2017

Using number recognition techniques to automate the revenue collection at mini bus terminals: a case of Dodoma municipal

Rweshobora, Kelvin Christopher

The University of Dodoma

USING NUMBER RECOGNITION TECHNIQUES TO AUTOMATE THE REVENUE COLLECTION AT MINI BUS TERMINALS: A CASE OF DODOMA MUNICIPAL

KELVIN CHRISTOPHER RWESHOBORA

MASTER OF SCIENCE IN INFORMATION TECHNOLOGY

THE UNIVERSITY OF DODOMA

OCTOBER, 2017

USING NUMBER RECOGNITION TECHNIQUES TO AUTOMATE THE REVENUE COLLECTION AT MINI BUS TERMINALS: A CASE OF DODOMA MUNICIPAL

By

Kelvin Christopher Rweshobora

A Dissertation submitted in partial fulfillments of the requirements for the degree of

Master of Science in Information Technology of the University of Dodoma

The University of Dodoma

October, 2017

CERTIFICATION

The undersigned certify that they have read and hereby recommend for acceptance by the University of Dodoma a dissertation entitled ***Using Number Recognition Techniques to Automate the Revenue Collection at Mini Bus Terminals: A Case of Dodoma Municipal*** in fulfillment of the requirements for the degree of Master of Science in Information Technology of the University of Dodoma.

.....

Dr. SALEHE MRUTU

(Supervisor)

.....

Dr. HECTOR J. MONGI

(Supervisor)

Date.....

DECLARATION

AND

COPYRIGHT

I Kelvin Christopher Rweshobora declare that, this dissertation is my own original work and that it has not been presented and will not be presented to any other University, for a similar or any other degree award.

Signature.....

ACKNOWLEDGEMENTS

First, I would like to thank the Almighty God for his guidance and provision of a good health during the report writing process. It is by his grace we are able to achieve our goals in life.

Secondly, I would like to give my heartfelt thanks to my supervisors Dr. Salehe Mrutu and Dr. Hector John Mongi, for their involvement in reviewing my research proposal and dissertation. Their constructive comments and encouragements enabled me to accomplish this work successfully.

Much gratitude should go to the management of the University of Dodoma, College of Informatics and Virtual Education, for giving me an opportunity to study at UDOM and also, provide a conducive environment for accomplishing my studies. Finally I would like to thank my class mates for their support and cooperation during the entire study year and dissertation.

However, it should be noted that, any shortcomings on this research are entirely belong to me.

DEDICATION

I dedicated this study to Christopher and Margreth.

ABSTRACT

The Number Recognition (NR), is a modern technique that uses optical character recognition on images to obtain desired characters. The technique involves four steps Image acquisition, pre-processing, character segmentation and Character recognition. This technique has been used in activities like license plate detection as Automatic Number Plate Recognition (ANPR).

The manual revenue collection process currently operating at the Dodoma Municipal mini bus terminals has deprived the Local Government Authority (LGA) of its potential income. This research attempts solve this problem by automating the current revenue collection process at the mini bus terminals in Dodoma Municipal by using, Number recognition techniques and OCR.

The research has adopted a case study research design and simulation in the problem investigation and proposed simulation development respectively. The research has used key informant interviews and observations to acquire a good understanding of the current operations at the mini bus terminals and the necessary requirements to achieve the main goal of the study. Simulation of the Number Recognition System (NRS) was achieved by using Matlab R2017a as a simulation tool on a Dell computer running windows 7 professional, Hard Disk Drive (HDD) 500 Gigabyte (GB), Random Access Memory (RAM) 4 Gigabyte.

The research, achieved a simulation for NRS with an accuracy of 0.988 for the character recognition of the captured Surface and Marine Transport Regulatory Authority (SUMATRA) numbers on mini buses. The researcher recommends a further study on the image acquisition process and messaging alert system, 2to completely automate the process.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

ABBREVIATIONS

ANPR	Automatic Number Plates Recognition
API	Application Programming Interface
DBMS	Database Management System
DOM	Dodoma
EFD	Electronic Fiscal Device
ERP	Electronic Road Pricing
ETR	Electronic Toll Route
GB	Gigabyte
HDD	Hard Disk Drive
IS	Information System
IT	Information Technology
ITM	Intelligent Template Matching
JPEG	Joint Photographic Expert Group
KEC	Korea Expressway Corporation
LG	Local Governments
LGA	Local Government Authorities
LGRCIS	Local Government Revenue collection Information system
MSER	Maximally Stable Extremal Region
NRS	Number Recognition System
NRT	Number Recognition Technique
OCR	Optical Character Recognition
ODBC	Open Database Connectivity
PPP	Potential of Private Public Partnership
RAM	Random Access Memory

RGB	Red Blue Green
RoI	Region of Interest
SUMATRA	Surface and Marine Transport Regulatory Authority
TRA	Tanzania Revenue Authority

CHAPTER ONE

INTRODUCTION

1.1 Background

Local government authorities (LGA) in Tanzania, are responsible for providing socio economical services to its people. To manage this task Tanzania has embarked on the decentralization with the aim of making the system of governance more accountable, more open and transparent, and more democratic (Mgonja & Tundui, 2012). For the local governments to be able to provide these services, they are supposed to have sources of revenue.

Although the decentralization aimed at enabling local governments to support themselves, many local government authorities failed to support their activities and depended on the central government due to unstable and poor planned sources of revenue (Government, 2013). Local governments have many sources of revenue which include Hotel, shops, restaurant, fish licensing fees and bus stand fees to c

Owing to this inability to support themselves, the local government authorities adopted an outsourcing revenue collection as a means to solve the previous problems of revenue collection which resulted into loss of revenue (Lukio, 2016).

Ngowi (2006) explains the potential of Private Public Partnership (PPP) as a way to solve the revenue collection problem. However, this did not entirely solve the problem as some council reported loss of revenue due to an outsourcing revenue collection.

Fjeldstad et.al (2005), reports the existence of corruption and fraud activities in the entire process of outsourcing and mismanagement of the whole process. The report

revealed that private collectors submitted less than what was collected and agreed in the contract. As reported by Mabhuye (2013), collection of market due dropped by average of 47.6%. This reveals the presence of fraud activities in the outsourcing of revenue collection to private collectors.

To prevent fraudulent activities, revenue collection in bus terminal is done with an agent using an Electronic Fiscal Device (EFD). The fee is collected as the mini bus leaves the station and a receipt is handed to the driver or the conductor. This situation gives room for corrupt collectors to benefit by cheating and not actually using the device.

1.2 Statement of the Problem

The revenue collection process in Dodoma Municipal mini bus terminals involves three parties, which are the Municipal council, revenue collecting agent and the bus fee payer who are the bus drivers or conductors. The manual process being used is of two types, electronic method which employs the EFD machine and the manual receipt book process.

The entire amount collected is then to be remitted to the Municipal council, for the agreed period of time on a daily or monthly basis. This is how the ideal system in the bus terminal revenue collection process should operate.

Several automated revenue collection systems have been deployed in developed countries like Canada, Italy, US and India which are Electronic Toll Route (ETR), Telepass, Electronic Road Pricing (ERP) and EZ TAG respectively (Kulkarni, 2014). The Automatic Toll Collection System deployed in India, aided the country in eradicating corruption, kept data centralized and automated the toll process. For

the car owners, it saved time and lowered vehicle operation cost and for the toll operator, it ensured better audit control by using a centralized account and better facilities (Salunke et al, 2013).

The revenue collection process in the Municipal Council currently employs EFD machine although the process is still outsourced to private collectors. The private collector remits the collected revenue to the council on the agreed time. Traces of fraud and corruption activities have been evident as some agents retaining the collected fees (Fjeldstad et al., 2009; Government, 2013).

Good examples are agents contracted to collect passenger and bus fee at the Ubungo bus terminal Dar es Salaam retained approximately 60% of the collected revenue, where as in Mwanza City Council, said 32% of the officially reported revenue collected was retained by the collecting agent.

The above reviewed literature shows an evidence of man being a weak link in areas that required high integrity, honesty and accuracy. This has led to the development of many automated system to counteract his/her weaknesses(Tyson's & Nelson, 2011).

Thus, this research is proposing a solution for the current revenue collection problem by simulating automatic collecting revenue at the mini bus terminals at Dodoma Municipal Council by using Number Recognition Techniques (NRT) and Optical Character Recognition. By so doing would eliminate man who has proven inefficient in the revenue collection process and increase the revenue collected by the respective local government authority.

1.3 Objectives of Study

The general objective of this research, is to automate the revenue collection process in Dodoma Municipal Council mini bus terminals by using NRT and OCR.

1.3.1 Specific Objectives

1. To investigate operations in the revenue collection process that requires automation.
2. To simulate the automatic revenue collection process.
3. Evaluating the performance of the simulated process.

1.4 Research Questions

1. Which revenue collection operation requires automation?
2. How can automatic revenue collection process be simulated?
3. How accurate is the simulated process?

1.5 Significance of the Study

The study, aims at revealing an alternative solution for increasing revenue collection at the bus terminals by eliminating fraud activities in the revenue collection process. Moreover, the designed simulation if developed and implemented would be able to assist the Local Government Authority in monitoring revenue collection, for their development and planning processes. The study will also benefit the academic world by revealing new opportunities to research on the topic at hand.

CHAPTER TWO

LITERATURE REVIEW

2.1 Introduction

This chapter, reviews several documents to observe what other scholars have written. A deep search on the related documents is carried out. It then describes the key terms in a broader perspective. The chapter includes local government authority revenue sources, application of computer systems in revenue collection and monitoring, image processing, number plate recognition systems character segmentation OCR payment techniques and the limiting factors in the current systems.

2.2 Sources of Revenue in LGA

Reliable sources of revenue, ensures the provision of social economic services e.g health, water, education and road services. Sources of revenue for the Local Government Authorities in Tanzania are either internal, external or loans (Shemdoe, 2015). This thesis focuses on the internal revenue collection specifically bus stand fees.

Losses and mismanagement of revenue collection process, has led to the introduction of the outsourcing of revenue collection process to private collectors. The sole purpose of outsourcing the process to the private collection agents, was to increase the revenue collection. On the contrary, the system tends to profit the private collectors (Lukio, 2016). This calls for another method to maximize the collection and management of the process.

2.3 Usage of Computerized Systems

The involvement of computers in our daily activities, has brought many possibilities and simplified man's work to a great extent. The ability of computer system to work efficiently has led the idea of automating functions, carried out by man that require accuracy and lots of working hours. In the security of any system both physical and electrical systems, it has been noted that the human was the weakest link (Tyson's & Nelson, 2011).

One of the efforts the government has done in enhancing the collection and management of revenue collection, is the deployment of Local Government Revenue Collection Information System (LGRCIS) in some major cities in Tanzania which monitors revenue collection (Mccluskey, 2015).

2.4 Image Processing and Recognition

Image processing and recognition, provides possible solution for many systems that require visual monitoring and identification. The field has had a fair share of research. Most of the research conducted is based on License plate recognition (goyal & Bhatia, 2016; Mitra & Banerjee, 2016).

