [Accessed Dec. 03, 2019].

[9] Monica Patel and Shital P. Thakkar, "Handwritten Character Recognition in English: A Survey" International Journal of Advanced Research in Computer and Communication Engineering Vol. 4, Issue 2, Feb. 2015.

[10] tesseract-ocr: 2019 [Online]. Available: https://tesseract-ocr.github.io/. [Accessed: May 16, 2020].

[11] E. Alpaydm, Introduction to Machine Learning, Cambridge: The MIT Press, 2010

[12] J. Brownlee, "Supervised and Unsupervised Machine Learning Algorithms," 16th March 2016. [Online]. Available: https://machinelearningmastery.com/supervised-and-unsupervised-machine-learning-algorithms/. [Accessed 16th May 2020].

[13] Andrey Golubev, Dmitry Kurtaev. “OpenCV”. 2019. [Online]. Available: https://opencv.org/about/

[14] tesseract-ocr: 2019 [Online]. Available: https://github.com/tesseract-ocr. [Accessed: May 16, 2020].

[15] Estevão S. Gedraite, Murielle Hadad. "Investigation on the Effect of a Gaussian Blur in Image Filtering and Segmentation". University of São Paulo, Av.Luciano Gualberto, travessa 3, nº 380, Brazil, 2011.

[16] Radhika Chandwadkar, Saurabh Dhole, Vaibhav Gadewar,Deepika Raut, Prof. S. A. Tiwaskar. "Comparison of Edge Detection Techniques". Computer Engineering Department, VishwakarmaInstitute of Information Technology, Pune. 2013.

[17] MichaelWirth, Matteo Fraschini, Martin Masek and Michel Bruynooghe, "Performance Evaluation in Image Processing". Department of Computing and Information Science, University of Guelph, Guelph, ON, Canada, 2006.

[18] Tesseract-MICR-OCR: 2019 [Online]. Available: https://github.com/BigPino67/Tesseract-MICR-OCR. [Accessed: June 21, 2020].

[19] Walid Hussein, Mostafa A. Salama and Osman Ibrahim, "Image Processing Based Signature Verification Technique to Reduce Fraud in Financial Institutions". Faculty of Informatics and Computer Science, the British University in Egypt, Cairo, Egypt, 2016.

[20] Darshana Mistry, Asim Banerjee, "Comparison of Feature Detection and Matching Approaches: SIFT and SURF", Department of Computer Engineering, Technical Associate, eInfochips Training and Research Academy, Ahmedabad, 2017.

[21] Shivani Dhiman, A.J Singh, "Tesseract Vs Gocr A Comparative Study", International Journal of Recent Technology and Engineering, Volume 2, Issue 4, September 2013. [Online]. Available: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.676.9685&rep=rep1&type=pdf

[22] Chirag Patel, Atul Patel, Dharmendra Patel, "Optical Character Recognition by Open Source OCR Tool Tesseract: A Case Study", Smt. Chandaben Mohanbhai, Patel Institute of Computer, Applications, Charotar University of Science and Technology, 2012

[23] Ashok Kumar. D, Dhandapani. S, "A Bank Cheque Signature Verification System using FFBP Neural Network Architecture and Feature Extraction based on GLCM", Department of Computer Science and Applications, Government Arts College, Tiruchirapalli – 620 022, Tamil Nadu, INDIA, 2014

[24] Chetana Vishwanath Khatode, Prof. Sunil S. Morade, "Verification of Signatures on Bank Cheques", International Journal of Science, Engineering and Technology Research, Volume 4, 2015

[25] P M Panchal1, S R Panchal2, S K Shah, "A Comparison of SIFT and SURF", International Journal of Innovative Research in Computer and Communication Engineering, Vol. 1, 2013

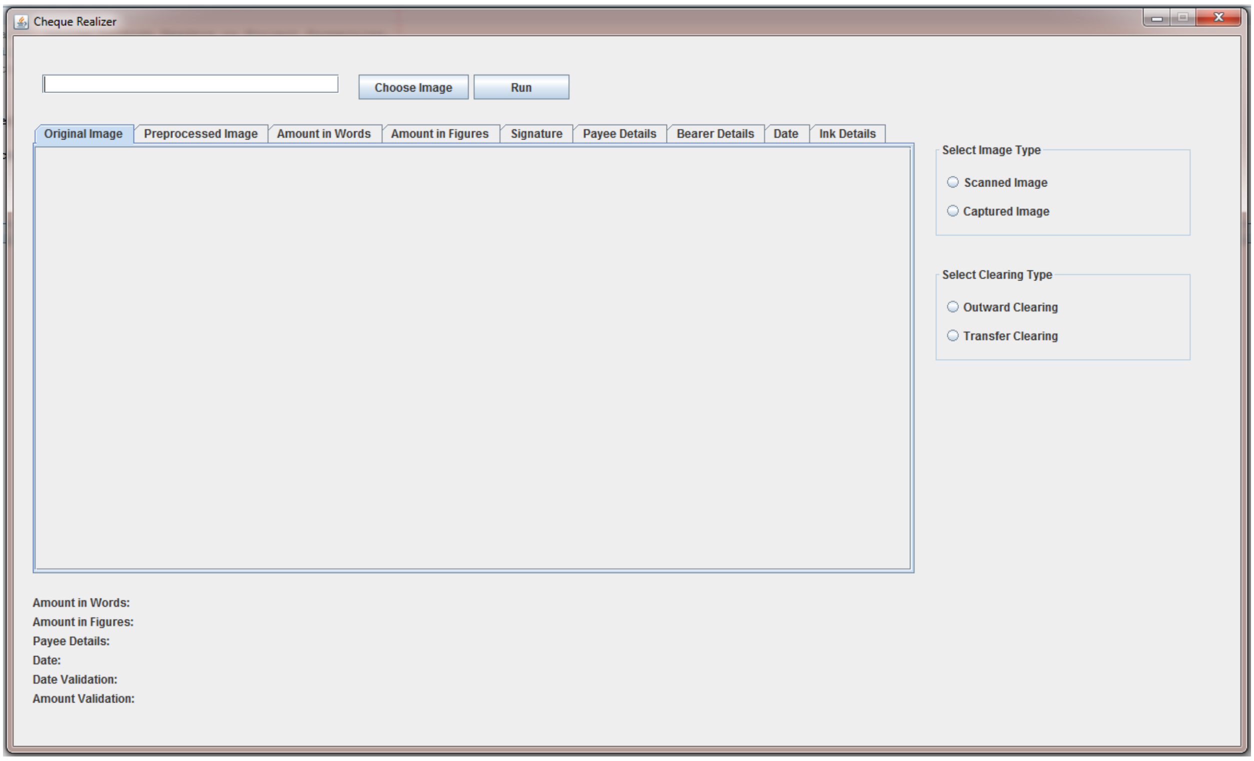
[26] David Fillingham, "A Comparison of Digital and Handwritten Signatures", Ethics and Law on the Electronic Frontier, Fall, 1997

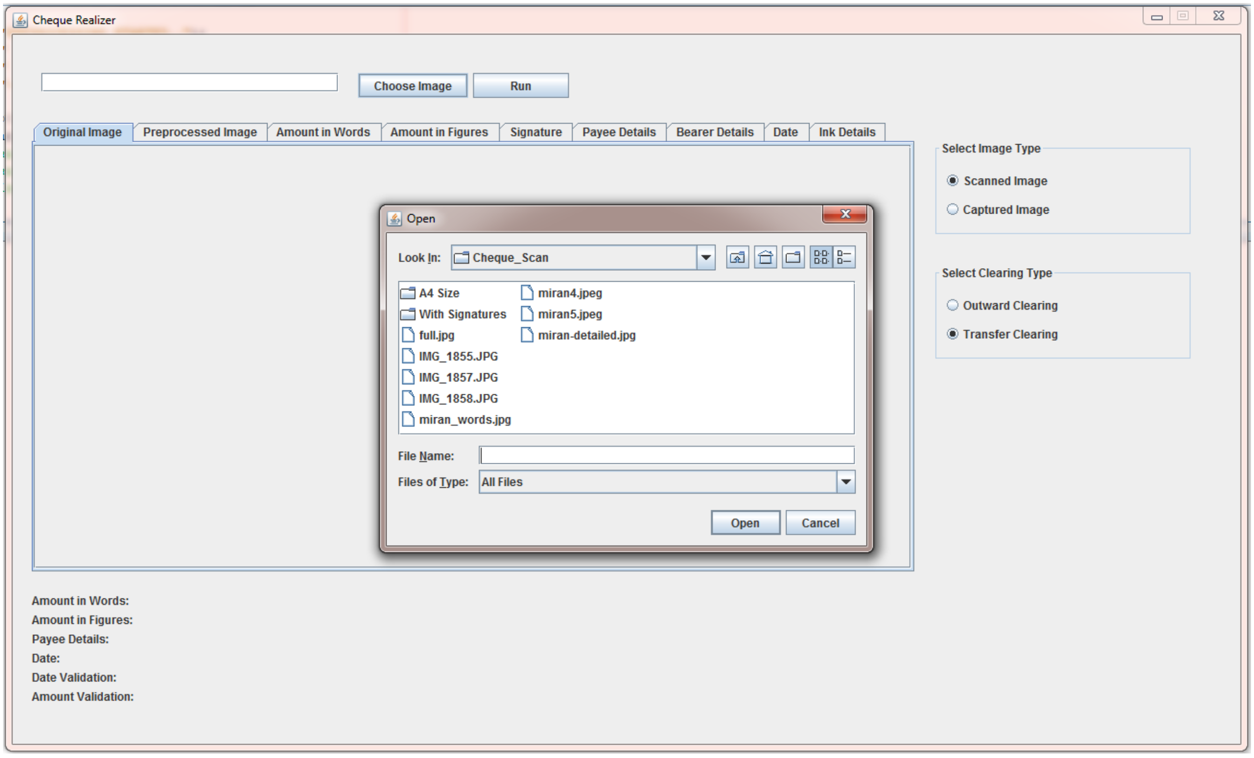
[27] Bača Miroslav, Koruga Petra, Fotak Tomislav, "Basic on-line handwritten signature features for personal biometric authentication", Faculty for Organization and Informatics, Centre for Biometrics Pavlinska 2, 42000 Varaždin, Croatia, 2011

