

WEB-BASED RECEIPT KEY INFORMATION EXTRACTION SYSTEM

BY

FONG JIA YEE

A REPORT

SUBMITTED TO

Universiti Tunku Abdul Rahman

in partial fulfillment of the requirements

for the degree of

**BACHELOR OF INFORMATION SYSTEMS (HONOURS) BUSINESS INFORMATION
SYSTEMS**

Faculty of Information and Communication Technology

(Kampar Campus)

JUNE 2022

REPORT STATUS DECLARATION FORM

Title: WEB-BASED RECEIPT KEY INFORMATION EXTRACTION SYSTEM

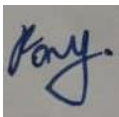
Academic Session: JUNE 2022

I FONG JIA YEE
(CAPITAL LETTER)

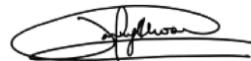
declare that I allow this Final Year Project Report to be kept in
Universiti Tunku Abdul Rahman Library subject to the regulations as follows:

1. The dissertation is a property of the Library.
2. The Library is allowed to make copies of this dissertation for academic purposes.

Verified by,



(Author's signature)



(Supervisor's signature)

Address:

LC 740, KAMPUNG BARU,

31700 MALIM NAWAR,

PERAK.

Ts Dr Tan Hung Khoon

Supervisor's name

Date: 5 SEPTEMBER 2022

Date: 17/9/2022

Universiti Tunku Abdul Rahman			
Form Title : Sample of Submission Sheet for FYP/Dissertation/Thesis			
Form Number: FM-IAD-004	Rev No.: 0	Effective Date: 21 JUNE 2011	Page No.: 1 of 1

FACULTY OF INFORMATION AND COMMUNICATION TECHNOLOGY

UNIVERSITI TUNKU ABDUL RAHMAN

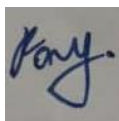
Date: 5 SEPTEMBER 2022

SUBMISSION OF FINAL YEAR PROJECT /DISSERTATION/THESIS

It is hereby certified that **FONG JIA YEE** (ID No: **19ACB04133**) has completed this final year project/ dissertation/ thesis* entitled “**WEB-BASED RECEIPT KEY INFORMATION EXTRACTION SYSTEM**” under the supervision of **Ts Dr Tan Hung Khoon** (Supervisor) from the Department of **Computer Science**, Faculty of **Information and Communication Technology**, and **Pn Noraini Binti Ibrahim** (Co-Supervisor) from the Department of **Digital Economy Technology**, Faculty of **Information and Communication Technology**.

I understand that University will upload softcopy of my final year project / dissertation/ thesis* in pdf format into UTAR Institutional Repository, which may be made accessible to UTAR community and public.

Yours truly,



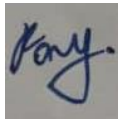
(*Fong Jia Yee*)

*Delete whichever not applicable

DECLARATION OF ORIGINALITY

I declare that this report entitled “**WEB-BASED RECEIPT KEY INFORMATION EXTRACTION SYSTEM**” is my own work except as cited in the references. The report has not been accepted for any degree and is not being submitted concurrently in candidature for any degree or other award.

Signature :



Name : FONG JIA YEE

Date : 13/6/2022

ACKNOWLEDGEMENTS

I would like to express my sincere thanks and appreciation to my supervisor, Ts Dr Tan Hung Khoon and my moderator, Puan Noraini Binti Ibrahim as they have given me this bright opportunity to engage in this joint project. Especially to Dr Tan, he gives me a lot of support and guidance throughout this project. He has been very helpful and insightful. A million thanks to you.

To my supportive partner of this project, Tan Ann Gee, for her patience, unconditional support and for standing by my side during hard times. Finally, I must say thanks to my parents and my family for their love, support, and continuous encouragement throughout the course.