The presence of a diverse nature of number plate format and languages, has hindered the development of a single application that would be deployed worldwide for license plate recognition. This has led to the development of different system for different countries (C. Patel et al., 2013). Several approaches have been used to overcome these problems, Argentina used Intelligent Template Matching (ITM), Australia approached the problem by using both fixed and mobile systems and Egypt, approached the problem by the creation of an organized database to be used for the number plate recognition system (Tran & Nguyen, 2014).

Many researchers, have ventured in the development of effective and efficient character recognition algorithms. Number Recognition Systems date back as far as 1976 where they were first studied in the scientific development branch Police in UK (Memon, 2012; Tran & Nguyen, 2014). The technique has been used in several areas such as the road, boarders and toll gates. License plate recognition, has assisted in solving several queries which include traffic monitoring, stolen vehicle monitoring and managing parking toll (Badr et al., 2011).

There have been many attempts to develop an efficient algorithm for character recognition part of the process. Some of these detection algorithms are Mathematical morphology, structuring element, media filtering and edge detection (Krishna, 2015). The License Plate Recognition process involves three steps which include Image license positioning, character segmentation and character recognition (Puranic, 2016; Tran & Nguyen, 2014).

2.5 Number Plate Recognition Systems World Wide

Introduced in the year 1976, Number Plates Recognition (NPR) has since found a wide commercial applications, making its research prospects demanding and scientifically interesting (Munuo, 2014). Countries like China, Europe, India and Malaysia have had a fair share of application from toll collection, intelligent traffic surveillance systems, law enforcement to car theft tracking (Ma et al., 2009; Munuo & Kisangiri, 2014; Pandey & Pandey, 2014; Tang et al., 2015).

Cosmo (2014) designed an algorithm that serve the private and commercial plates, for Tanzania mainland plates which are yellow and white. The aim was to exclude diplomatic plates which have different color and format.

Cosmos study conducted in Tanzania like other studies carried out in developed and developing countries, aimed at minimizing human involvement meanwhile aiding laws enforcement agents, road toll fee collection processes and to set a ground for other automated systems, such as automated gate controls for authorized/ non-authorized vehicle and record keeping for entry and exit time.

Some of the difficulties that Cosmos encountered, are the variation of Tanzania plates shape, size and color which pose challenges in the detection process figure 2.1.

Figure 2. 1: Tanzania Number Plates (Source: Munuo (2014))

This research has used minibus identification number assigned by Surface and Marine Transport Regulatory Authority (SUMATRA) for the detection process. A few challenges existed since some of numbers may be clear while others are not as shown in figure 2.2.

Figure 2. 2: SUMATRA Bus Identification Number (Source: Own Processing)

Image processing technology has facilitated these studies to a greater extent and enabled the acquisition of new knowledge and the simplification of man's work.

2.6 Image Localization

To acquire a good image, a high resolution camera should be placed in a strategic position to avoid any obstacle. A strategic position in this case, would mean the distance from the vehicle and the height from the ground for a good image capture. License plate localization, is the vital part of NPR as it localizes the number plate from the rest of the vehicle (Naik et al., 2017).

Wen et.al (2011) defined image localization as the detection of the area of interest that is similar to the license plate. Any technique involving localization focuses on edges or the boundary of the region enclosing the number plate. Difficulty of extraction depends on the background and noises in the image. The acquired image passes through different processes to extract the desired region. The resultant output image is fed to the segmentation function.

2.7 Character Segmentation

The second stage in the process, is the separation of each character so that it appears separately for the recognition process. Choudhary (2014) defined the process as a decomposition of an image of a sequence of characters into a set of individual characters. Character segmentation is basically slicing the extracted image according to individual characters that lie on the plate image (Pervez, 2013).

Grayscale conversion enables us to change all the pixels to either just black or white pixels depending on whether the pixels are above or below the defined threshold. This threshold selection has to be a function of the intensity range of the pixels in the image. The average of the minimum and maximum threshold values, is normally enough to optimize the conversion (Siam, 2014). Siam uses horizontal and vertical projections for identification of each number and finally, uses the boundary box to

segment characters. Pervez (2013) also points out that, character segmentation can be based on projection methods, vertical and horizontal projection.

In the two types of segmentation, vertical segmentation on the number plate is performed to obtain vertically segmented characters. To obtain the characters from the plate, horizontal segmentation is performed (Kranthi et al., 2012).

2.8 Optical Character Recognition

OCR as the third stage, is the process of converting printed character into encoded text (S. G. Patel & Vallabhbhai, 2013). Simin (2013) outlined three methods, correlation, structure analysis and neural network. Among the three methods, correlation is straight forward and reliable while the other two are tolerant to font difference and tilt.

Alkhushair (2012) outlines four best approaches for pattern recognition as template matching, statistical classification, syntactic or structural matching and neural network. Among the mentioned approaches, template matching is less resource intensive and time consuming for implementation (goyal & Bhatia, 2016).

The purpose of Optical Character Recognition (OCR) is to categorize optical patterns often contained in a digital image matching to alphanumeric or other characters(Hansen, 2002). The process of OCR involves several steps including segmentation, feature extraction, and classification as disused above.

2.9 Word Confidences

OCR has three properties that can be used to show the accuracy of the obtained results. These properties are OcrCharacter. Confidence, OcrCharacter.WordIsCertain and the OcrCharacter. Leading Spaces Confidence properties (Leadtools, 2017). A

confidence value in OCR ranges from [0, 1] and it is interpreted as a probability (Matlab documentation, R2017a). This value can be used to identify the location of the required text within the image by eliminating the words with low confidences.

2.10 Payment Techniques

Payment mechanisms differ from one payment process to another. The choice of the payment method depends on the ease of access and time required to complete the process. For a busy area like a bus terminal, time is a critical factor to consider. Putting this in mind, several toll collecting booths in the world deployed different methods with accordance to the nature of activities. Payment methods may include cash, pass cards or credit card (Ma et al., 2009). When the user is paying the toll in cash, it requires much time which causes a delay at the toll booth.

Korea Expressway Corporation (KEC) has employed an electronic payment system that enables the driver to use the prepaid/deferred card to directly pay the toll fee. This reduced the delay at toll booth to about 3~4 seconds from 11 seconds as shown in figure 2.3 (Maetal., 2009).

Figure 2. 3: Flow Chart on E-Payment (Source: Ma et al (2009))

2.12 Methodology Used

In this research several methodological techniques were used to achieve the defined objectives.

2.12.1 Research design

According to Trochim (2005), a research design "provides the glue that holds the research project together. A design is used to structure the research, to show how the entire major parts of the research project works together to try to address the central research problem." Thus, a research design is like a recipe. Yin (1994), defines the case study research method “as an empirical inquiry that investigates a contemporary phenomenon within its real-life context; when the boundaries between phenomenon and context are not clearly evident; and in which multiple sources of evidence are used”. John Gerring (2004) in his paper “What is case study and what is it for”, argues that case study is well defined as a thorough study of a single unit aiming at generalizing the findings across a larger set of units.

2.12.2 Design science

Design science, refers to creation and evaluation of Information Technology (IT) artifacts, intended to solve identified problems. It involves rigorous process to solve observed problems, to make research contributions, to evaluate the designs, and to communicate the results to appropriate audiences (Elio et al., 2011; Hevner & Chatterjee, 2010).

A design science research, approaches for developing Information System (IS) research artifacts focus on first clarifying the goals of the artifacts (constructs, methods, models, or instantiations) and then, on building and carefully evaluating

the utility of the artifacts, and to a lesser degree, their reliability and validity (McLaren & Buijs, 2011).

2.12.3 Data collection

Data collection is the gathering of relevant information for the study. The choice of the collection method depends on time, finance and the researcher skills (Kothari, 2014).

Key informant interviews are qualitative in-depth interviews with people whom understand what is going on in the community (UCLA Center for Health Policy Research, 2016). A key informant interview, is a loosely structured conversation with people whom have specialized knowledge about the topic you wish to understand (Education Development Center, 2004)

2.12.4 Sampling technique

Sampling, is the art of selecting a small group to represent the whole universe (Kumar & Phrommathed, 2005). Young (2009) defines sampling as "A statistical sample is a miniature picture or cross –section of the entire group or aggregate from which the sample is taken."

2.12.5 Simulation

Ayash (2007) defines simulation method as a way to "help investigate systems out of experimental domain". He goes on to give reasons for the usage of simulation methods as, the presence of financial, environmental or time constraints. Simulation has to mimic the actual environment in order for valid results to be drawn from it.

2.13 Limiting Factors for the Existing Systems

Several factors have been noted which hindered the performance of a good number recognition system. Among them, are environmental and others are physical. Lighting or illumination of the number plates can hinder the performance. On the physical aspect, the tilt of the number region, ambiguity of the numbers, variation of font size and missing characters also, affect the performance of the system (Puranic, 2016; Tran & Nguyen, 2014).

CHAPTER THREE

METHODOLOGY

3.1 Research Methodology

This chapter, discusses on the research methods used in the study. It gives a detailed explanation on the research design, research site, and data collection method. It then goes on to show the validity of the tools and method used, reliability and finally how obtained data were analyzed.

3.1 Research Site

In this research the researcher studied the Dodoma mini bus terminal of Jamatini and Sabasaba, to gain the understanding of how the current revenue collection system works. The two terminals were ideal for the study because both are revenue collection stations for the municipal. The research site is located in Dodoma region at Dodoma Municipal Council which is bordered by Chamwino district council in the East and Bahi district in the West. The two terminals Jamatini and Sabasaba are located at Tambukareli and Waridani wards respectively.

3.2 Research Design

This study has adopted a case study research design and design science. The researcher studied two mini bus terminals at Dodoma Municipal. This choice was influenced by the nature of the research site which is a single lane of exit. Single lane is essential for the control the number of mini bus leaving the mini bus terminal. This feature enables the designed system to capture each mini bus detail and perform its automated revenue collection process. Time and budget constraints have also contributed to the choice of case study as a research design. Therefore, by studying one unit (Dodoma municipal), the researcher aims to represent large set of units

(other municipals and regions). Thus case study is cost efficient and relevant for the study.

The researcher has also used design science approach by creating an artifact that mimics the operation of image acquisition and automatic revenue payment. The artifact was created with the focus on a modified four step frame work that was initially designed by (Kopparapu, 2015).

3.3 Data Collection Methods and Procedures

In this research, the researcher has used quantitative research methods in gathering the requirements for the process to be simulated. The researcher relied on both primary data and secondary data. To collect these data, observations, key informant interviews and library study were conducted.

Common techniques used in conducting such interviews are telephone interviews and face to face interviews(UCLA Center for Health Policy Research, 2016). This research has used face to face interview technique in gathering information.

3.4 Investigating the Revenue Collection Process

To obtain the primary data on the operations of the current revenue collection process, the researcher employed observations, key informants and document reviews as research tools. The researcher has visited the research sites for the key informant interview as well as observations to obtain data for the study.