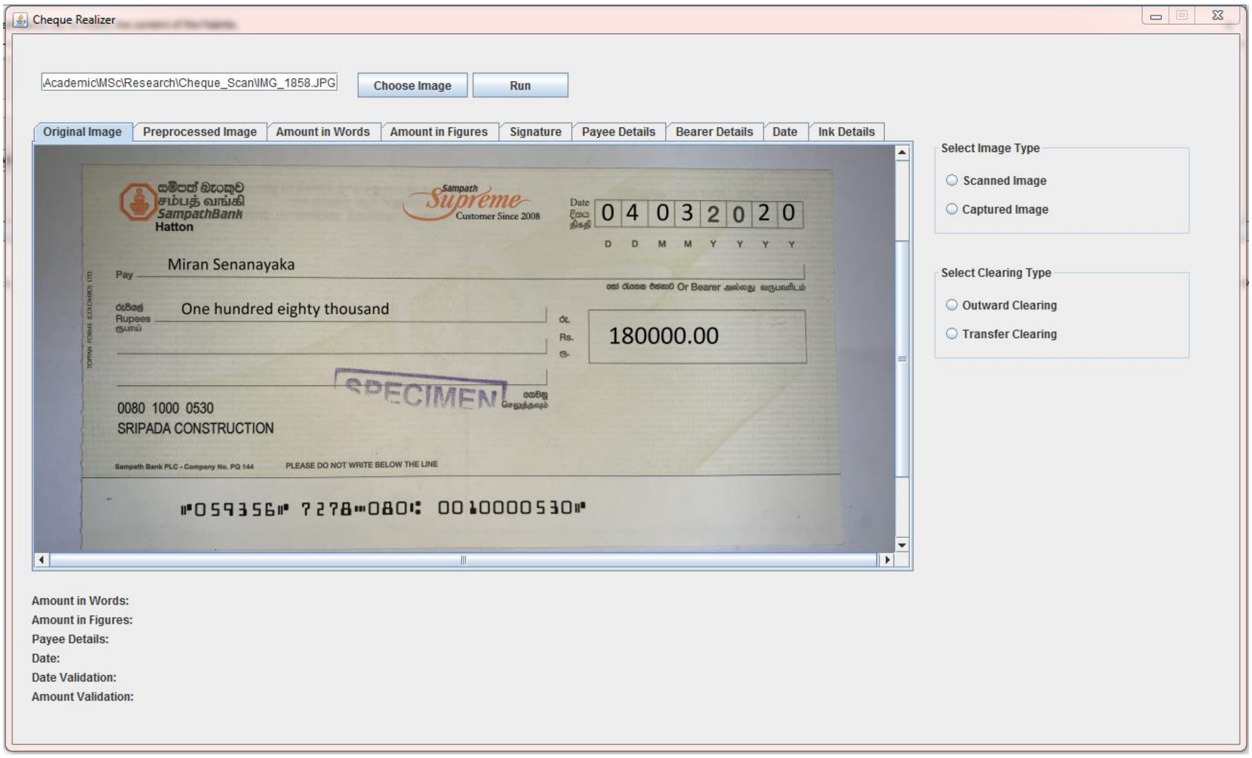
[28] ABM Ashikur Rahman, Md Redwan Karim, Rafsanjany Kushol and Md. Hasanul Kabir, "Performance Comparison of Feature Descriptors in Offline Signature Verification," *IUT Journal of Engineering and Technology (JET)*, vol 14, No. 1, December 2018