ABSTRACT

Due to public financial awareness, it is significant for public to manage receipts properly, especially to extract key information from the receipts and store them properly. However, it is difficult for them to deal with numerous receipts due to some inevitable reasons. Therefore, the receipt processing should be automated. In this project, the main objective is to develop a receipt key information extraction system. Since this is a joint project performed by Fong Jia Yee and Tan Ann Gee, the task allocation has been done in which I am responsible in integrating three models to perform key information extraction and developing web-based receipt key information extraction system whereas Ann Gee is responsible in developing mobile-based receipt key information extraction system. The task to be completed together is to train the selected models. In the end, my developed system will be able to upload multiple receipts to be processed and store the records for review purpose whereas the developed system of Ann Gee will be able to upload one receipt and obtain the extracted key information from the receipt.

TABLE OF CONTENTS

TITLE PAGE	i
REPORT STATUS DECLARATION FORM	ii
FYP THESIS SUBMISSION FORM	iii
DECLARATION OF ORIGINALITY	iv
ACKNOWLEDGEMENTS	v
ABSTRACT	vi
TABLE OF CONTENTS	vii
LIST OF FIGURES	x
LIST OF TABLES	xiv
LIST OF ABBREVIATIONS	xv
CHAPTER 1 INTRODUCTION	1
1.1 Key Information Extraction of Receipts	1
1.2 Automated Receipt Processing	4
1.3 Objectives	5
1.4 Project Scope and Direction	5
1.5 Contributions	6
1.6 Report Organization	6
CHAPTER 2 LITERATURE REVIEW	8
2.1 Text Localization	8
2.1.1 Connectionist Text Proposal Network (CTPN)	8
2.2 Optical Character Recognition (OCR)	10
2.3 Keyword Information Extraction	11
2.3.1 Convolutional Universal Text Information Extractor (CUTIE)	11
2.3.2 Bidirectional Encoder Representations from Transformers (BERT)	14
2.3.3 Character-Aware Neural Language Model	17
2.4 Limitations of Previous Studies	18

CHAPTER 3 SYSTEM DESIGN	20
3.1 System Overview	20
3.2 System Architecture Diagram	20
3.2.1 Keyword Information Extraction Module	20
3.2.2 User Management Module	24
3.3 System Block Diagram	26
3.3.1 User Management Module	26
3.4 Use Case Diagram for User Management Module	29
3.4.1 Use Case Diagram for Project I	29
3.4.2 Use Case Diagram for Project II	30
3.5 Activity Diagram for User Management Module	32
3.5.1 Activity Diagram for Project I	32
3.5.2 Activity Diagram for Project II	34
3.6 Database Design	44
3.7 User Interface Design	46
CHAPTER 4 SYSTEM IMPLEMENTATION	52
4.1 Hardware Setup	52
4.2 Software Setup	52
4.2.1 Visual Studio Code with WSL	52
4.2.2 Python	53
4.2.3 EasyOCR	53
4.2.4 Flask	53
4.2.5 MySQL Server	53
4.2.6 Planning and Design Tools	53
4.3 Models Training	54
4.4 System Operation (with Screenshot)	55
4.5 Implementation Issues and Challenges	81
4.6 Concluding Remark	81
CHAPTER 5 SYSTEM EVALUATION AND DISCUSSION	82
5.1 System Testing and Performance Metrics	82
5.2 Testing Setup and Result	82

5.3	Project Challenges	96
5.4	Objectives Evaluation	96
5.5	Concluding Remark	97
CHAPTER 6 CONCLUSION AND RECOMMENDATION		98
6.1	Conclusion	98
6.2	Recommendation	98
REFERENCES		99
WEEKLY LOG		101
POSTER		114
PLAGIARISM CHECK RESULT		115
FYP2 CHECKLIST		121