From the observations, the researcher observed the working processes of the current revenue process and identified areas that required automation so as to improve the revenue collection. Thereafter, key informant interviews with practitioners (the fee collectors) were conducted to obtain insights of the difficulties they encountered and

areas they thought would simplify their work and improve the revenue collection process. The researcher obtained insights of the difficulties faced by the fee collectors and even the trickery ways used by fee payers to evade fee payment.

Finally, the researcher conducted a document review to obtain the relevant information for the creation of dataset required in the model. To obtain the additional data to be used, the researcher used a digital camera for capturing images of the side bus numbers (SUMATRA number) that were used in the designing of the images used in the model to be designed.

3.5 Mini Buses Selection

The researcher has used a purposive sampling technique to achieve the targeted population. The target population in this study was mini buses at Jamatini and Sabasaba bus terminals. Sampled mini buses were chosen based on the clarity/ visibility of their SUMATRA number. This sampling technique was appropriate for the research as it saved time and enabled the researcher to get the right minibus with correct clearly visible SUMATRA numbers.

Figure 3. 1: Captured SUMATRA Number (Source: Own Processing)

3.6 Revenue Collection Process Simulation

To address the second objective of this study, the researcher has used simulation method as part of the methodology.

The researcher has used Matlab as a simulation tool for the thesis. The software choice was influenced by the presence of a rich literature of image recognition systems developed and the presence of required libraries, for the image processing to demonstrate a complete simulation of the revenue collection process, database to represent the electronic payment process was designed by using Microsoft access database and linked to Matlab by using Open Database Connectivity (ODBC) drivers. This research has used a laptop computer for the demonstration and coding of the system. The computer specifications will be a 4 GB RAM computer with an Intel i3 processor or higher.

Number Plate Recognition as discussed in the literature, involves three to four stages/ steps. This research adopts the four step frame work shown in figure 3.2.

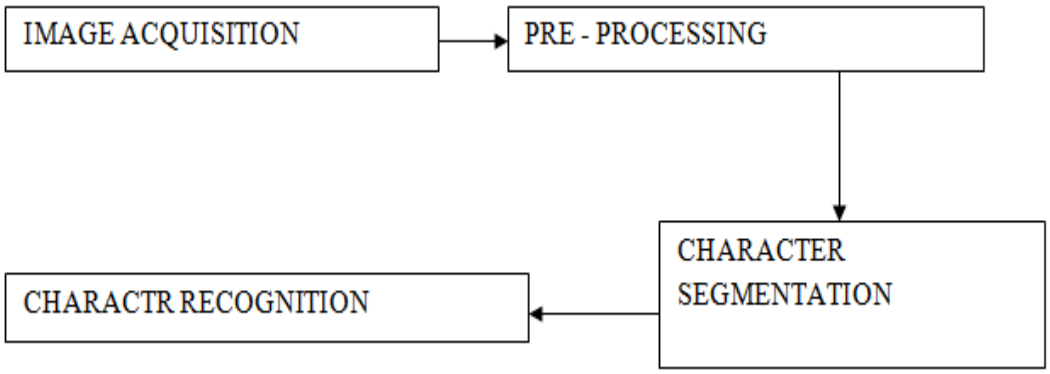


Figure 3. 2: Proposed Flow Diagram (Source: Modified from Kopparapu (2015).

3.6.1 Image Acquisition

The researcher has used a mobile phone TECNO Y4 to capture image in the Red Green Blue (RGB) format for the study figure 3.1. Images were saved in a Joint Photographic Expert Group (JPEG also abbreviated as JPG) file format. The image acquisition process, involved the retrival of images from a file in the computer system. These images were custome designed to meet the requirements for the image acquisition process.

3.6.2 Pre-processing

After capturing the image, it is then processed to convert it from the RGB to gray scale. Since machine best understands grey images which are defined as a two dimention function *f(x,y)* where *x* and *y* are spatial coordinates and *f* is the lignt intensity at the region. A gray scale image is a black and white representation of the converted true color image. Gray scale conversion was achieved throuh binalization. Detect MSER features function which was then applied to the gray scale image to obtain Region of Intrest (RoI) through thresholding.

3.6.3 Character Segmentation

After obtaining the ROI, areas with the same color threshold were detected and considered as character segments. A mask is created to compare the region detected with the regions in the origional image.The create Mask method returns a binary image the same size as the input image, containing1s inside the ROI and 0s everywhere else (Matlab documentation, R2013). This mask enables us to easily perform OCR from the captured image.

3.6.4 Recognition of Character

The final steps, is the identification of the recognized numerals and match them to the pre existing database and check if the number exist. The numbers were recognised and extracted from the image using the OCR function of matlab.

Only numbers were checked because the Sumatra mini bus identification number contains a number and a region short code. The rest of the characters e.g Dodoma (DOM) and SUMATRA that apear are simmilar for all the vehicles thus they are not of great interest in identifying the vehivles tax payer details.

3.6.5 Database

The researcher, developed a simple database that contained two tables to mimic the operation in the payment process. One represented SUMATRA and the other representing the revenue collecting agency e.g. Tanzania Revenue Authority (TRA). The earlier contained information of the tax payer including the full name, SUMATRA number and amount pre-deposited for the day transaction at the station. The later has first name and amount deducted from the pre-deposited account in the SUMATRA table.

The researcher used Microsoft Access database. Prefference to using this Database Management System (DBMS) was due to familiality, the ease of use and availability of the Open Database Connectivity (ODBC) drivers, to link the database to Matlab for simulation purposes.

ODBC is an open standard Application Programming Interface (API) for accessing a database. ODBC statements in a program, allows access to files in a number of different databases, including Access, dBase and DB2.

3.6.6 Payment Method

The mode of payment for the proposed model, was pre-paid e-payment where the mini bus owner has a payment account. The mini bus owner deposits amount in his or her account for the daily payment at the mini bus terminal. This pre-deposited amount can be seen in the Sumatra figure 4.6. Therefore, each trip made, a specified amount id deducted from this account.

Unlike the current payment by cash system, the electronic payment enable the automatic payment of the gate fee /toll once the mini bus leaves the bus terminal. The payment problem solves the delay and change issuing problems as stated by the collecting agents.

3.7 Research Limitations

Some of the limitations that the researcher encountered were, getting permission to take photos and observation time. For the permission to take photos the researcher used data collection letter to prove the usage of the gathered data was for the educational purposes only. The researcher had to arrive as early as 5:00 am in the morning at the mini bus terminal and left late at night. The researcher used two sets of questionnaires one for tax collecting agents and the other for mini bus drivers.

3.8 Ethical Considerations

The researcher ensured that, a permission from the bus drivers to take photos of the SUMATRA number on mini buses. The obtained photos were only used for the research purpose and not otherwise. He also observed for the consent of fee collecting companies to interview their agents.

CHAPTER FOUR

FINDINGS AND DISCUSSION

4.1 Introduction

This chapter, presents the discussions of the findings relating to the objectives of the study, to answer the study questions. The chapter is organized in three parts. The first part discusses the findings from the observation and the key informant interview. The second part discusses the simulation of the revenue collection process and finally the third part discusses the achievements from the simulation developed.

4.2 Observation and Key Informant Results

This study objective, involve the responses from the questions answered in the key informant interviews and results from the physical observation which were conducted by the researcher.

The aim of the key informant interviews was to achieve the insights of the day to day working of the current manual revenue collection process, at the Dodoma Municipal mini bus terminals. The researcher visited the two mini bus terminals which manually collect gate fees, namely Jamatini mini bus terminal and Sabasaba mini bus terminal and four mini bus drivers.

4.2.1 Collecting Agents

Both the collecting agents said that, only the SUMATRA numbers were needed for the record keeping, although they did not appear on the receipt since were written on a separate notebook. The Respondents also said that, only one agent was assigned at a single bus terminal. To improve the revenue collection process, respondents suggested that, the receipts which were produces by the EFD machines should include the SUMATRA numbers for the respective mini bus. This would in turn,

assist them in the submission of the revenue collected and simplify the identification tax evaders. One other problem was issuing change from the amount paid by the driver or conductor of the mini bus.

4.2.2 Mini Bus Drivers

Interviewed bus drivers said that, only SUMATRA numbers which resided on the side of the mini bus were required for the payment process to be completed. They also pointed out the amount charged per trip was 500/- Tsh which was collected by the collecting agent as they exited the mini bus terminal.

The Mini bus drivers' explanation points out that SUMATRA numbers were a key factor for the accomplishment of the payment process.

Table 4. 1: Respondents for Key Informant Interview

Respondent name	Title	Research site
R1	Bus driver	Jamatini
R2	Bus driver	Jamatini
R3	Bus driver	Sabasaba
R4	Bus driver	Sabasaba
R5	Collecting agent	Jamatini
R6	Collecting agent	Sabasaba

Source: Own Processing

Terms R1 to R6 represent interviewed respondents. This name coding was used to preserve their identities.

4.2.3 Observations

From the observations, the researcher observed that, each mini bus left the terminal was charged a gate fee for every trip that it made. The driver or conductor was then issued an EFD receipt for the payment from the collecting agent and the agent also notes down the SUMATRA number on a separate note book.

From the information obtained through key informant interview, a separate note book was required for recording the paying mini buses simply because the receipt issued from the EFD machine did not include the SUMATRA number. The receipt includes information like receipt number, date and time issued and name of the issuer figure 4.1.

Figure 4. 1: Sample EFD Receipts (Source: Own Processing)

Therefore, from the observations and interviews conducted, the researcher deduced that, processes which required the automation were the payment method and the image capture to extract the SUMATRA number. Thus, the required data for the

development of the model include SUMATRA number, mini bus owners name and the deposited amount.

4.3 Simulation Revenue Collection Process

This study objective 4.3 relies on findings from the first objective. To simulate the revenue collection process, a process for the identification of a mini bus SUMATRA number was required which had functions in table 4.2 and a database with two tables one to represent SUMATRA and the other to represent TRA figure 4.6 and 4.7 respectively.

Table 4. 2: Simulation Functions and Description

MODEL FUNCTIONS	DESCRIPTION
Image capture	Used to retrieve image from the designated folder.
Clear	Clears the display axis and text fields for a new image capture.
Exit	Used to close the system.
Alert	Give a notification for insufficient funds and absence of the number in the database.
Status	Displays the payment status of the mini bus.
Top-up	Used to increment the user account balance.
Payment	Deducts the required amount from SUMATRA mini bus account and increment TRA mini bus account.

Source: Own Processing

4.3.1 SUMATRA Number Identification

The proposed simulation, was developed with a focus on the four step frame work as adopted from (Kopparapu, 2015). The four steps include Image acquisition, Image pre-processing, character segmentation and character recognition.

4.3.2 Image Acquisition

Image was acquired from a set of pre-designed images. Images are custom designed to meet the system specifications which are font style (Arial Unicode MS), font size (36) and character spacing (Normal). To acquire the image from a file in the computer, the imread() function of Matlab was used. The image should reside in the same location or a path has to be defined.