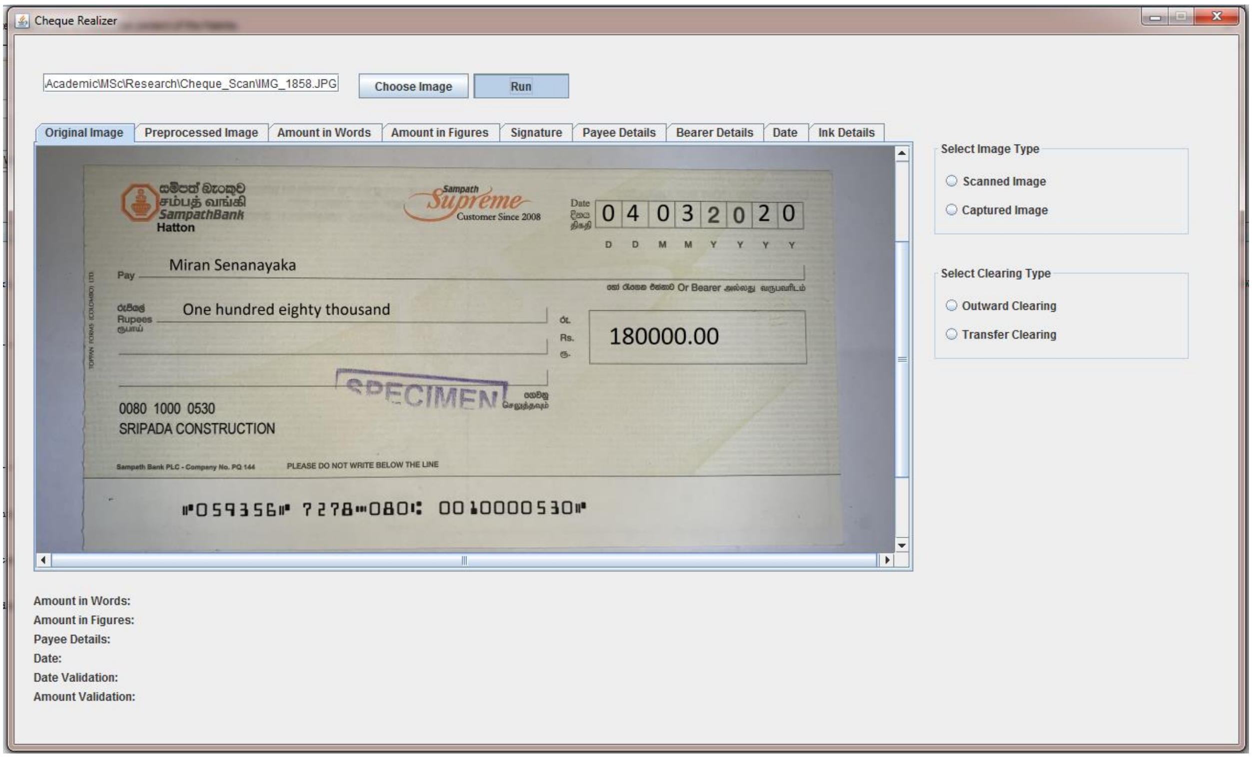
Appendix A

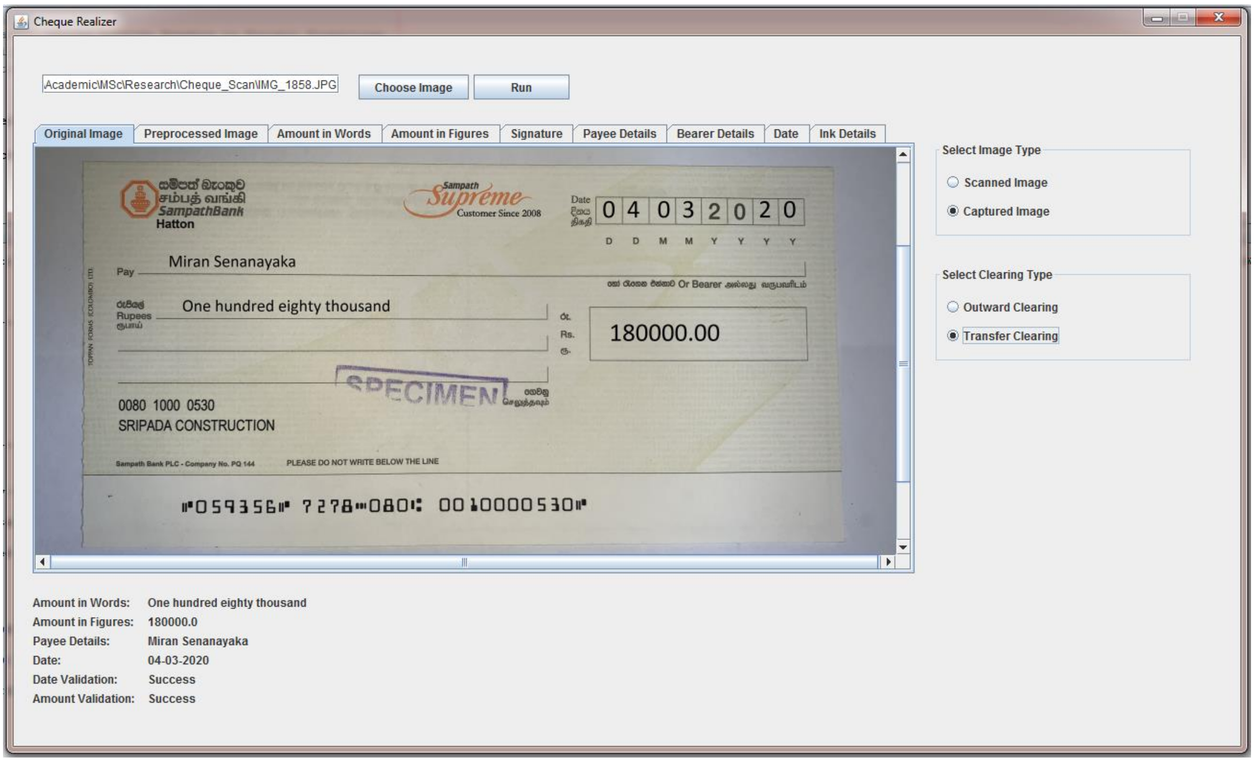
USER INTERFACES OF CAPTURED IMAGE

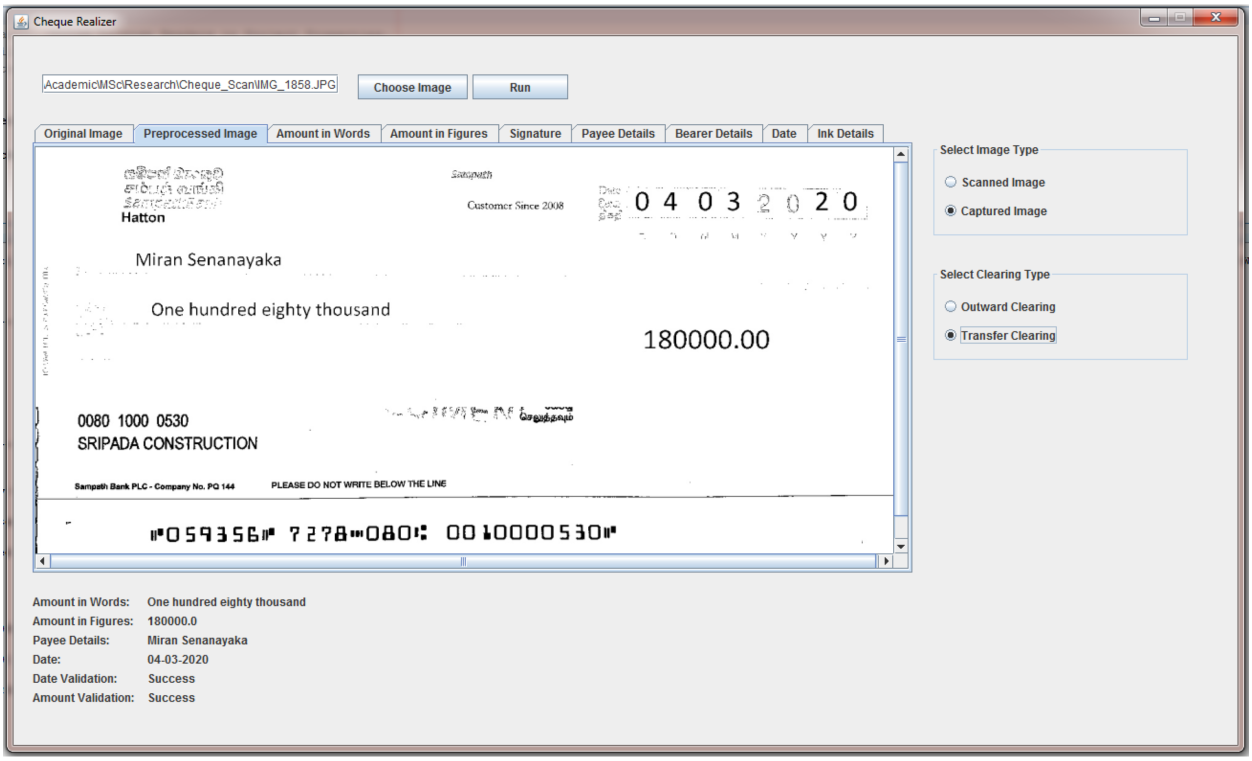


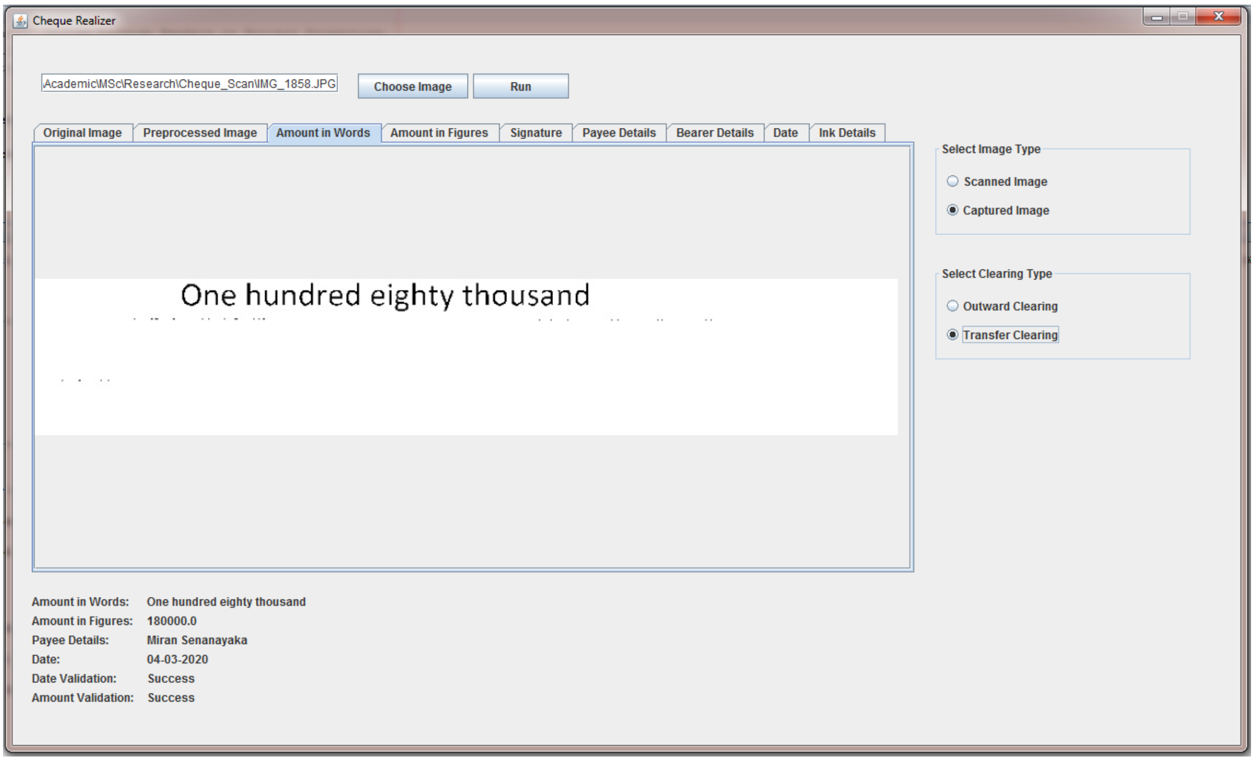


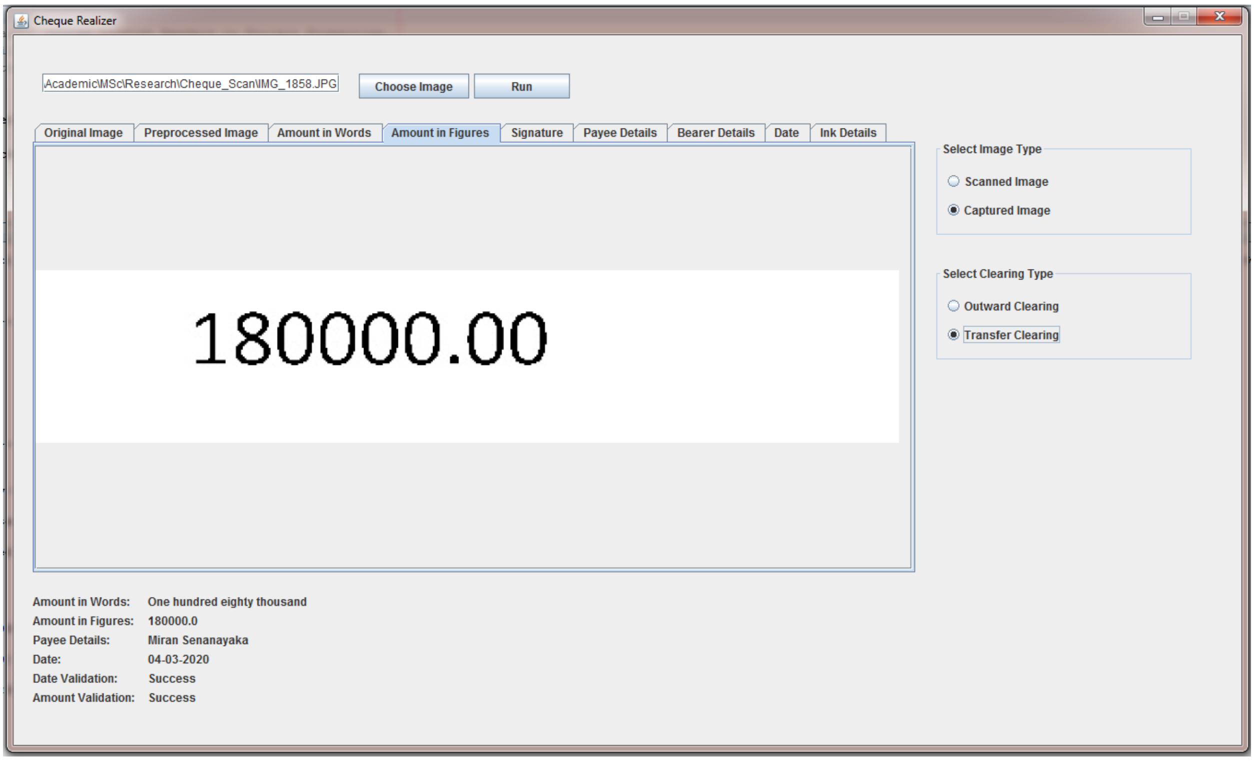

Cheque Realizer
Academic\MSc\Research\Cheque_Scan\IMG_1858.JPG
Choose Image
Run
Original Image
Preprocessed Image
Amount in Words
Amount in Figures
Signature
Payee Details
Bearer Details
Date
Ink Details
Select Image Type
Scanned Image
Captured Image
Select Clearing Type
Outward Clearing
Transfer Clearing
Amount in Words:
Amount in Figures:
Payee Details:
Date:
Date Validation:
Amount Validation:

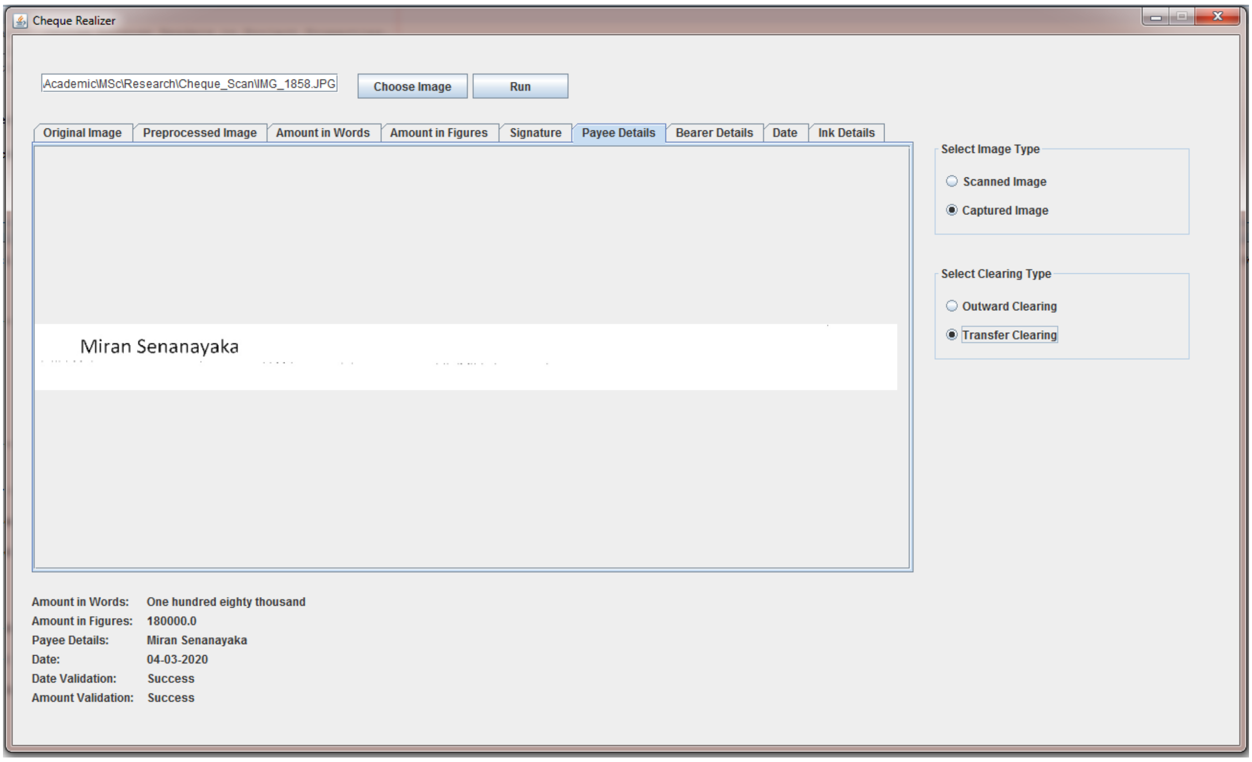

Cheque Realizer
Academic\MSc\Research\Cheque_Scan\IMG_1858.JPG
Choose Image
Run
Original Image
Preprocessed Image
Amount in Words
Amount in Figures
Signature
Payee Details
Bearer Details
Date
Ink Details
Select Image Type
Scanned Image
Captured Image
Select Clearing Type
Outward Clearing
Transfer Clearing
Amount in Words:
Amount in Figures:
Payee Details:
Date:
Date Validation:
Amount Validation:

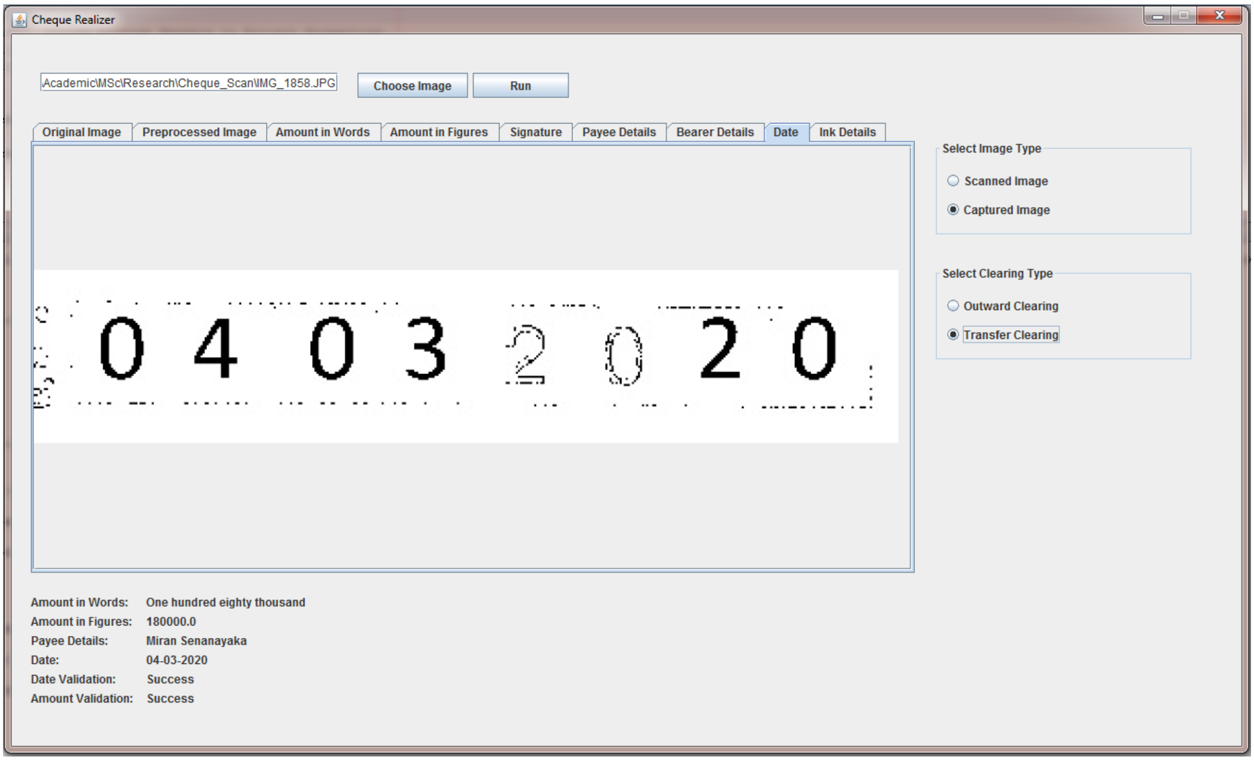

Cheque Realizer
Academic\MSc\Research\Cheque_Scan\IMG_1858.JPG
Choose Image
Run
Original Image
Preprocessed Image
Amount in Words
Amount in Figures
Signature
Payee Details
Bearer Details
Date
Ink Details
Select Image Type
Scanned Image
Captured Image
Select Clearing Type
Outward Clearing
Transfer Clearing
Hatton
Customer Since 2008
Date
04032020
D D M M Y Y Y Y
Pay
Miran Senanayaka
Or Bearer
Rupees
One hundred eighty thousand
Rs.
180000.00
SPECIMEN
0080 1000 0530
SRIPADA CONSTRUCTION
PLEASE DO NOT WRITE BELOW THE LINE
Amount in Words: One hundred eighty thousand
Amount in Figures: 180000.0
Payee Details: Miran Senanayaka
Date: 04-03-2020
Date Validation: Success
Amount Validation: Success

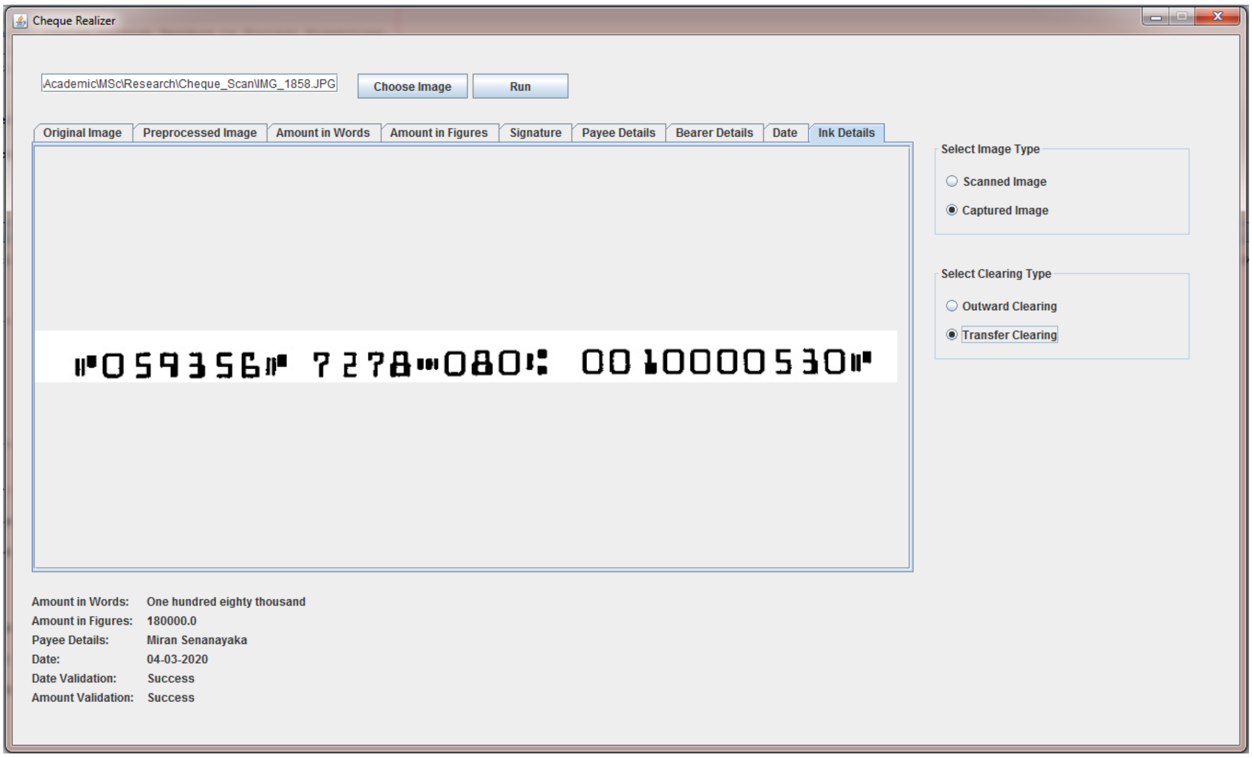

Cheque Realizer
Academic\MSc\Research\Cheque_Scan\IMG_1858.JPG
Choose Image
Run
Original Image
Preprocessed Image
Amount in Words
Amount in Figures
Signature
Payee Details
Bearer Details
Date
Ink Details
Select Image Type
Scanned Image
Captured Image
Select Clearing Type
Outward Clearing
Transfer Clearing
Hatton
Customer Since 2008
0 4 0 3 2 0 2 0
Miran Senanayaka
One hundred eighty thousand
180000.00
0080 1000 0530
SRIPADA CONSTRUCTION
PLEASE DO NOT WRITE BELOW THE LINE
Amount in Words: One hundred eighty thousand
Amount in Figures: 180000.0
Payee Details: Miran Senanayaka
Date: 04-03-2020
Date Validation: Success
Amount Validation: Success


Cheque Realizer
Academic\MSc\Research\Cheque_Scan\IMG_1858.JPG
Choose Image
Run
Original Image
Preprocessed Image
Amount in Words
Amount in Figures
Signature
Payee Details
Bearer Details
Date
Ink Details
One hundred eighty thousand
Select Image Type
Scanned Image
Captured Image
Select Clearing Type
Outward Clearing
Transfer Clearing
Amount in Words: One hundred eighty thousand
Amount in Figures: 180000.0
Payee Details: Miran Senanayaka
Date: 04-03-2020
Date Validation: Success
Amount Validation: Success

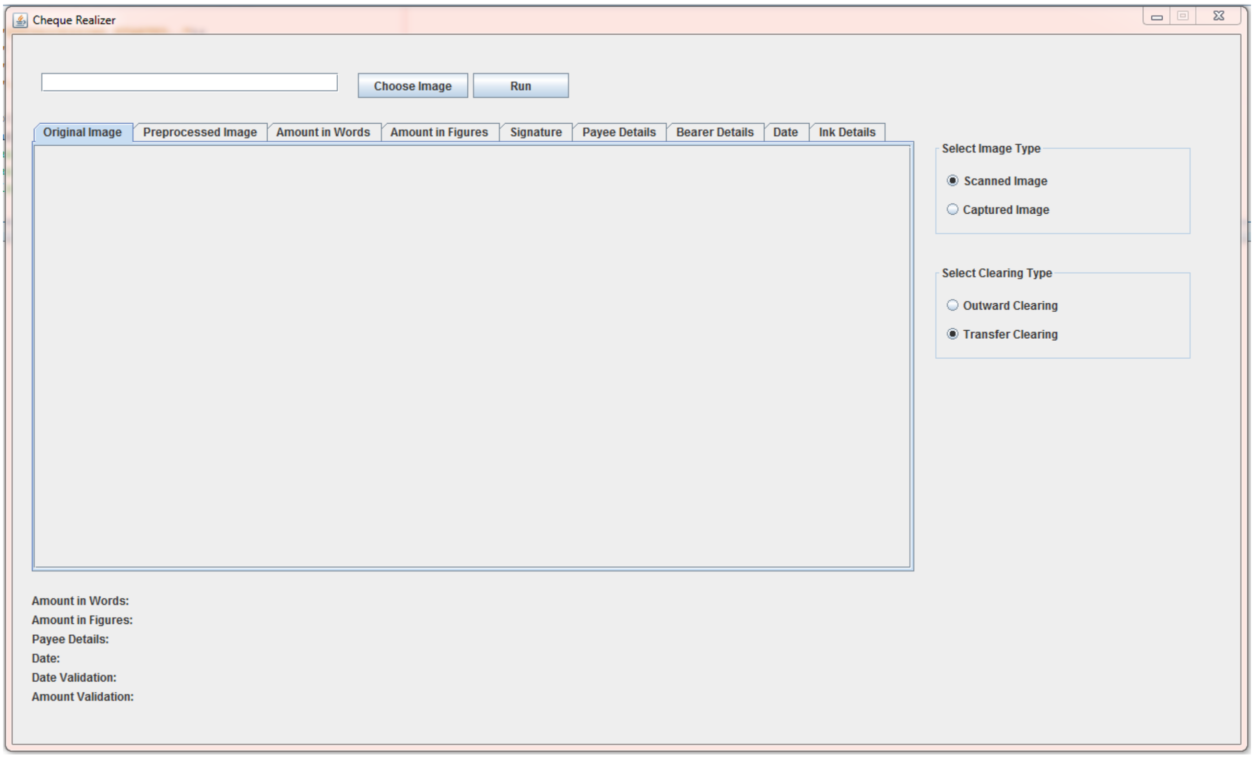

Cheque Realizer
Academic\MSc\Research\Cheque_Scan\IMG_1858.JPG
Choose Image
Run
Original Image
Preprocessed Image
Amount in Words
Amount in Figures
Signature
Payee Details
Bearer Details
Date
Ink Details
180000.00
Select Image Type
Scanned Image
Captured Image
Select Clearing Type
Outward Clearing
Transfer Clearing
Amount in Words: One hundred eighty thousand
Amount in Figures: 180000.0
Payee Details: Miran Senanayaka
Date: 04-03-2020
Date Validation: Success
Amount Validation: Success