LIST OF FIGURES

Figure Number	Title	Page
Figure 1.1.1	Key information of a receipt	2
Figure 1.1.2	Documents with different types and layouts	3
Figure 1.1.3	Handwritten receipt with normal angle	3
Figure 1.1.4	Computer-generated receipt with normal angle	3
Figure 1.1.5	Computer-generated receipt with abnormal angle	3
Figure 1.1.6	Date and its value are arranged horizontally	4
Figure 1.1.7	Date and its value are arranged vertically	4
Figure 2.1.1.1	Connectionist Text Proposal Network (CTPN) architecture	8
Figure 2.3.1.1	Equations to calculate the mapping position of texts in the grid	11
Figure 2.3.1.2	Equations to augment gridded texts to shape with different rows and columns	11
Figure 2.3.1.3	High-Resolution Network (HRNet) architecture	12
Figure 2.3.1.4	Structure of CUTIE-B model	12
Figure 2.3.1.5	Flow of tokens applying CUTIE-B	13
Figure 2.3.2.1	Transformer encoder for Masked Language Model (MLM)	15
Figure 2.3.2.2	Transformer encoder for Next Sentence Prediction (NSP)	16
Figure 2.3.3.1	Character-Aware Neural Language model architecture	18
Figure 3.2.1.1	Connectionist Text Proposal Network (CTPN) architecture	20
Figure 3.2.1.2	Crop pre-processing flow	21
Figure 3.2.1.3	EasyOCR architecture	22
Figure 3.2.1.4	Character-Aware Neural Language model architecture	23
Figure 3.2.2.1	Web-based receipt key information extraction system architecture	24
Figure 3.3.1.1	Web-based receipt key information extraction system block diagram	26
Figure 3.4.1.1	Use case diagram for Project I	29
Figure 3.4.2.1	Use case diagram for Project II	30

Figure 3.5.1.1	“Upload files” activity diagram for Project I	32
Figure 3.5.1.2	“View processed images” activity diagram for Project I	33
Figure 3.5.1.3	“Download text files” activity diagram for Project I	33
Figure 3.5.2.1	“Register” activity diagram for Project II	34
Figure 3.5.2.2	“Login” activity diagram for Project II	35
Figure 3.5.2.3	“Logout” activity diagram for Project II	36
Figure 3.5.2.4	“Forgot password” activity diagram for Project II	36
Figure 3.5.2.5	“Upload files” activity diagram for Project II	37
Figure 3.5.2.6	“View uploaded images” activity diagram for Project II	38
Figure 3.5.2.7	“Download text files” activity diagram for Project II	39
Figure 3.5.2.8	“View uploaded records” activity diagram for Project II	40
Figure 3.5.2.9	“View record details” activity diagram for Project II	41
Figure 3.5.2.10	“View deleted records” activity diagram for Project II	42
Figure 3.5.2.11	“Delete records” activity diagram for Project II	43
Figure 3.5.2.12	“Restore records” activity diagram for Project II	44
Figure 3.7.1	“login” user interface design.	46
Figure 3.7.2	“register” user interface design.	47
Figure 3.7.3	“validate email” user interface design (first stage to reset password).	48
Figure 3.7.4	“reset password” user interface design (second stage to reset password).	49
Figure 3.7.5	“upload” user interface design.	49
Figure 3.7.6	“uploaded records” user interface design.	50
Figure 3.7.7	“deleted records” user interface design.	50
Figure 3.7.8	“record details” user interface design.	51
Figure 4.3.1	Training and evaluation outputs for both Connectionist Text Proposal Network (CTPN) and Character-Aware Neural Language models.	54
Figure 4.4.1	“login” interface	55
Figure 4.4.2	“register” interface	55
Figure 4.4.3	Invalid email format message	56
Figure 4.4.4	Password and confirm password do not match message	56
Figure 4.4.5	Successfully registered message	57