>> **image=imread('num6.jpg');** # Read image for processing

>> **inshow(image);** # Displays true color image

On reading the image, the type must be explicitly defined .JPEG or .JPG. From the above code snip, the following image was obtained figure 4.2.

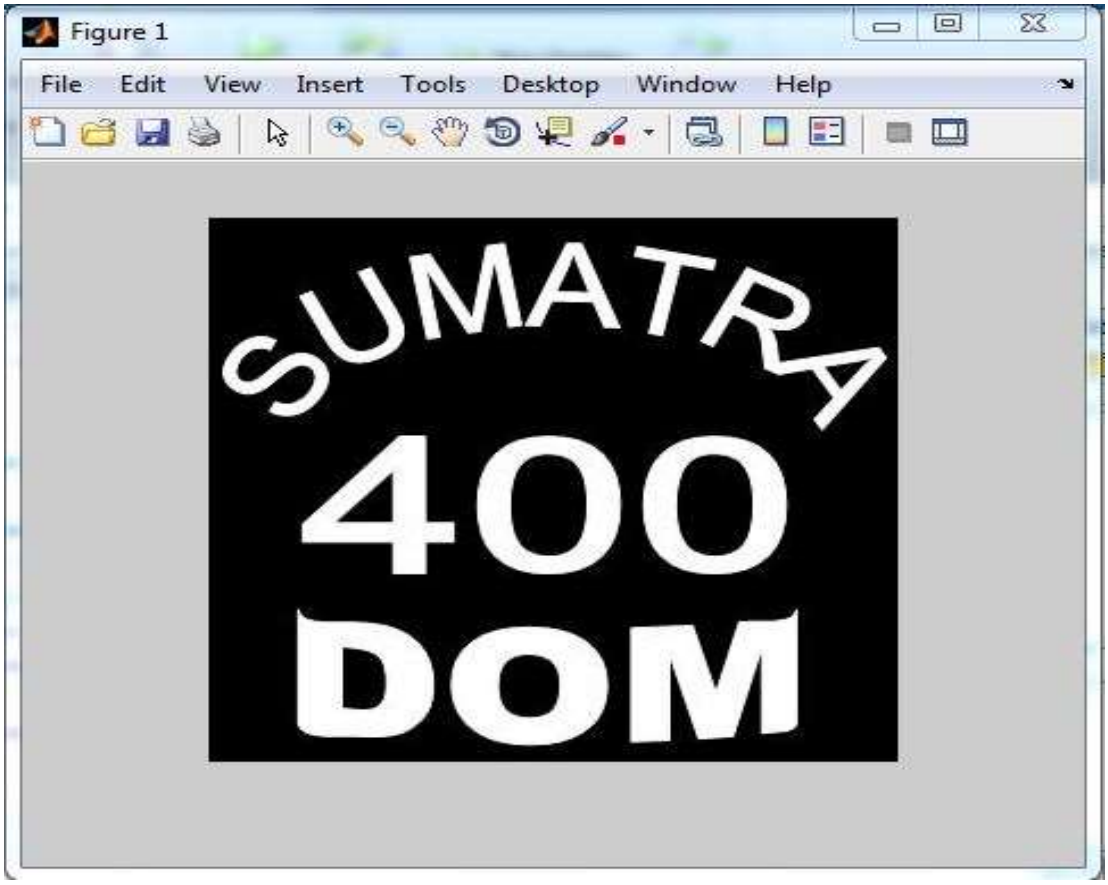


Figure 4. 2: Acquired Image (Source: Own Processing)

4.3.3 Pre-processing

The acquired true color image Red Green Blue (RGB) is converted to the grayscale intensity image by the rgb2gray ( ) function that converts RGB images to grayscale by eliminating the hue and saturation information while retaining the luminance. This is achieved by forming a weighted sum of the R, G, and B components:

0.2989 * R + 0.5870 * G + 0.1140 * B

>> **grayimage=rgb2gray(image); #** Converts the true color image to a gray scale image

>> **imshow(grayimage); #** Displays the grayscale image figure 4.3.

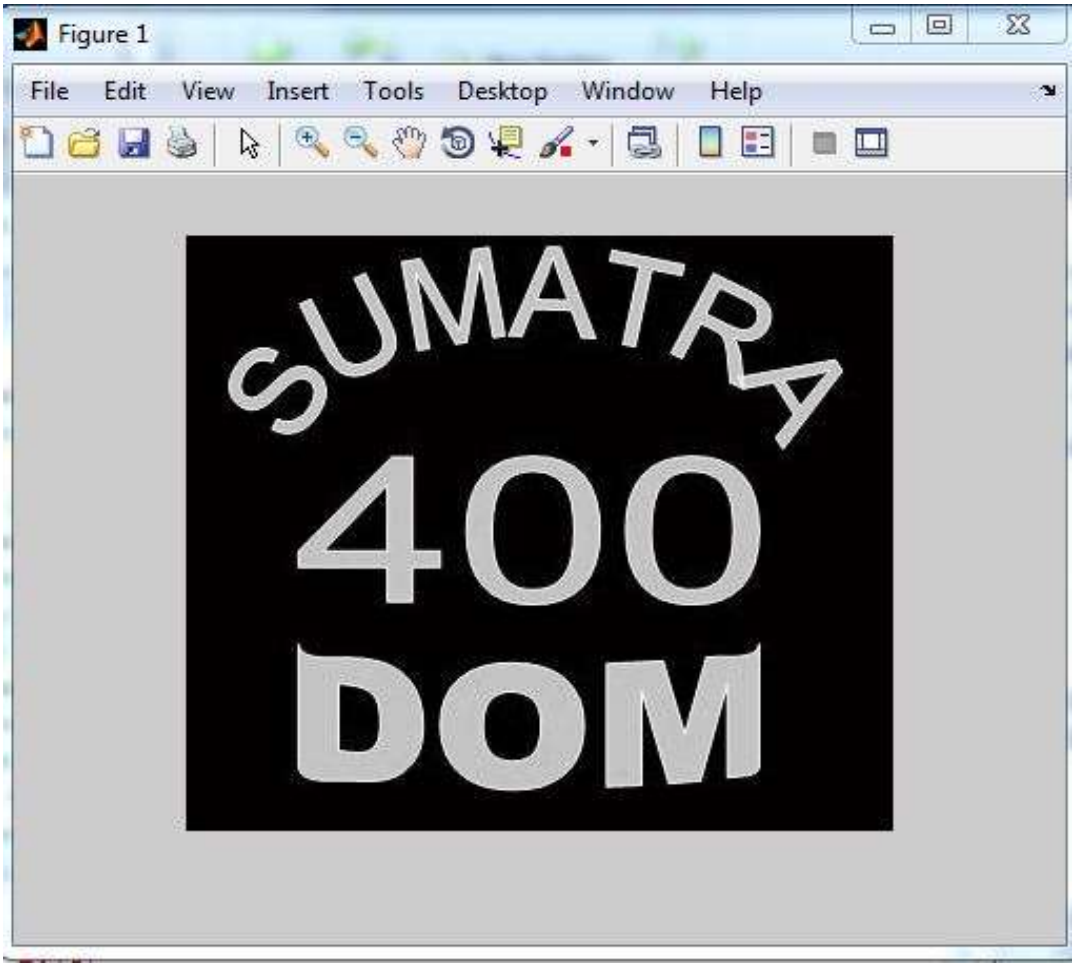


Figure 4. 3: Gray Scale Image (Source: Own Processing)

4.3.4 Character Segmentation

The image is then checked for the region with a specified threshold value for the intensity to detect areas/ regions which are likely to contain the characters. Through threshold, the none required regions of the image is filtered out for the next stage. Thus non character part of the image is segmented from the ROI through the create mask method.

>>**imgRegions=detectMSERFeatures(grayimage,'ThresholdDelta',0.5);** # threshold conversion

smtRegionsPixels=vertcat(cell2mat(imgRegions.PixelList));

imshow(image); hold on; # image display

plot(imgRegions,'showPixelList',true,'showEllipses',false); plotting the threshold results on the image figure 4.4.

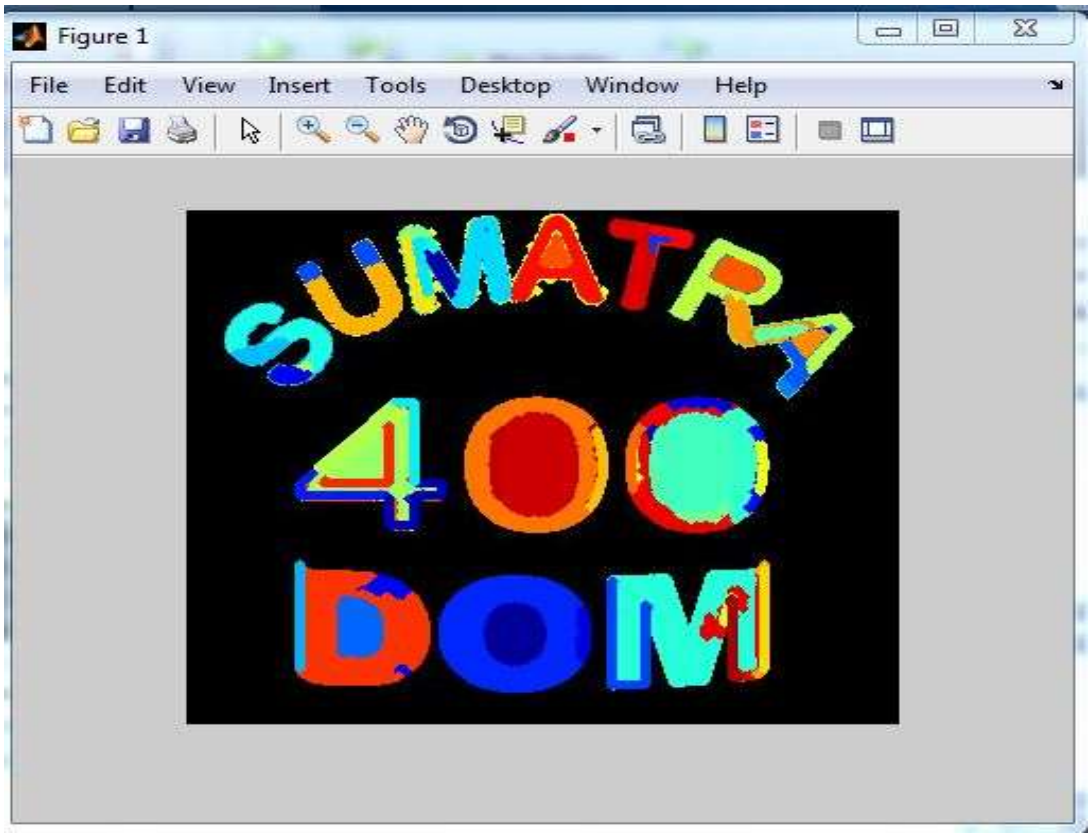


Figure 4.4: Threshold Image (Source: Own Processing)

4.3.5 Character Recognition

After the characters are segmented through the threshold and Maximally Stable Extremal Regions (MSER) feature detection, OCR function is applied to recognize and capture the characters which are then stored in the ocrtext object. By using the ocrtext.Word ( ) function, the required text can be displayed or used by other function in the developed system model. The "**OCRTXT**" object captures all text on the image. To obtain the required text, the "**.Word**" property is used. >> text=ocrtxt.Word ( ). In order to obtain the desired character, the object **OCRTXT.Word(2:3,1)** was used to specify the character from the row two and three of the first column in the OCRTXT object figure 4.5.