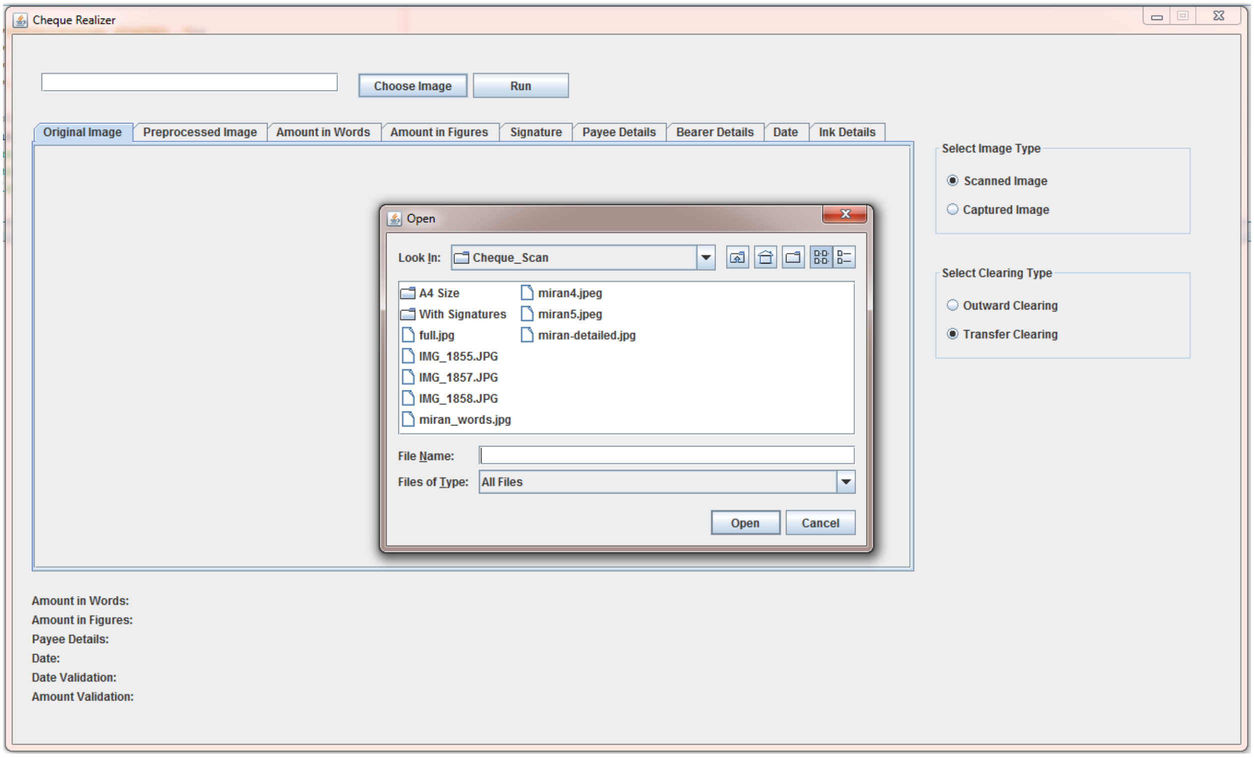

Cheque Realizer
Academic\MSc\Research\Cheque_Scan\IMG_1858.JPG
Choose Image
Run
Original Image
Preprocessed Image
Amount in Words
Amount in Figures
Signature
Payee Details
Bearer Details
Date
Ink Details
Miran Senanayaka
Select Image Type
Scanned Image
Captured Image
Select Clearing Type
Outward Clearing
Transfer Clearing
Amount in Words: One hundred eighty thousand
Amount in Figures: 180000.0
Payee Details: Miran Senanayaka
Date: 04-03-2020
Date Validation: Success
Amount Validation: Success

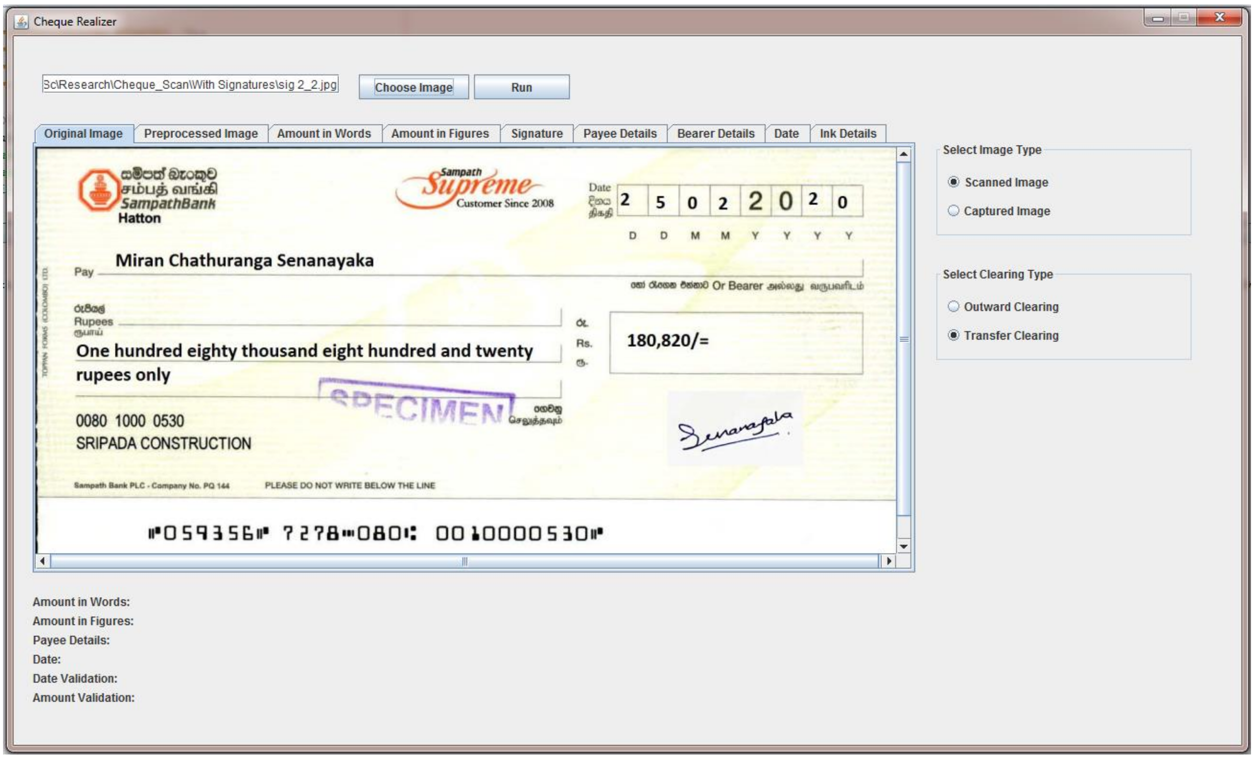

Cheque Realizer
Academic\MSc\Research\Cheque_Scan\IMG_1858.JPG
Choose Image
Run
Original Image
Preprocessed Image
Amount in Words
Amount in Figures
Signature
Payee Details
Bearer Details
Date
Ink Details
0 4 0 3 2 0 2 0
Select Image Type
Scanned Image
Captured Image
Select Clearing Type
Outward Clearing
Transfer Clearing
Amount in Words: One hundred eighty thousand
Amount in Figures: 180000.0
Payee Details: Miran Senanayaka
Date: 04-03-2020
Date Validation: Success
Amount Validation: Success


Cheque Realizer
Academic\MSc\Research\Cheque_Scan\IMG_1858.JPG
Choose Image
Run
Original Image
Preprocessed Image
Amount in Words
Amount in Figures
Signature
Payee Details
Bearer Details
Date
Ink Details
Select Image Type
Scanned Image
Captured Image
Select Clearing Type
Outward Clearing
Transfer Clearing
Amount in Words: One hundred eighty thousand
Amount in Figures: 180000.0
Payee Details: Miran Senanayaka
Date: 04-03-2020
Date Validation: Success
Amount Validation: Success