Figure 4.4.6	Login with invalid email or password	57
Figure 4.4.7	Invalid email or password message	58
Figure 4.4.8	Login with valid email and password	58
Figure 4.4.9	“upload” interface or home page	59
Figure 4.4.10	“uploaded records” page with no record	59
Figure 4.4.11	“deleted records” page with no record	60
Figure 4.4.12	At least one file required to be uploaded message	60
Figure 4.4.13	Files to be uploaded	61
Figure 4.4.14	Progress bar when waiting for files to be processed	61
Figure 4.4.15	Results when files have been processed	62
Figure 4.4.16	Expanded results	62
Figure 4.4.17	“uploaded records” page with records	63
Figure 4.4.18	“uploaded records” page with records (cont.)	63
Figure 4.4.19	“uploaded records” page with records (cont.)	64
Figure 4.4.20	Uploaded image in new tab	64
Figure 4.4.21	Extracted key information stored in text file	65
Figure 4.4.22	Search by company details	65
Figure 4.4.23	Search by address details	66
Figure 4.4.24	Search by date details	66
Figure 4.4.25	Search by total details	67
Figure 4.4.26	“record details” page	67
Figure 4.4.27	“uploaded records” page after record ID 1 has been deleted	68
Figure 4.4.28	“deleted records” page after record ID 1 has been deleted	68
Figure 4.4.29	“uploaded records” page after deleting other records	69
Figure 4.4.30	“deleted records” page after deleting other records	69
Figure 4.4.31	Upload the records have been deleted from “record details” page	70
Figure 4.4.32	“uploaded records” page after uploading the records	70
Figure 4.4.33	Record has not been restored message	71
Figure 4.4.34	“uploaded records” page after deleting two records	71
Figure 4.4.35	“deleted records” page after deleting two records	72
Figure 4.4.36	Successfully restored and record has not been restored messages	72

Figure 4.4.37	“uploaded records” page after two records have been restored	73
Figure 4.4.38	Successfully deleted message	73
Figure 4.4.39	“validate email” interface (first stage to reset password)	74
Figure 4.4.40	Reset password with invalid email	74
Figure 4.4.41	Not an existing email message	75
Figure 4.4.42	Reset password with valid email	75
Figure 4.4.43	“reset password” interface (second stage to reset password)	76
Figure 4.4.44	Password and confirm password do not match message	76
Figure 4.4.45	Reset password with matched password and confirm password	77
Figure 4.4.46	Successfully reset password message	77
Figure 4.4.47	Register another new user account	78
Figure 4.4.48	Expanded results when file has been processed	78
Figure 4.4.49	“uploaded records” page with records	79
Figure 4.4.50	Expanded results after uploading duplicated image	79
Figure 4.4.51	Text files for uploaded image and recorded image	80
Figure 4.4.52	“user” table in “kie” database	80
Figure 4.4.53	“record” table in “kie” database	80

LIST OF TABLES

Table Number	Title	Page
Table 3.6.1	Database structure	44
Table 4.1.1	Laptop specification	52
Table 5.2.1	Register testing	82
Table 5.2.2	Login testing	83
Table 5.2.3	Logout testing	84
Table 5.2.4	Reset password testing	84
Table 5.2.5	Upload files testing	85
Table 5.2.6	View uploaded images testing	89
Table 5.2.7	Download text files testing	89
Table 5.2.8	View uploaded records testing	89
Table 5.2.9	View record details testing	91
Table 5.2.10	View deleted records testing	91
Table 5.2.11	Delete records testing	93
Table 5.2.12	Restore records testing	94
Table 5.4.1	Objectives evaluation	96

LIST OF ABBREVIATIONS

<i>OCR</i>	Optical Character Recognition
<i>CTPN</i>	Connectionist Text Proposal Network
<i>SMEs</i>	Small and medium-sized enterprises
<i>CNN</i>	Convolutional Neural Network
<i>LSTM</i>	Long Short-Term Memory
<i>BLSTM</i>	Bidirectional LSTM
<i>RNN</i>	Recurrent Neural Network
<i>CUTIE</i>	Convolutional Universal Text Information Extractor
<i>BERT</i>	Bidirectional Encoder Representations from Transformers
<i>ASPP</i>	Atrous Spatial Pyramid Pooling
<i>HRNet</i>	High-Resolution Network
<i>NLP</i>	Natural Language Processing
<i>MLM</i>	Masked Language Model
<i>NSP</i>	Next Sentence Prediction
<i>MSS</i>	Max Sum Similarity
<i>MMR</i>	Maximal Marginal Relevance
<i>API</i>	Application Programming Interface