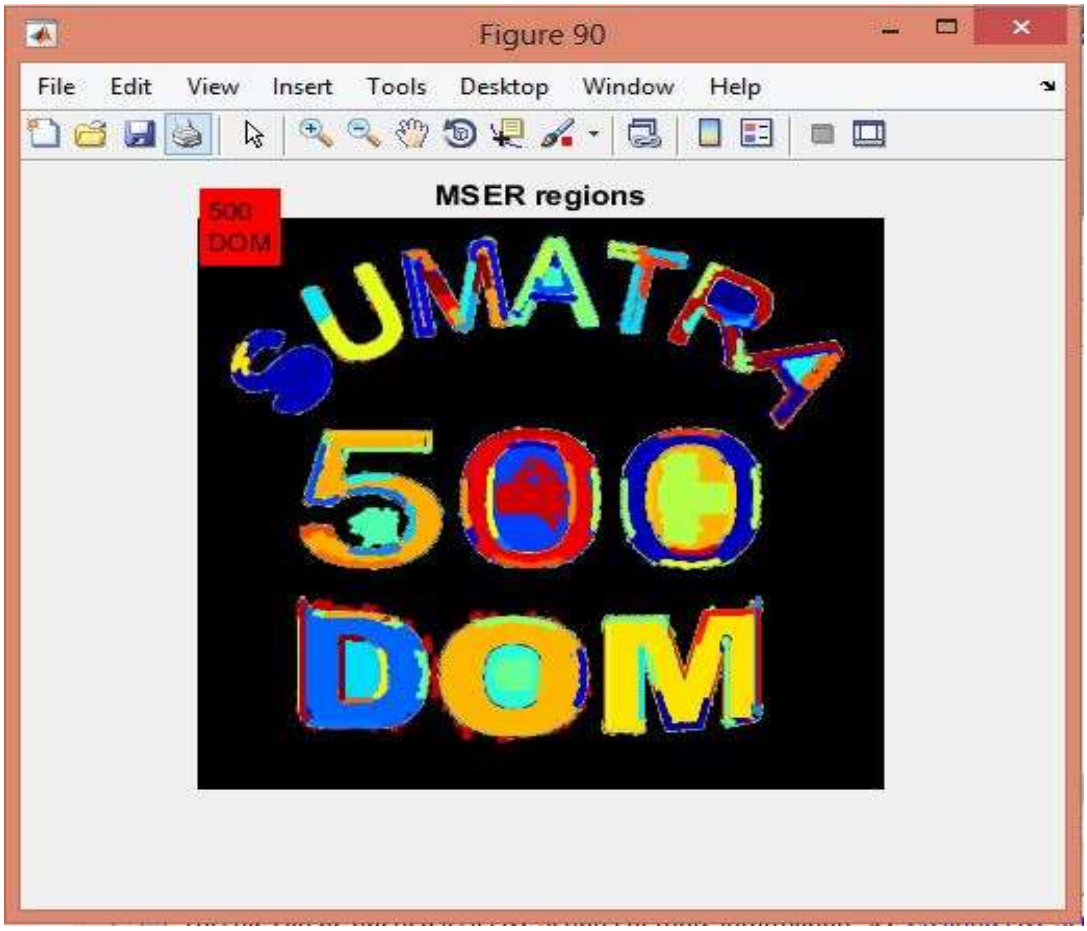


Figure 4. 5: Detected Characters (Source: Own Processing)

Figure 4.5 displays the output of Maximally Stable Extremal Regions (MSER) and OCR on the grayscale image.

The following table 4.3 shows the Word Confidence for all the images that were passed through the model.

Table 4. 3: OCR Results for Recognition Accuracy

S/N	SUMATRA NUMBER IMAGE	RECOGNISED NUMBER	OCR WORD CONFIDENCES
1		400 DOM	0.8908 (Sumatra number) 0.8978 (region)
2		500 DOM	0.8908 (Sumatra number) 0.8978 (region)
3		490 DOM	0.9111 (Sumatra number) 0.8776 (region)
4		7000 DOM	0.9278 (Sumatra number) 0.8706 (Region)
5		900 DOM	0.8912 (Sumatra number) 0.8741 (region)

Source: Own Processing

The above table 4.3 displays the probability of returning the accurate reading from the captured SUMATRA number. The word confidences ranging from 0 to 1, shows the probability of the captured number being correct. The above table 4.3 shows that, the designed model has a recognition accuracy of 0.90234 which is an average of the five attempts, to recognize different SUMATRA numbers.

Finding the average of recognition show the extent to which the obtained results are accurate. The sum of the overall Word Confidences is obtained and divided by the

number of images captured and processed to obtain the degree of accuracy of the recognition process Equation 1.1.

$$\text{AVERAGE RECOGNITION WORD CONFIDENCES} = \frac{\text{SUM OF WORD CONFIDENCE}}{\text{NUMBER OF PROCESSED IMAGES}}$$

Equation 4.1: Average of Word Confidences

4.3.6 Database for the Simulation

The database designed for the proposed simulation contains two tables.

1. The SUMATRA table which contains fields (SID, snumber, amount and Region) Table 4.2.
2. The TRA table which contains the fields (TRAID, snumber and Amount Paid) Table 4.3.

Table 4. 4: SUMATRA Table Details

FIELD NAME	KEY	DATA TYPE	DESCRIPTION
SID	Primary key	Auto number	Primary key in Sumatra table
SNumber		Number	Sumatra Number
OwnerN		Text	Name of the Snumber
amount		Currency	Per-paid amount for a day transaction.
region		Text	Region short name in which the mini bus operates.

Source: Own Processing

Table 4. 5: TRA Table Details

FIELD NAME	KEY	DATA TYPE	DESCRIPTION
TRAID	Primary key	Auto number	Primary in the TRA table.
SNumber		Number	Sumatra Number
AmountPaid		Currency	Amount automatically paid by the mini bus.
NOTrips		Number	Number of trips the mini bus had gone.

Source: Own Processing

SUMATRA

SID	snumber	ownerN	amount	region
1	500	KELVIN	1100	DOM
2	400	juma	5300	DOM
3	490	stuwart	1550	DOM
4	900	Angle	2700	DOM
5	7000	simon	5700	DOM

Figure 4. 6: Design View of SUMATRA Table (Source: Own Processing)

TRA

TRAID	snumber	AmountPaid
1	500	0
2	400	0
3	490	0
4	900	0
5	1000	0
6	7000	0

Figure 4. 7: Design View of TRA Table (Source: Own Processing)

Figure 4. 8: Database Tables (Source: Own Processing)

4.3.7 Payment Process

Payment process involves the linking of the SUMATRA number acquisition system and database. Microsoft Access database was linked to the system in Matlab via the ODBC driver. A data source "Test" was created to access data from Microsoft Access database. Below is the series of payment process flow to illustrate the process. The two flow charts illustrate the two scenarios.

1 Processes occurring when the recognized Snumber exists in the SUMATRA database and when the Snumber does not exist in the SUMATRA database figure 4.9.

2 Processes occurring during the payment process. When the recognized mini bus had sufficient balance for transaction and when it had insufficient balance for the transaction figure 4.10.

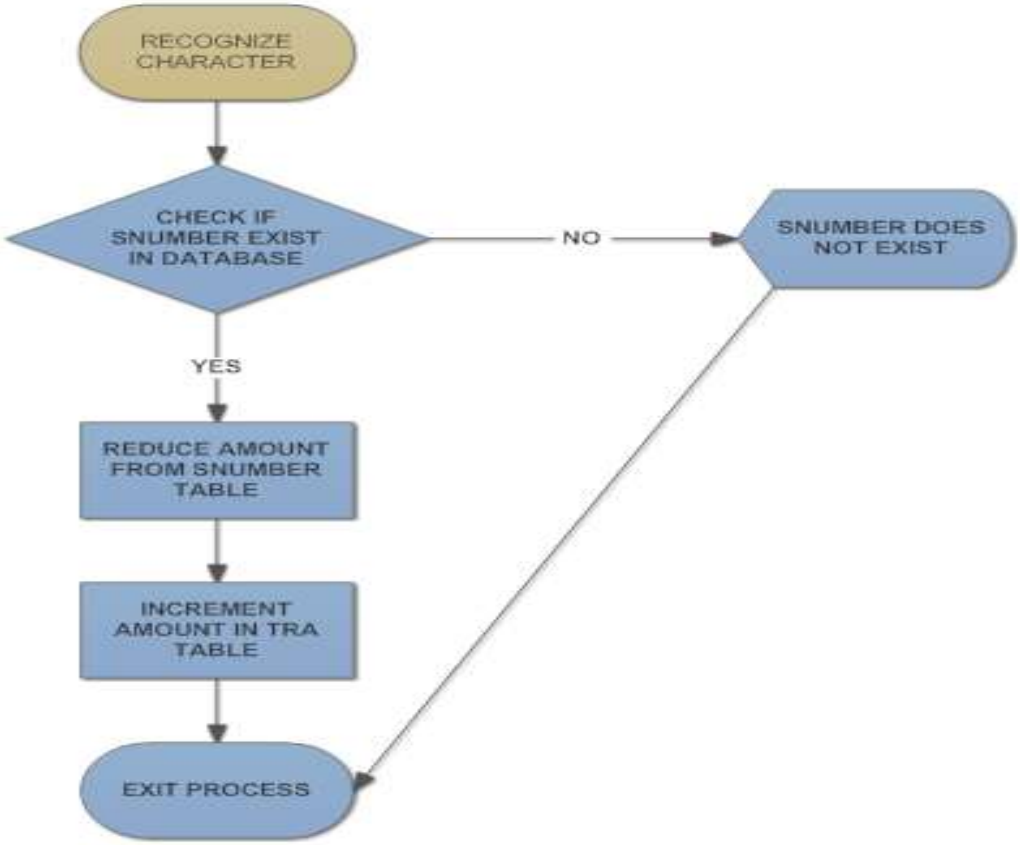


Figure 4. 9: Payment Process (Source: Own Processing)

The figure 4.9 illustrates the proposed e-payment process for the simulation. In the figure, (snumber) stands for SUMATRA number which is a unique number assigned to every mini bus operating in the Dodoma Municipal. Therefore, if the recognized number does not match the snumber in the database, the model system alerts the responsible person, to take action against the mini bus with the identified SUMATRA number.