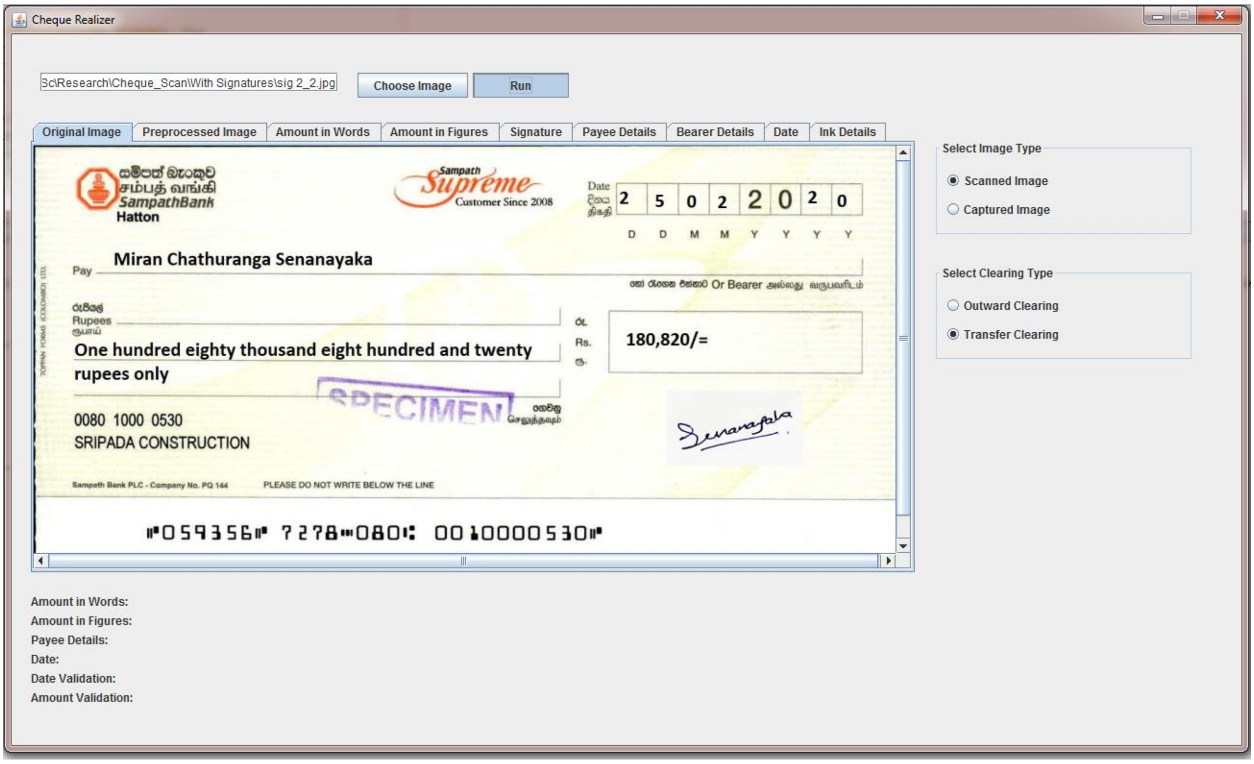
USER INTERFACES OF SCANNED IMAGE



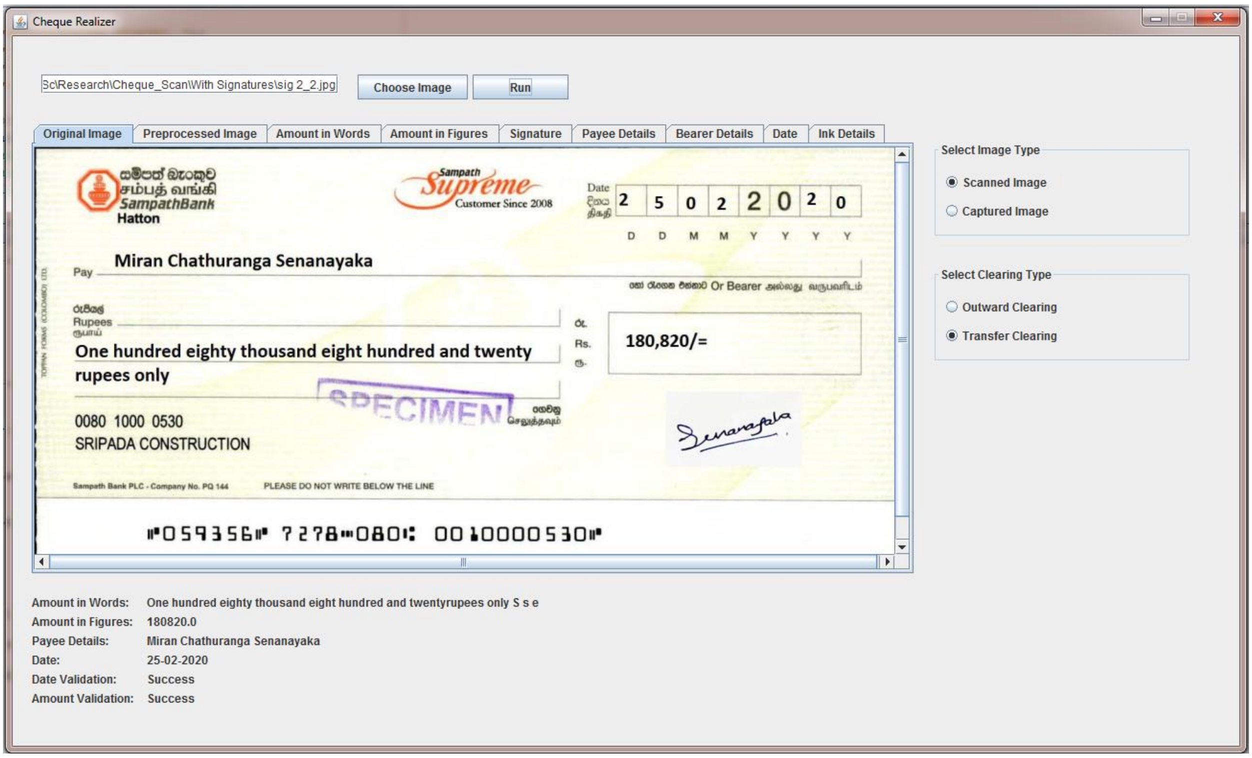


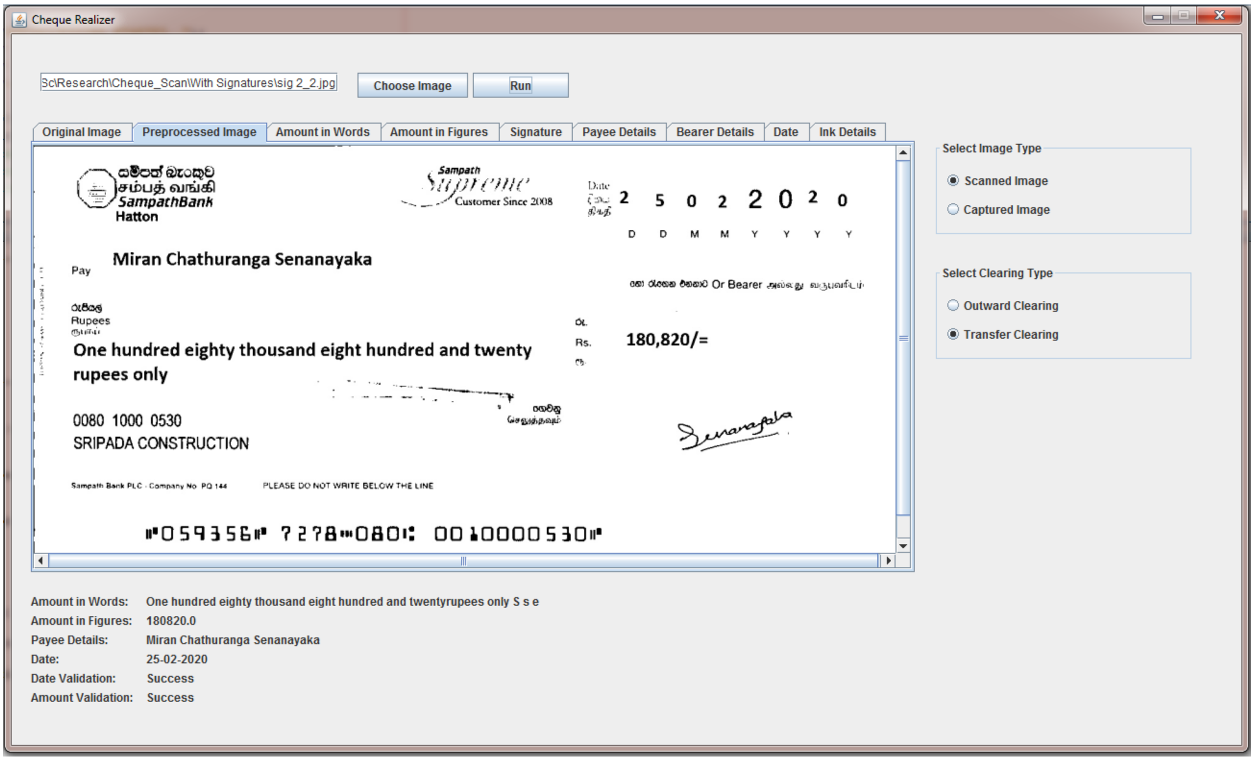


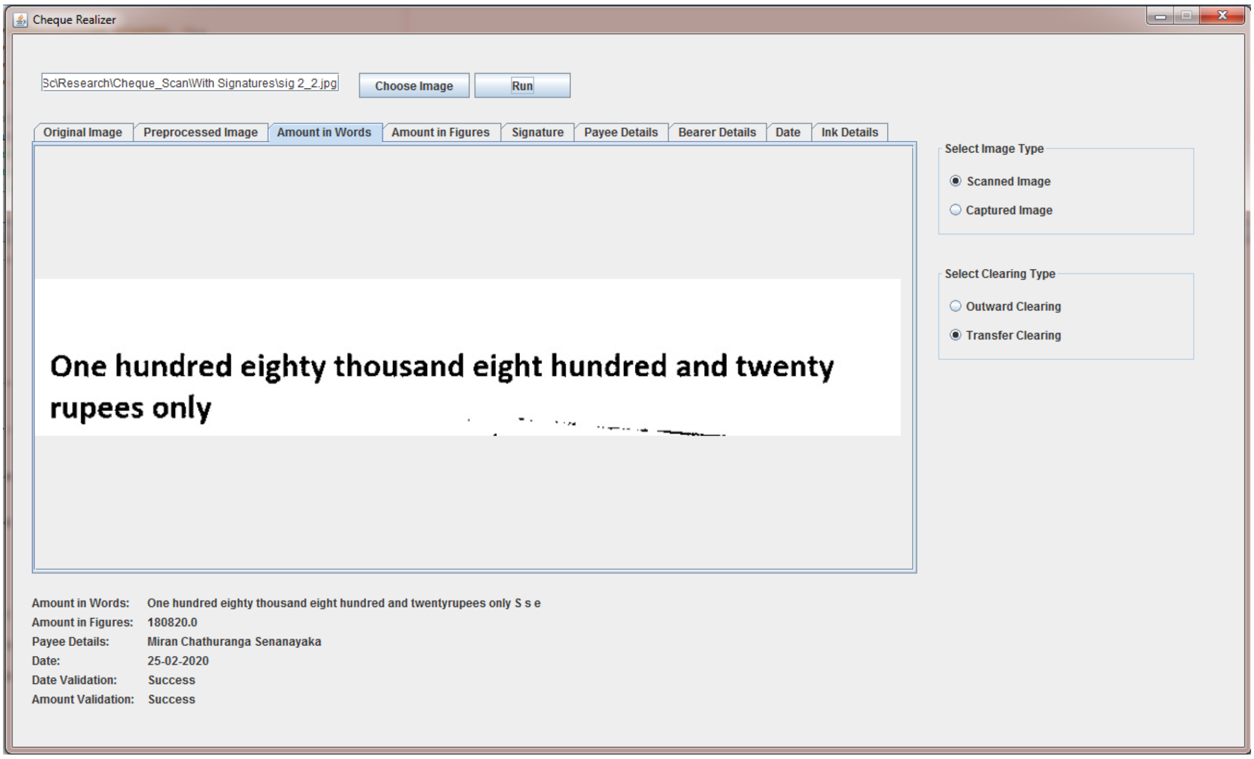
Cheque Realizer
Sc\Research\Cheque_Scan\With Signatures\sig 2_2.jpg
Choose Image
Run
Original Image
Preprocessed Image
Amount in Words
Amount in Figures
Signature
Payee Details
Bearer Details
Date
Ink Details
SampathBank
Hatton
Sampath Supreme
Customer Since 2008
Date
2 5 0 2 2 0 2 0
D D M M Y Y Y Y
Pay
Miran Chathuranga Senanayaka
Or Bearer
Rupees
One hundred eighty thousand eight hundred and twenty
rupees only
Rs.
180,820/=
SPECIMEN
0080 1000 0530
SRIPADA CONSTRUCTION
Sampath Bank PLC - Company No. PQ 144
PLEASE DO NOT WRITE BELOW THE LINE
⑈059356⑈ 7278⑆080⑆ 0010000530⑈
Select Image Type
Scanned Image
Captured Image
Select Clearing Type
Outward Clearing
Transfer Clearing
Amount in Words:
Amount in Figures:
Payee Details:
Date:
Date Validation:
Amount Validation:



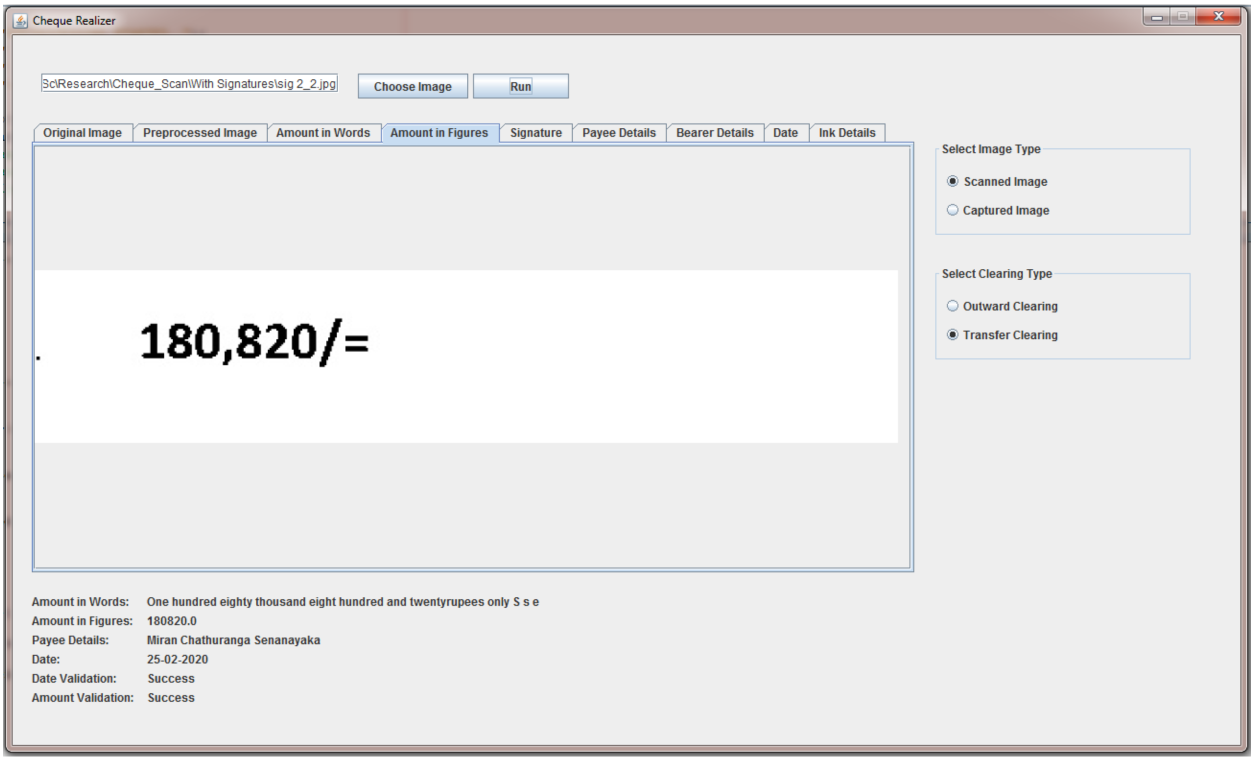
Cheque Realizer
Sc\Research\Cheque_Scan\With Signatures\sig 2_2.jpg
Choose Image
Run
Original Image
Preprocessed Image
Amount in Words
Amount in Figures
Signature
Payee Details
Bearer Details
Date
Ink Details
SampathBank
Hatton
Sampath Supreme
Customer Since 2008
Date
2 5 0 2 2 0 2 0
D D M M Y Y Y Y
Pay
Miran Chathuranga Senanayaka
Or Bearer
Rupees
One hundred eighty thousand eight hundred and twenty
rupees only
Rs.
180,820/=
SPECIMEN
0080 1000 0530
SRIPADA CONSTRUCTION
Sampath Bank PLC - Company No. PQ 144
PLEASE DO NOT WRITE BELOW THE LINE
⑈059356⑈ 7278⑆080⑆ 0010000530⑈
Select Image Type
Scanned Image
Captured Image
Select Clearing Type
Outward Clearing
Transfer Clearing
Amount in Words:
Amount in Figures:
Payee Details:
Date:
Date Validation:
Amount Validation:


Cheque Realizer
Sc\Research\Cheque_Scan\With Signatures\sig 2_2.jpg
Choose Image
Run
Original Image
Preprocessed Image
Amount in Words
Amount in Figures
Signature
Payee Details
Bearer Details
Date
Ink Details
Select Image Type
Scanned Image
Captured Image
Select Clearing Type
Outward Clearing
Transfer Clearing
Amount in Words: One hundred eighty thousand eight hundred and twentyrupees only S s e
Amount in Figures: 180820.0
Payee Details: Miran Chathuranga Senanayaka
Date: 25-02-2020
Date Validation: Success
Amount Validation: Success

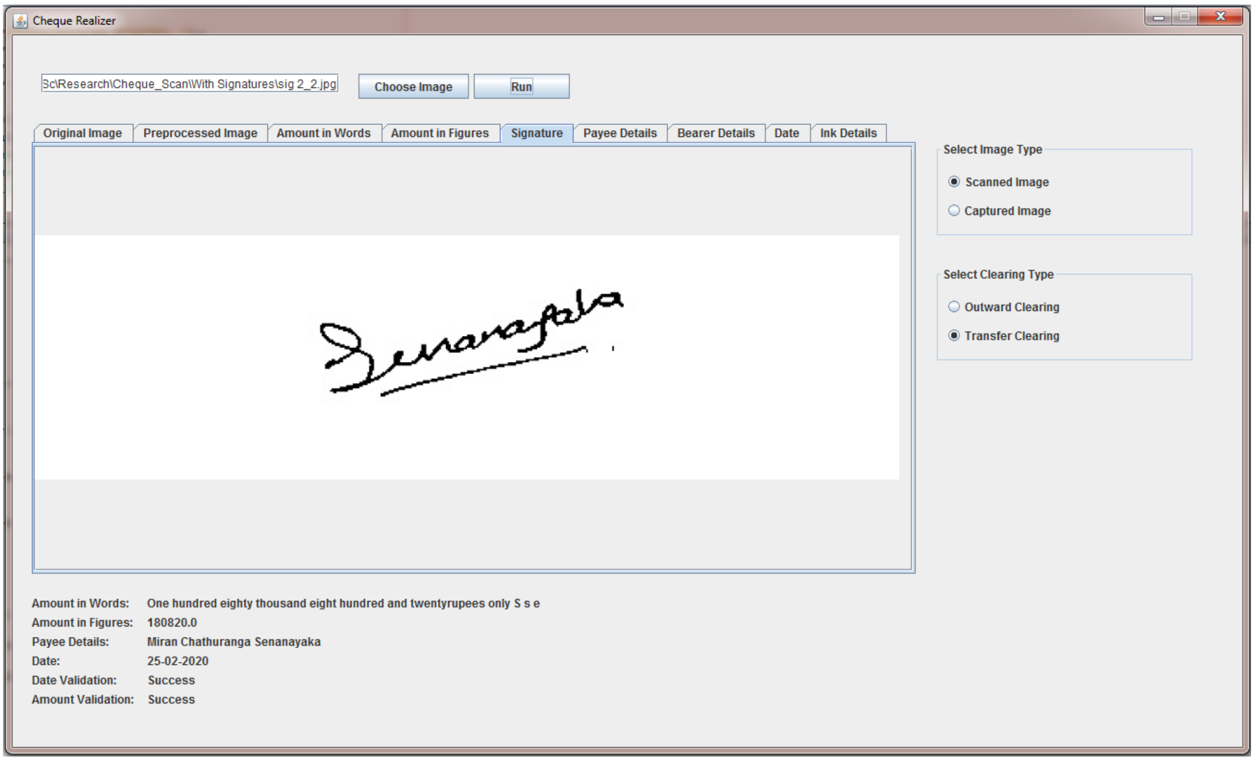

Cheque Realizer
Sc\Research\Cheque_Scan\With Signatures\sig 2_2.jpg
Choose Image
Run
Original Image
Preprocessed Image
Amount in Words
Amount in Figures
Signature
Payee Details
Bearer Details
Date
Ink Details
Select Image Type
Scanned Image
Captured Image
Select Clearing Type
Outward Clearing
Transfer Clearing
Amount in Words: One hundred eighty thousand eight hundred and twentyrupees only S s e
Amount in Figures: 180820.0
Payee Details: Miran Chathuranga Senanayaka
Date: 25-02-2020
Date Validation: Success
Amount Validation: Success

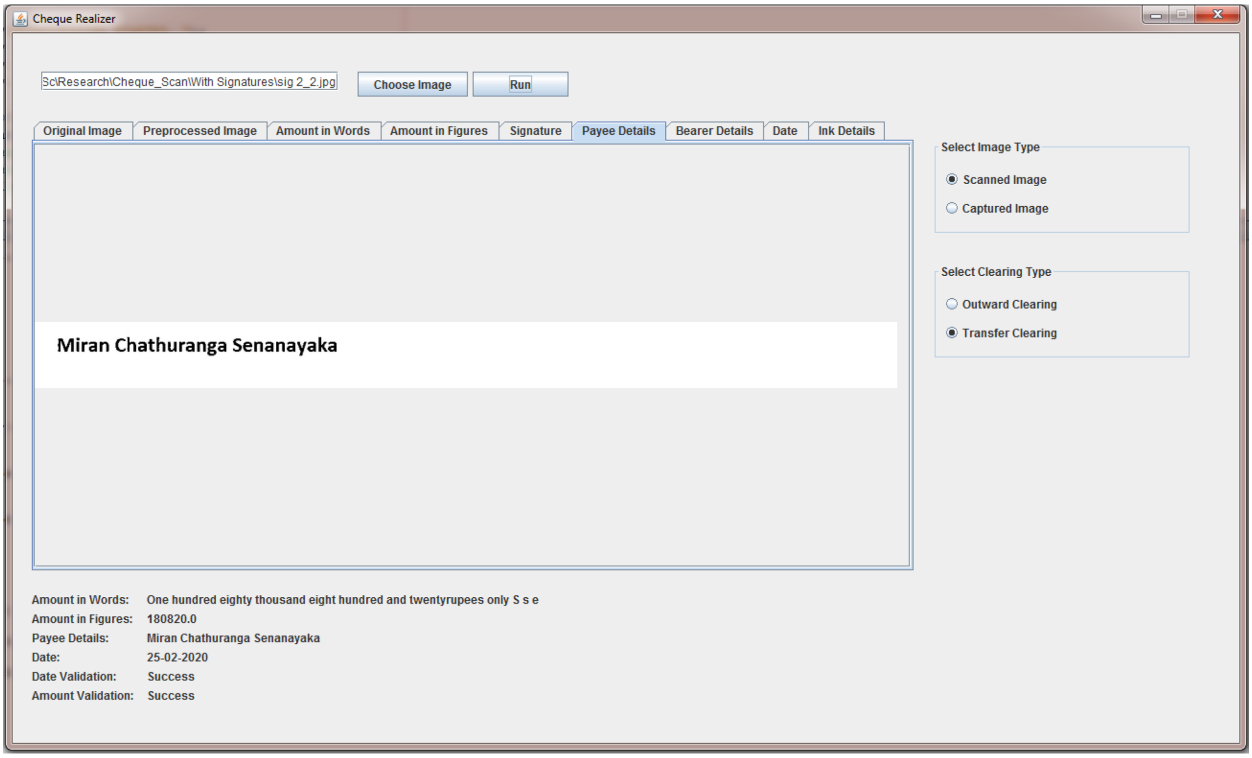


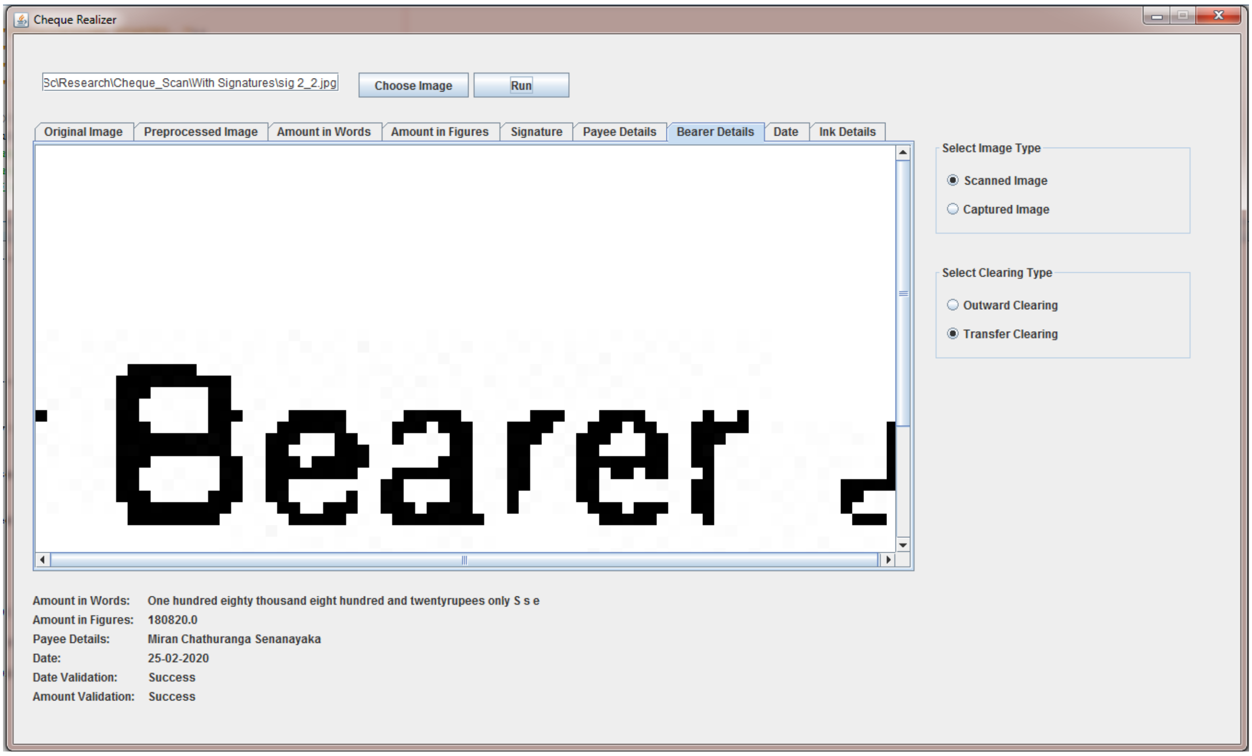
Cheque Realizer
3c\Research\Cheque_Scan\With Signatures\sig 2_2.jpg
Choose Image
Run
Original Image
Preprocessed Image
Amount in Words
Amount in Figures
Signature
Payee Details
Bearer Details
Date
Ink Details
One hundred eighty thousand eight hundred and twenty rupees only
Select Image Type
Scanned Image
Captured Image
Select Clearing Type
Outward Clearing
Transfer Clearing
Amount in Words: One hundred eighty thousand eight hundred and twentyrupees only S s e
Amount in Figures: 180820.0
Payee Details: Miran Chathuranga Senanayaka
Date: 25-02-2020
Date Validation: Success
Amount Validation: Success