Chapter 1

Introduction

People typically receive slips of paper when they buy items in a physical store, they are known as receipts. A receipt is a record required to be provided to the buyer as a reference to prove the existence of the transaction between both the seller and buyer parties and avoid them getting into unnecessary disputes when the buyer is willing to refund, return an item, or claim a warranty without a receipt. Keeping receipts has become a common practice of majority people regardless of business receipts or personal receipts as they assist people in monitoring and staying updated the monetary transactions. If there are numerous receipts required to be managed, even if a lot of manpower is provided, the process of manually extracting details from receipts is time-consuming. Besides, human error is unavoidable as people are more likely to make mistakes when extracting details from receipts compared with the system. The main challenge is due to the faded thermal receipt nowadays as characters on the receipt will fade when the receipt exposed for an extended period of time. Therefore, in order to minimize such problems, the automated processing of receipts should be required where it is a process of extracting key information from receipts in a smooth manner. In this project, the key information of the receipt is reference number, transaction date and time, item or service name and its corresponding price, total amount paid, etc.

1.1 Key Information Extraction of Receipts

Key information extraction is a process of extracting texts from a scanned document where extracted texts have both spatial and semantic features. The input of key information extraction is a scanned document, and its output is extracted texts from the document.

**Berghotel
Grosse Scheidegg**
3818 Grindelwald
Familie R.Müller

Rech.Nr. 4572 30.07.2007/13:29:17
Bar Tisch 7/01

2xLatte Macchiato	à	4.50	CHF	9.00
1xGloki	à	5.00	CHF	5.00
1xSchweinschnitzel	à	22.00	CHF	22.00
1xChässpätzli	à	18.50	CHF	18.50

Total : CHF **54.50**

Incl. 7.6% MwSt 54.50 CHF: 3.85

Entspricht in Euro 36.33 EUR
Es bediente Sie: Ursula

MwSt Nr.: 430 234
Tel.: 033 853 67 16
Fax.: 033 853 67 19
E-mail: grossescheidegg@bluewin.ch

Figure 1.1.1: Key information of a receipt.

Based on the receipt provided in figure 1.1.1, the value “Berghotel Grosse Scheidegg” will be extracted and mapped to the key “Hotel name”. Other key information examples such as the value “30.07.2007” and its key “Date of the transaction”, the value “13:29:17” and its key “Time of the transaction”, etc.

CHAPTER 1

Some key information extraction problems are discovered such as different types of document structures, a large number of potentially related words, the order or word-to-word distance of texts in the line-base-aligned long paragraph tends to differ significantly due to the layout variant.



a. Cash receipt.



b. Bank slip.



c. Bus ticket.

Figure 1.1.2: Documents with different types and layouts.

Receipts may come in different points. However, processing receipts is by no means an easy task. Receipts may come from different layouts. For each layout, some receipts may be handwritten whereas others are computer generated. For each case, some receipts are captured with the phone whereas some are from the phone screen, some are from a normal angle whereas some are from an abnormal angle.

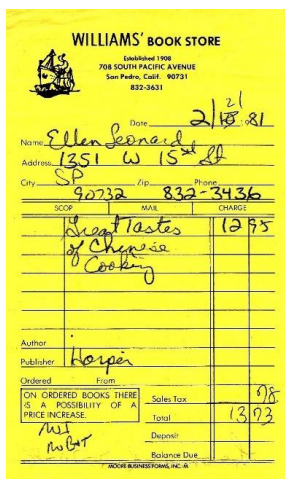


Figure 1.1.3: Handwritten receipt with normal angle.



Figure 1.1.4: Computer-generated receipt with normal angle.



Figure 1.1.5: Computer-generated receipt with abnormal angle.