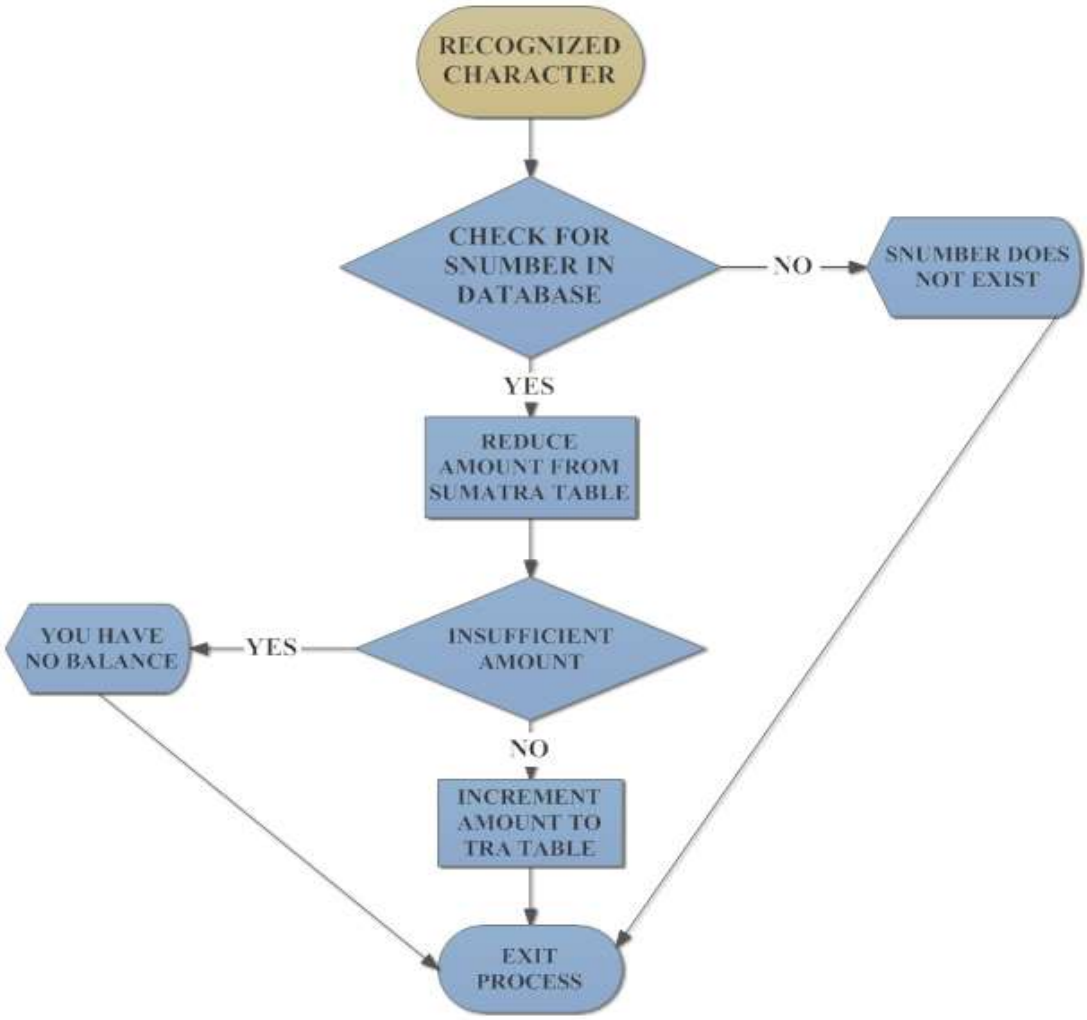


Figure 4. 10: Payment Process in Case of Insufficient Funds (Source: Own Processing)

The above figure 4.10 illustrates the payment process in case the mini bus pre-paid account had no amount to deduct from. Once the comparison occurs and the amount in the database is less than 500/- Tsh; which is the amount charged per trip, the system alerts the responsible person. Then the system terminates ready for the next recognized snumber, comparison for the payment process.

4.4 Evaluation of the Simulation Performance

This study has simulated the revenue collection process, with a focus on the SUMATRA mini bus numbers' recognition and revenue payment process. Two scenarios were examined, first when a recognized number did not exist in the database Appendix 6 and secondly when the mini bus account had insufficient balance for the money transfer to occur Appendix 5.

Both the scenarios are demonstrated as shown in figure 4.9 and 4.10 respectively. An appropriate alert message were generated and displayed for the responsible person to take an action when required to.

CHAPTER FIVE

CONCLUSION AND RECOMMENDATION

5.1 Introduction

This chapter concludes the research by giving a brief summary of the findings from the proceeding chapters, with a focus on the results obtained from the testing of the designed simulation. The chapter then, advances to show the limitations encountered and recommendations for the stake holders and a further study.

5.2 Conclusions

Number recognition and character recognition are growing technologies with a number of applications from the health sector, engineering, education to business. The potential to solve these problems has seen these technologies being applied in our daily lives, including a face recognition for security, object identification and number plate recognition for traffic purposes.

This study has employed number recognition technique and OCR to obtain SUMATRA numbers on the side of mini buses, at the Dodoma Municipal mini bus terminals at the Jamatini and Sabasaba for the automation of revenue collection. The developed simulation has the potential to solve revenue collection problems facing LGA's. Deployment of such processes at the desired terminals would increase the revenue collection and eliminate if not reduce, fraud activities which the current manual system is prone to.

5.3 Research Limitations

Several factors, can limit the simulation performance when use of the directly captured images from the mini bus include:

1 Variation of the character properties like font size and font face used in the creation of the Sumatra number.

2 Variation in the placement / printing of the number on the mini bus.

These limitations are due to the absence of a standard description of the SUMATRA number character properties. The orientation and positioning of the SUMATRA number on the side of the mini bus is also not uniform for all the mini buses in the municipal.

5.4 Recommendations

This research main objective was to simulate the automation of the revenue collection process at the Dodoma Municipal mini bus terminals of Jamatini and Sabasaba.

Due to limitations encountered, the research used own processed images and proposed a Microsoft publisher template, for the creation of the SUMATRA numbers that will be similar for all the mini busses, operating at the Municipal. The template will eliminate variation among SUMATRA numbers on mini busses. The research also recommends a standard or guide from SUMATRA for both number and position of the number, on the side of the mini bus. The guide will create conducive environment for the development and implementation of the simulated revenue collection process.

5.5 Further Research

Despite the accomplishment, the researcher recommends further improvements in the designed simulation, so as to make it robust. Further study should be done on the database system to be used, image acquisition techniques so as to obtain images

from moving vehicles. Other research area that is worth looking at is the integration of a messaging system to automatically inform the mini bus owner of insufficient balance in real time.

REFERENCES

Ayash, E. M. M. (2007), Research Methodologies in Computer Science and Information Systems, 1–4.

Badr, A., Abdel, M. M., Thabet, A. M., & Abdelsadek, A. M. (2011), Automatic number plate recognition system. Annals of the University of Craiova, Mathematics and Computer Science Series, 38(1), 62–71.

Burke, J. G., O'Campo, P., Peak, G. L., Gielen, A. C., McDonnell, K. A., & Trochim, W. M. K. (2005), An Introduction to Concept Mapping as a Participatory Public Health Research Method. Qual Health Res, 15(10), 1392–1410. https://doi.org/10.1177/1049732305278876

Choudhary, A. (2014), A Review of Various Character Segmentation Techniques, 4(6), 559–564.

Education Development Center. (2004), Key Informant Interviews. Data Collection Methods: Getting Down to Basics, 1–6. Retrieved from https://www. oasas.ny.gov/prevention/needs/documents/KeyInformantInterviews.pdf

Elio, R., Hoover, J., Nikolaidis, I., Salavatipour, M., Stewart, L., & Wong, K. (2011), About Computing Science Research Methodology, 9.

Fjeldstad, O.-H., Katera, L., & Ngalewa, E. (2005), Outsourcing Revenue Collection : Experiences from Local Government Authorities in Tanzania. Public Administration, (until).

Fjeldstad, O.-H., Katera, L., & Ngalewa, E. (2009), Outsourcing Revenue Experiences from Local Authorities in Tanzania Outsourcing Revenue Collection to Private Agents : Experiences from Local Authorities in Tanzania. Retrieved from https://www. africaportal. org/dspace/articles/ outsourcing-revenue-collection-private-agents-experiences-local-authorities-tanzania

GERRING, J. (2004), apsa. American Political Science Review, 98(2), 341–354. https://doi.org/10.1017/S0003055404001182

Government, L. (2013), The United Republic of Tanzania Priminister's Office , Final Report A Study on LGAs Own Source Revenue Collection.

goyal, A., & Bhatia, R. (2016), Automated Car Number Plate Detection System to detect far number plates. IOSR Journal of Computer Engineering, 18(4), 34–40. https://doi.org/10.9790/0661-1804033440

Hansen, J. (2002), A Matlab Project in Optical Character Recognition (OCR). DSP Lab, University of Rhode Island, 6.

Hevner, A., & Chatterjee, S. (2010), Design Research in Information Systems, 22, 9–23. https://doi.org/10.1007/978-1-4419-5653-8

Kopparapu, S. K. (2015), Mobile Phone Based Vehicle License Plate Recognition for Road Policing, 1–7. Retrieved from http://arxiv.org/abs/1504.01476

Kothari, C. R. (2014), No Title.

Kranthi, S., Pranathi, K., & Srisaila, A. (2012), Automatic number plate recognition. Indian Journal of Science and Technology, 2(3), 1–18. Retrieved from https://pdfs.semanticscholar.org/f31c/17e64e3e78f83ab28478e20196b56b84176a.pdf

Krishna, P. S. (2015), Automatic Number Plate Recognition by Using Matlab, 2(4), 1–7.

Kulkarni, A. H. (2014), Study of Different Electronic toll Collection Systems and Proposed toll Snapping and Processing System, 4(3), 1357–1360.

Kumar, & Phrommathed. (2005), Research Methodology.

Leadtools. (2017), OCR Confidence Reporting. Retrieved from https://www.leadtools.com/help/leadtools/v19/dh/to/leadtools.topics.forms.ocr~fo.topics.ocrconfidencereporting.html

Lukio, M. (2016), Outsourcing or Giving All Out ? Experience from Tanzania Local Government Revenue Collection, 6(3), 1–11. https://doi.org/ 10.5296/jpag.v6i3.8695

Ma, O.-C., Yang, J.-H., Manager, G., Choi, Y.-J., Jang, B.-S., Lee, J.-T., & Manager, S. (2009), Toll Collection System and Freeway Traffic Management System in Incheon Bridge project Kwang-Ayoul Moon, 1–17.

Mabhuye, P. P. (2013), The Performance of Outsourced Revenue Collection System in Local Government Authorities Case Study of Kasulu District.

Mccluskey. (2015), The role of ICT in delivering efficient revenue collection in developing countries: The Tanzanian experience, 5(5).

McLaren, T., & Buijs, P. (2011), A Design Science Approach for Developing Information Systems Research Instruments. 6th International Conference on Design Science Research in Information Systems and Technology, 1–10. Retrieved from http://www. rug.nl/staff/p. buijs/design_ science_ approach_for_developing_isr_instruments.pdf

Memon, F. A. (2012), Vehicle Number Recognition System for Automatic.