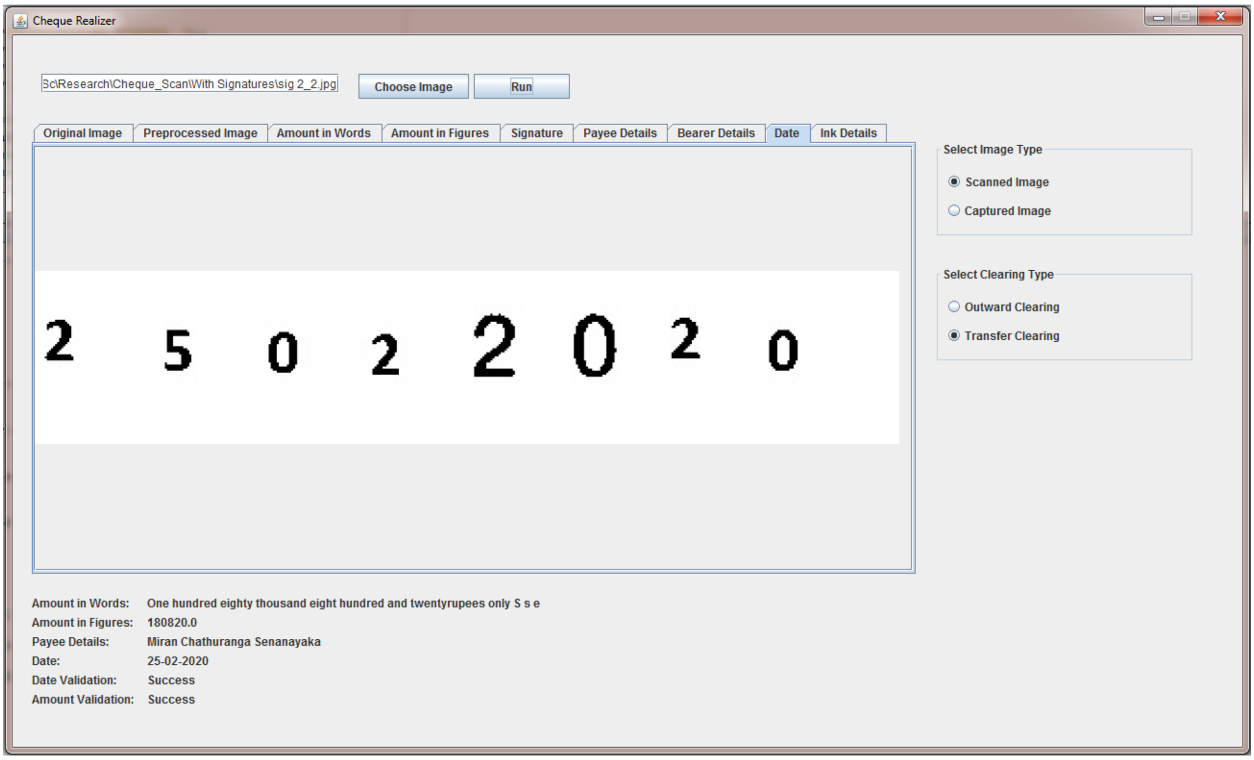


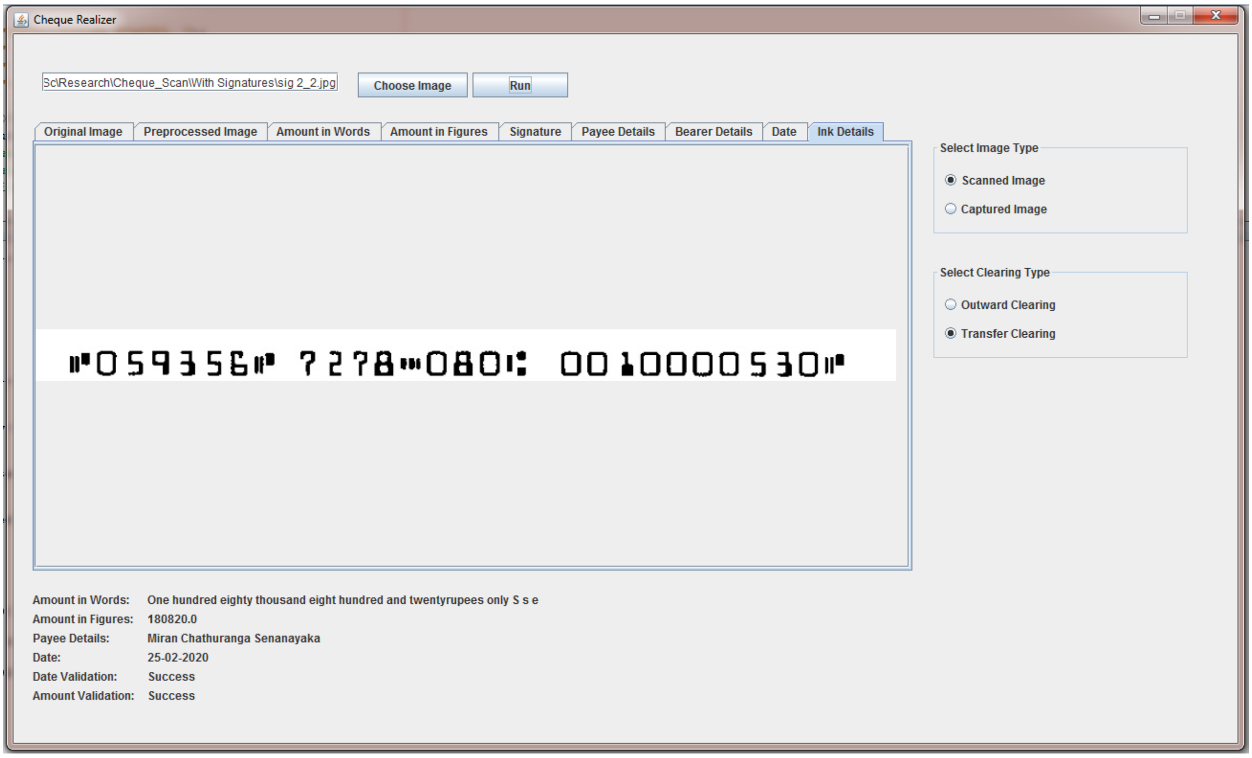
Cheque Realizer
3c\Research\Cheque_Scan\With Signatures\sig 2_2.jpg
Choose Image
Run
Original Image
Preprocessed Image
Amount in Words
Amount in Figures
Signature
Payee Details
Bearer Details
Date
Ink Details
180,820/=
Select Image Type
Scanned Image
Captured Image
Select Clearing Type
Outward Clearing
Transfer Clearing
Amount in Words: One hundred eighty thousand eight hundred and twentyrupees only S s e
Amount in Figures: 180820.0
Payee Details: Miran Chathuranga Senanayaka
Date: 25-02-2020
Date Validation: Success
Amount Validation: Success


Cheque Realizer
3c\Research\Cheque_Scan\With Signatures\sig 2_2.jpg
Choose Image
Run
Original Image
Preprocessed Image
Amount in Words
Amount in Figures
Signature
Payee Details
Bearer Details
Date
Ink Details
Select Image Type
Scanned Image
Captured Image
Select Clearing Type
Outward Clearing
Transfer Clearing
Amount in Words: One hundred eighty thousand eight hundred and twentyrupees only S s e
Amount in Figures: 180820.0
Payee Details: Miran Chathuranga Senanayaka
Date: 25-02-2020
Date Validation: Success
Amount Validation: Success


Cheque Realizer
3c\Research\Cheque_Scan\With Signatures\sig 2_2.jpg
Choose Image
Run
Original Image
Preprocessed Image
Amount in Words
Amount in Figures
Signature
Payee Details
Bearer Details
Date
Ink Details
Select Image Type
Scanned Image
Captured Image
Select Clearing Type
Outward Clearing
Transfer Clearing
Miran Chathuranga Senanayaka
Amount in Words: One hundred eighty thousand eight hundred and twentyrupees only S s e
Amount in Figures: 180820.0
Payee Details: Miran Chathuranga Senanayaka
Date: 25-02-2020
Date Validation: Success
Amount Validation: Success


Cheque Realizer
Sc\Research\Cheque_Scan\With Signatures\sig 2_2.jpg
Choose Image
Run
Original Image
Preprocessed Image
Amount in Words
Amount in Figures
Signature
Payee Details
Bearer Details
Date
Ink Details
Bearer
Select Image Type
Scanned Image
Captured Image
Select Clearing Type
Outward Clearing
Transfer Clearing
Amount in Words: One hundred eighty thousand eight hundred and twentyrupees only S s e
Amount in Figures: 180820.0
Payee Details: Miran Chathuranga Senanayaka
Date: 25-02-2020
Date Validation: Success
Amount Validation: Success


Cheque Realizer
Sc\Research\Cheque_Scan\With Signatures\sig 2_2.jpg
Choose Image
Run
Original Image
Preprocessed Image
Amount in Words
Amount in Figures
Signature
Payee Details
Bearer Details
Date
Ink Details
2 5 0 2 2 0 2 0
Select Image Type
Scanned Image
Captured Image
Select Clearing Type
Outward Clearing
Transfer Clearing
Amount in Words: One hundred eighty thousand eight hundred and twentyrupees only S s e
Amount in Figures: 180820.0
Payee Details: Miran Chathuranga Senanayaka
Date: 25-02-2020
Date Validation: Success
Amount Validation: Success


Cheque Realizer
Sc\Research\Cheque_Scan\With Signatures\sig 2_2.jpg
Choose Image
Run
Original Image
Preprocessed Image
Amount in Words
Amount in Figures
Signature
Payee Details
Bearer Details
Date
Ink Details
Select Image Type
Scanned Image
Captured Image
Select Clearing Type
Outward Clearing
Transfer Clearing
Amount in Words: One hundred eighty thousand eight hundred and twentyrupees only S s e
Amount in Figures: 180820.0
Payee Details: Miran Chathuranga Senanayaka
Date: 25-02-2020
Date Validation: Success
Amount Validation: Success