CHAPTER 1

Optical Character Recognition (OCR) tool is not able to distinguish types of scanned documents. It blindly converts texts containing images into characters regardless of tables, fields, and alignment of them [1]. In other words, it is able to detect and recognize texts from the scanned document, however the extracted texts are meaningless due to some reasons. For example, key information can be expressed by different terms such as reference number and order number. Some texts are difficult to be categorized such as the name of the company. Besides, keys and values may have no association between them because they are arranged in different ways such as horizontally in which the value appears to the right of its key or vertically in which the value appears to the bottom of its key. These machine-readable data cannot be further processed since the machine does not understand the meaning. Therefore, the extracted texts which are the OCR output should be converted into machine-actionable data.



Figure 1.1.6: Date and its value are arranged horizontally.



Figure 1.1.7: Date and its value are arranged vertically.

1.2 Automated Receipt Processing

There are numerous existing models. After selecting the suitable existing text localization, OCR, and keyword information extraction models, a web-based receipt key information extraction system should be developed in this project to allow the keyword information extraction from a provided receipt. Since each model has its own task to be performed, they should be connected properly in a way that the output of the text localization model (first

CHAPTER 1

model) is able to be accepted by the OCR model (second model) and the output of the OCR model is able to be accepted by the keyword information extraction models (third model). In the end, the input of the connected models should be an image which is a receipt, and the final output should be the keyword information extracted from the receipt in the format of text file. Since the models work well, the models should be implemented when developing the web-based receipt key information extraction system.

1.3 Objectives

The main objective of this project is to develop a web-based receipt key information extraction system. Before developing the web-based receipt key information extraction system, three tasks should be performed:

- i. Text localization – To receive an input image which is a receipt and generate the output which are the coordinates of the detected texts in the format of text file.
- ii. Optical Character Recognition (OCR) – To generate the output which are the extracted texts from the receipt based on the output provided by the text localization model.
- iii. Keyword information extraction – To receive the output from text localization and OCR models and generate the output which is the keyword information extracted from the receipt.

After completing these tasks with the assistance of the existing models, they should be deployed to be used in the development of the web-based receipt key information extraction system. The system should be able to handle the following tasks:

- i. Upload multiple receipts – To develop a web-based system which is able to upload multiple receipts for keyword information extraction from them simultaneously.
- ii. Store keyword information extraction records – To develop a web-based system which is able to store keyword information extraction records for review purpose.

1.4 Project Scope and Direction

This is a joint project performed by two students, namely Fong Jia Yee and Tan Ann Gee. Our goal is to develop a receipt key information extraction system. Nowadays, web and mobile applications are popular. People have their own preferences to select a platform to extract the keyword information from receipts due to different conditions. Thus, in this project, I am required to deal with the development of web-based receipt key information extraction system

CHAPTER 1

whereas Ann Gee is responsible in managing the development of mobile application receipt key information extraction system.

Connectionist Text Proposal Network (CTPN) model for text localization and Character-Aware Neural Language model for keyword information extraction are selected to be trained in order to generate the trained model to be used in the development of receipt key information extraction system. Both of us are responsible in managing the training of these models and the dataset used is from ICDAR 2019 Competition.

The receipts which are able to be handled by these models should be large in image size and the language of the receipt is limited to English as the dataset used to train the models are English receipts. The models are not required to retrain as the receipt language used in Malaysia is same with the receipt language in the provided dataset.

After completing the training, I am required to connect CTPN, EasyOCR, and Character-Aware Neural Language models to get the final output which is the keyword information extracted from the input receipt.

1.5 Contributions

The developed system is able to help the automation and digitalization of businesses, especially small and medium-sized enterprises (SMEs). When receipts are uploaded to the system, the problems created by faded thermal receipts can be eliminated. This is because receipts should be in image file format before scanning it. In this moment, softcopies of receipts in image file format have been collected. Since the system can process multiple images at once, it is more timesaving, efficient, and effective compared with processing multiple images manually. With the aid of the system, majority errors can be removed as its output accuracy is high. In this moment, softcopies of extracted texts from receipts in text file format have been collected if they are downloaded.