Mgonja, B. E. S., & Tundui, C. (2012), Institutional Impacts of the Local Government Reform Program on Good Local Governance in Tanzania. International Journal of Academic Research in Business and Social Sciences, 2(5), 206–222.

Mitra, D., & Banerjee, S. (2016), Automatic Number Plate Recognition System : A Histogram Based Approach, 11(1), 26–32. https://doi.org/10.9790/1676-11142632

Munuo, C. H. (2014), Vehicle Plate Number Detection and Recognition Using Improved Algorithm, 5(10), 31–41.

Munuo, C. H., & Kisangiri, M. (2014), Vehicle Number Plates Detection and Recognition using improved Algorithms : A Review with Tanzanian Case study, 3(5), 5828–5832.

Naik, S. S., Bidya, T., Singha, C., & Beham, A. R. (2017), Review on Different Techniques for Open Road Tolling System Using Pattern Recognition, 2(3), 311–314.

Ngowi, H. P. (2006), Public-Private Partnership (PPPs) in the Management of Municipalities in Tanzania – Issues and Lessons of Experience Mzumbe University. African Journal of Public Administration and Management, 17(2), 1–18.

Nur Aishah Rodziah Alkhushairi, S. (2012), CAR PLATE DETECTION SYSTEM.

Pandey, M., & Pandey, M. K. (2014), Automatic Vehical Registration Plate Recognition System Using Soft Computing Techniques. IJCER, 3(5), 242–247. Retrieved from http://ijcer.org/index.php/ojs/article/view/630

Patel, C., Shah, D., & Patel, A. (2013), Automatic Number Plate Recognition System (ANPR): A Survey. International Journal of Computer Applications, 69(9), 21–33. https://doi.org/10.5120/11871-7665

Patel, S. G., & Vallabhbhai, S. (2013), Vehicle License Plate recognition using Morphology and neural network, 2(1), 1–7.

Pervez, U. (2013), Automatic Number Plate Recognition System for, 69(December), 16–18.

Puranic, A. (2016), Vehicle Number Plate Recognition System : A Literature Review and Implementation using Template Matching, 134(1), 12–16.

Salunke, P., Malle, P., Datir, K., & Dukale, J. (2013), Automated Toll Collection System Using RFID. Journal of Computer Engineering, 9(2), 61–66.

Shemdoe, E. (2015), Factors Affecting Collection of Revenue in Muleba District Council in Kagera Region Tanzania: The Case Study of Muleba District Council in Kagera Region.

Siam, M. Z. (2014), Practical Design of an Automatic License Plate Recognition Using Image Processing Technique. International Journal of Engineering and Innovative Technology (IJEIT), 4(4), 170–179.

Simin, N., Choong, F., & Mei, C. (2013), Automatic Car-plate Detection and Recognition System, 113–114.

Tang, Y., Zhang, C., Gu, R., Li, P., & Yang, B. (2015), Vehicle Detection and Recognition for Intelligent Traffic Surveillance System. Multimedia Tools and Applications, (243), 1–16. https://doi.org/10.1007/s11042-015-2520-x

Tran, H.-N., & Nguyen, T. (2014), Supervised Learning Method for Mobile International Number Plate Recognition. Proceedings of the 16th International Conference on Information Integration and Web-Based Applications & Services, 212–216. https://doi.org/ 10.1145/ 2684200. 2684307

Tyson's, C., & Nelson, K. (2011), InfoSec Reading Room tu , A ho ll r igh ts. SANS Institute, (September), 21.

UCLA Center for Health Policy Research. (2016), Using Key Informants. Key Informants, 10.

Wen, Y., Lu, Y., Yan, J., & Zhou, Z. (2011), An Algorithm for License Plate Recognition Applied to Intelligent Transportation System An Algorithm for License Plate Recognition Applied to Intelligent Transportation System, (October). https://doi.org/10.1109/TITS.2011.2114346

Yin, B. R. K. (1994), Case Study Research., 1–5.

APPENDICES

Appendix 1: Interview Questions

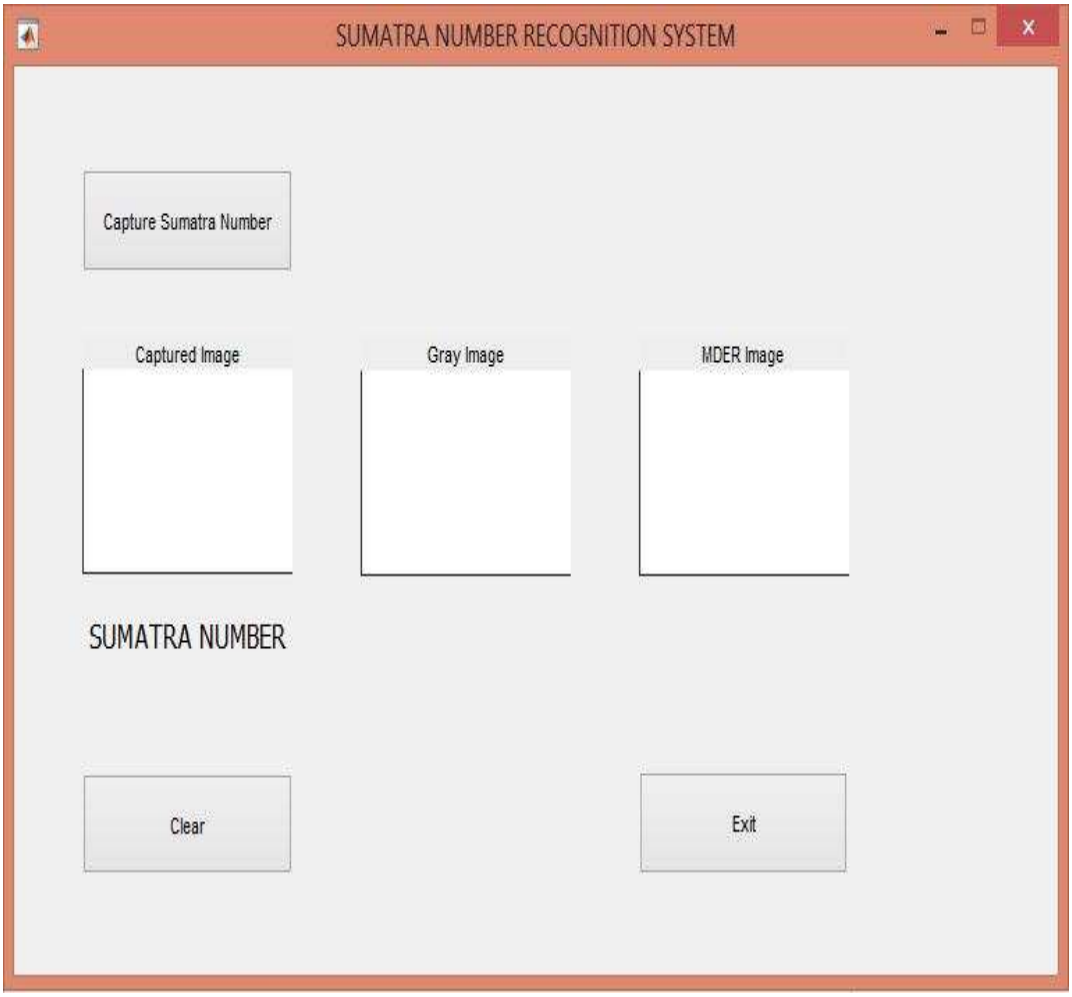
Revenue collection agent

1. What information do you need to issue the correct receipt?
2. What is the average time it takes to issue a receipt?
3. Are there fee avoiders?
    a. If yes. What do you do about it?
4. What are your working hours?
5. How many collectors are assigned in a single station?
6. During your absence, who collects the fees?
7. What sort of challenges do you face in your line of work?
8. What do you think could be the solution to the problems you face?

Mini bus Drivers

1. What is required from you by the collecting agent?
2. How much do you pay for a trip?
3. Do you pay for all trips?

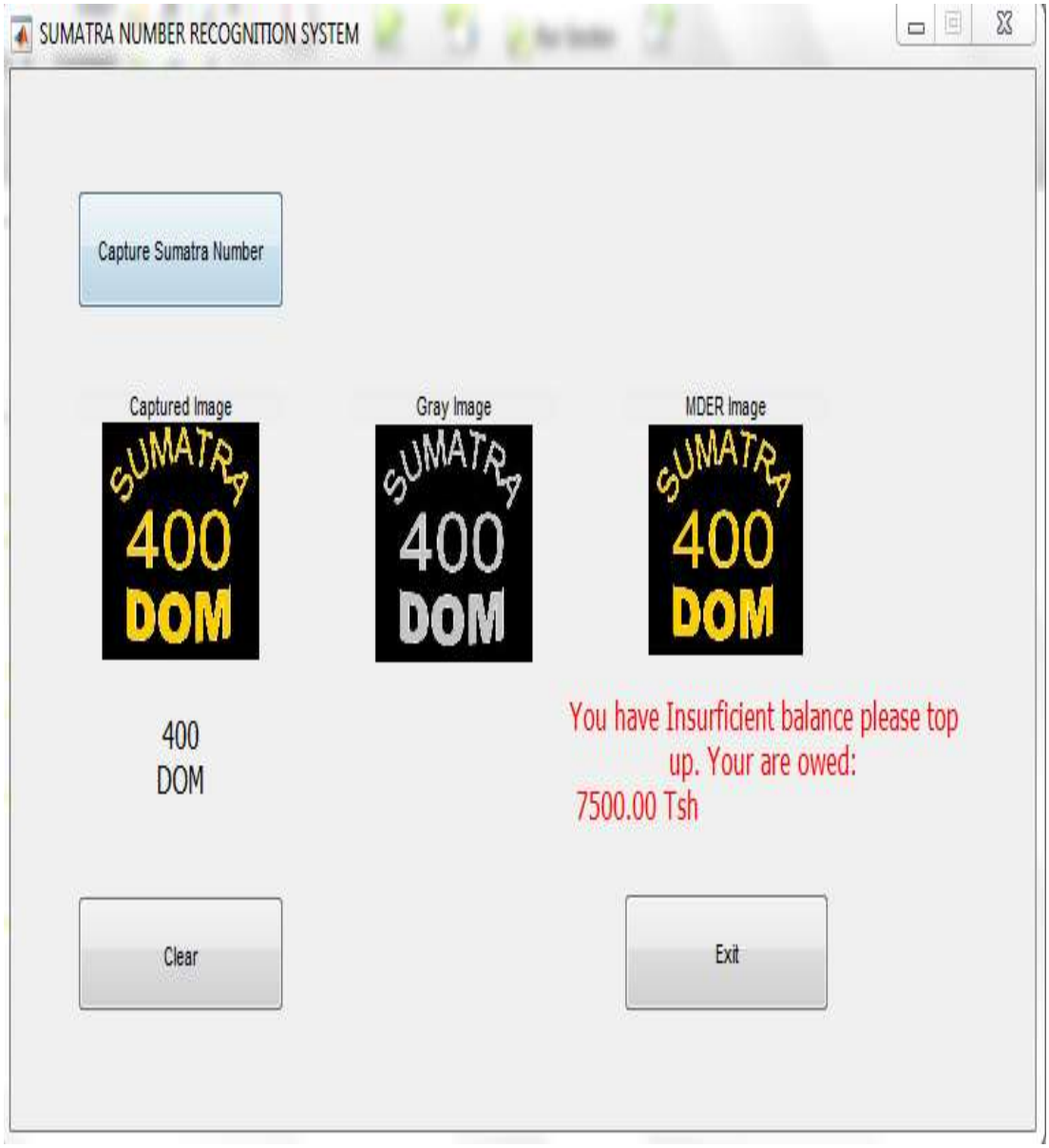
Appendix 2: Interface



Appendix 3: Account Top-up Interface

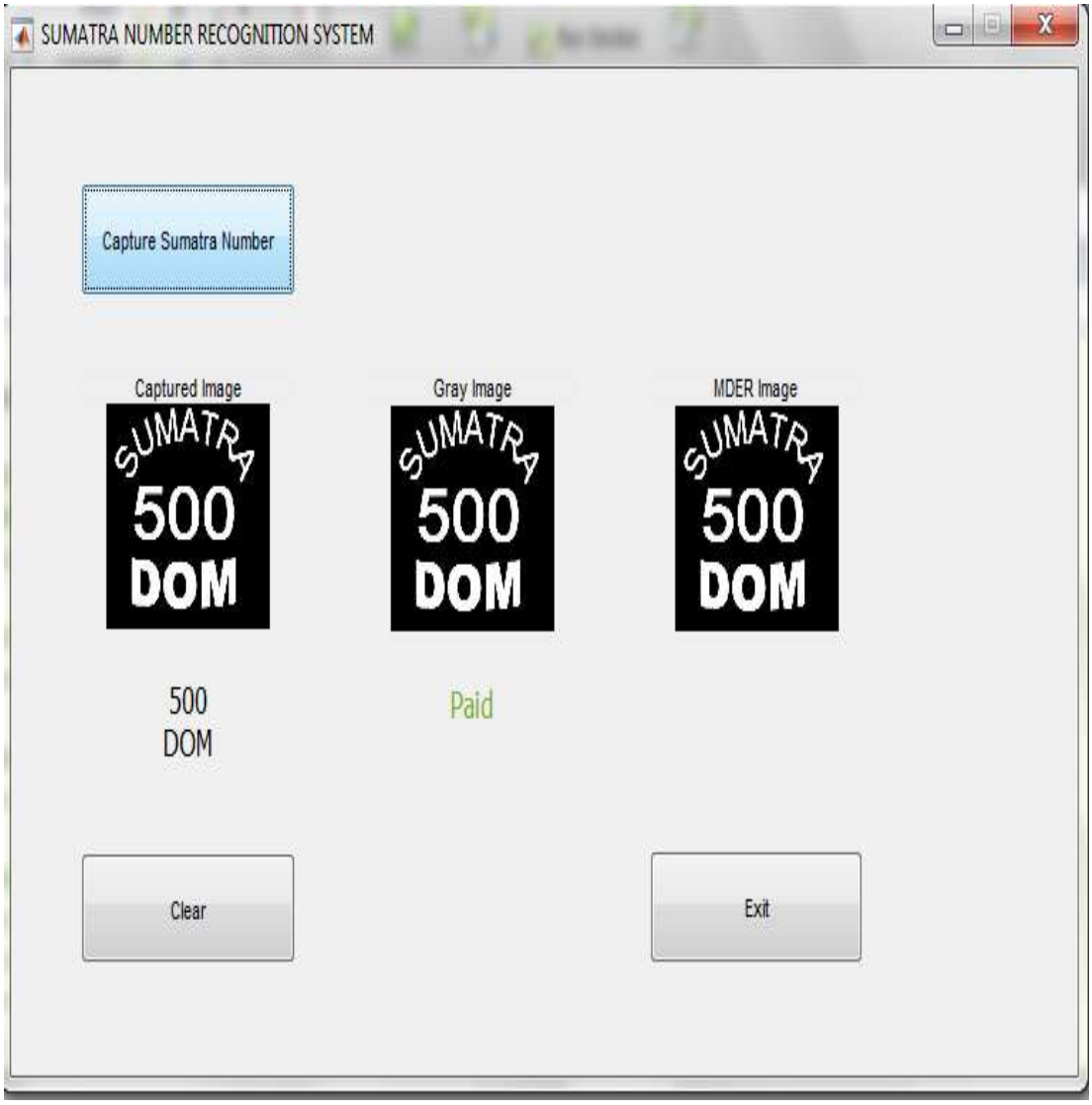
Appendix 4: EFD Receipt

Appendix 5: Insufficient Amount



Appendix 6: Not Registered Mini Bus

Appendix 7: Successfully Paid



Appendix 8: Account Top-up Codes

```
snum= char(get(handles.BOname,'string'));
conn=database('sumatra','admin','admin');
query=['select amount from sumatra where snumber in (',snum,')'];
curs=exec(conn,query);
curs=fetch(curs);
balance=curs.Data;
if cell2mat(balance) ~= 'No Data'
amount=str2double(get(handles.Amt,'string'));
adding=cell2mat(balance) + amount;
colname={'amount'};
whereclause=['where snumber in (',snum,')'];
update(conn, 'sumatra', colname, adding,whereclause);
 else
   disp('The sumatra number is not registered')
   end
set(handles.BOname,'string','');
set(handles.Amt,'string','')
```

Appendix 9: Number Recognition System Code

```
axes(handles.axes1);
smtr=imread(uigetfile('.jpg'));
imshow(smtr);
%% convert to gray scale
axes(handles.axes2);
smtrgry=rgb2gray(smtr);
imshow(smtrgry);
imshow(smtrgry);
%% detect MSER regions
axes(handles.axes3);
smtRegions=detectMSERFeatures(smtrgry,'ThresholdDelta',0.5);
smtRegionsPixels=vertcat(cell2mat(smtRegions.PixelList));
imshow(smtr); hold on;
plot(smtRegions,'showPixelList',true,'showEllipses',false);
%convert to binary mask
mserMask= false(size(smtr));
ind = sub2ind(size(mserMask),smtRegionsPixels(:,2),smtRegionsPixels(:,1));
mserMask(ind)=true;
ocrtxt=ocr(smtrgry);
%% single line display
recoText=ocrtxt.Words(2:3,1);
imshow(smtr);
set(handles.snumber,'string',recoText);
final=char(ocrtxt.Words(2,1));
```

```
conn=database('sumatra','admin','admin');
query = ['SELECT amount FROM sumatra where snumber in (',final,')'];
curs=exec(conn,query);
curs=fetch(curs);
jibu=curs.Data;
%% reduction
if cell2mat(jibu)~= 'No Data'
colname={'amount'};
reduction={(cell2mat(jibu))-500};
tname='sumatra';
whereclause=['where snumber = (',final,')'];
update(conn,tname ,colname,reduction,whereclause);
else
  alert=sprintf('Sumatra number is not registered');
    set(handles.alert,'string',alert);
return;
end
%% UPDATE TRA
squery=['select AmountPaid from TRA where snumber  in (',final,')'];
curs=exec(conn,squery);
curs=fetch(curs);
if (cell2mat(jibu))>=500
colname={'AmountPaid'};
pamont={(double(cell2mat(curs.Data)))+500};
tname='TRA';
```

```
whereclause=['where snumber=(',final,')'];
update(conn,tname,colname,pamont,whereclause);
cols={'NOTrips'};
ntrip={(cell2mat(pamont))/500};
update(conn,tname,cols,ntrip,whereclause);
status=('Paid');
set(handles.stats,'string',status);
else
    deni=abs(cell2mat(reduction));
 alert=sprintf('You have Insurficient balance please top up. Your are owed: \n %.2f
Tsh',deni);
set(handles.alert,'string',alert);
end
function clrBTN_Callback(~, ~, handles)
cla(handles.axes1);
cla(handles.axes2);
cla(handles.axes3);
set(handles.snumber,'string','');
set(handles.stats,'string','');
set(handles.alert,'string','');
 close all;
return;
```

Appendix 10: Comments Provided by External Examiner.

Chapter	Comments	Response	Page
1.	Use this chapter to tell exactly what you did, how you did and for which reasons to be able to respond to the different research questions.	Not applicable here. It's presented in the methodology chapter.	1
	Punctuation	Placed a full stop.	
	Re-phrase	The sentence is rephrased	2 & 3
	You have not talked anything on number recognition system previous works.	This is talked about in the literature review chapter	
	Wording and gamma errors	"m" , "the" & "to be" added and "a" removed	
	Need to modify specific objective No. one to capture NRT and OCR	Not applicable	4
	Can you simulate before designing?	Not applicable. Modified Kopparapu's framework used to simulate.	
	Align with reviewed specific objectives	Not applicable.	
2.	Can computers work tirelessly? and replace "pose as a run to" with "provides possible"	Removed tirelessly and replaced "pose as a run to" with "provides possible"	6
	Image source.	Changed image source to MUNUO (2014)	8
	Change words.	Removed the words "instead", "others" and re-phrased the last sentence.	
	Replace "end", add "that" and "ly"	Replaced "end " with "resolution" added "the" and "separately"	9
	Add "as shown in" and remove "below"	Added "as shown in" and removed "below"	11
	State the research gap as seen from the literature review	Not applicable.	12
3.	Drop or change to research methodology instead of literature review.	Sent to literature review and used some literature to justify the choice of methodology.	13, 14, 15 &16.
	Move closer to where mentioned in the text.	Image is in the right place.	17

	Remove " below" Drop turn to methodology.	Removed. Returned to literature review.	
	Punctuation and word error	Punctuated and corrected "ate to are"	18
	Elaborate more on the process How was it done What if a car is not from Dodoma and use facility?	Process elaborated. Explained how it works An error message will pop to indicate the number is not in database. And cars out of Dodoma can't operate in Dodoma.	19
	Elaborate how it works What is this?	Elaborated the how the process works. Removed the phrase.	20
4.	Does the chapter present?	Yes it does present findings	21
	Remove "below" punctuation	Removed. Sentence punctuated.	23
	Not referred in text body. There is no segmentation process described.	Referred in 4.3 Segmentation process is described.	24
	Remove "below" and "above"	Removed	25, 26, 27, 28, 29, 30, 33,34, 35
	Figure label Figure numbering	Re-labeled Corrected numbering.	29
	Why generation of images Refer to table number	Reason explained in the limitations of the existing numbers. Referred to table number	30
	Label equation using appropriate format	Labeled appropriately.	31
	Make sure figures are referred in the text body	Figures are referred.	32
	Check correctness of numbering of figures	Figures are numbered correctly	33
	Improve quality of figure	Quality improved.	35
5.	You did not seem to use actual images from the mini buses! I am not sure that this is true to its full extent.	Explained in the limitation on image quality and properties. Proposed standards for their generation/ creation. Re-phrased to the research context.	38