1.6 Report Organization

This project report has been organized into 6 chapters and the sequence is Chapter 1 Introduction, Chapter 2 Literature Review, Chapter 3 System Design, Chapter 4 System Implementation, System 5 System Evaluation and Discussion, and Chapter 6 Conclusion and Recommendation.

Chapter 1 Introduction introduces key information extraction of receipts, automated receipt processing, objectives, project scope and direction, contributions, and report organization.

CHAPTER 1

Chapter 2 Literature Review includes the review of text localization, OCR, and keyword information extraction. Some models which are able to handle text localization and keyword information extraction have been analyzed. In the end, limitations of previous studies have been stated.

Chapter 3 System Design describes the system overview and shows the system design of the project by system architecture diagram, system block diagram, use case diagram, activity diagram, database design, and user interface design.

Chapter 4 System Implementation describes the hardware setup and software setup for this project. Besides, it has described models training, system operations, implementation issues and challenges, and concluding remark for this chapter.

Chapter 5 System Evaluation and Discussion consists of system testing and performance metrics, testing setup and result, project challenges, objectives evaluation, and concluding remark for this chapter.

Chapter 6 Conclusion and Recommendation includes conclusion and recommendation after this project has been completed.

Chapter 2

Literature Review

2.1 Text Localization

2.1.1 Connectionist Text Proposal Network (CTPN)

Connectionist Text Proposal Network (CTPN) is a model proposed by combining the advantages of Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM). In order to make the output of text localization more accurate and reliable, the tasks involved are detecting text in fine-scale proposals, recurrent connectionist text proposals, and side-refinement [2].

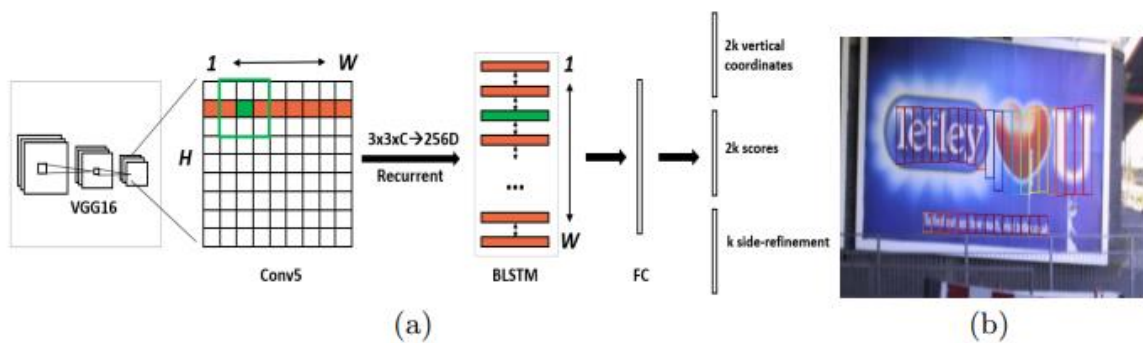


Figure 2.1.1.1: Connectionist Text Proposal Network (CTPN) architecture [2].

Based on (a) in figure 2.1.1.1, a 3 x 3 spatial window is slid through the last convolutional maps (Conv5) of the VGG16 model. The consecutive windows in each row are recurrently connected by bidirectional LSTM (BLSTM). 3 x 3 x C which is the convolutional feature of each window is the input of 256D BLSTM in which 128D for each LSTM. The Recurrent Neural Network (RNN) layer is connected to 512D fully connected layer. It is followed by the output layer which is used to predict the text or non-text scores, y-axis coordinates and side-refinement offsets of k anchors.

(b) in figure 2.1.1.1 is the CTPN outputs consecutive fixed-width fine-scale text proposals.

i. Detecting Text in Fine-scale Proposals

Anchor regression mechanism is adopted to detect multi-scale objects with a single-scale window in which the single-scale window can predict objects with a wide range of scales and aspect ratios by employing a number of flexible anchors. However, there is a major difference

between detecting generic objects and detecting texts in which the generic object is allowed to be inferred from even a minor portion of it due to its clearly defined closed boundary and center whereas the text has no clearly defined closed boundary as it may consist of multi-level components which are difficult to differentiate from one another. Detecting the text as a single object makes the text detection accuracy unreliable as it is defined at the word or text line level. It is difficult to determine where a word begins and ends as each character in a word is isolated. Therefore, an assumption is provided in which the prediction of the vertical location of each proposal is more accurate compared with the prediction of its horizontal location. Besides, the detection of a general fixed-width text proposal is more reliable than the detection of a single character as it can be mistaken for a portion of a character or numerous characters easily. Thus, fine-scale text proposal is designed. The standard defines a text proposal in the input image as having a fixed width of 16 pixels. The k anchors have the same horizontal location and a fixed width of 16 pixels whereas their respective vertical locations vary by k different heights. In the end, each predicted text proposal in the input image has a bounding box with size of $h \times 16$.

ii. Recurrent Connectionist Text Proposals

It lacks robustness when considering each text proposal separately as the text is strongly sequential in nature in which the sequential context information is significant for making reliable decision to recognize the task on word image cropped. Thus, hidden layers of RNN are used to encode the sequential context information directly in the convolutional layer to build a seamless in-network connection of the fine-scale text proposals. Besides, bidirectional LSTM is used to further extend the RNN layers to encode the recurrent context information in both directions. Thus, the whole image width can be covered by the connectionist receipt field.

Below is the explanation of using BLSTM instead of standard LSTM.

“My phone is broken. I am planning to ___ a new phone.”

By using BLSTM, the possibility to fill in the blank by the word “buy” is higher as it is able to capture the information before where the phone is broken and after the blank where a new phone is mentioned.

iii. Side-refinement

Accurate localization in the vertical direction can be predicted using the fine-scale detection and RNN connection. The image is divided into a series of equal 16-pixel width proposals in

the horizontal direction. Inaccurate localization may occur when the text proposals in both horizontal sides are not completely covered by a ground truth text line area, or when some side proposals are disposed of.

Thus, side-refinement is proposed to estimate the offset of each proposal in both left and right horizontal sides. They are used to refine the final text line bounding box.

2.2 Optical Character Recognition (OCR)

Optical Character Recognition (OCR) is a conversion of texts containing images into machine-readable texts. An OCR algorithm consists of three basic steps, which are preprocessing the input image, followed by text detection, and ended by text recognition [3]. There are two significant parts of OCR, namely preprocessing and segmentation.

After scanning or taking photo of the selected document, preprocessing is taken place. Some preprocessing techniques are introduced such as binarization, skew correction, noise removal, and thinning and skeletonization [4]. Binarization is used to convert a colored image into a binary image to minimize the ambiguity when detecting texts. Skew correction is used to tilt an image to be oriented at a specific angle with horizontal to ease information extraction. Noise removal is used to remove small dots with higher intensity to smooth the image. Thinning and skeletonization is used to uniform the stroke width.

Segmentation is carried out after performing image preprocessing to break the entire processed image into sub parts to further process them. There are three levels of segmentation, namely line level segmentation, word level segmentation, and character level segmentation [5]. Before going into segmentation, two histogram projection methods which are horizontal histogram projection and vertical histogram projection should be introduced. Foreground pixels mean pixels which represent the meaningful information whereas background pixels are the opposite of foreground pixels. For horizontal histogram projection, foreground pixels number along image rows is counted. For vertical histogram projection, foreground pixels number along image columns is counted. Line level segmentation is used to segment the image into lines by horizontal histogram projection in which the text line has high foreground pixels number whereas the gap between lines has high background pixels numbers. High background pixels number represents low peak in the histogram and columns with this characteristic are selected as segmenting lines to separate words. Character level segmentation is used to segment the image into individual characters by vertical histogram

projection. It is more complicated compared with line and word level segmentations due to ligatures in cursive handwriting.

2.3 Keyword Information Extraction

2.3.1 Convolutional Universal Text Information Extractor (CUTIE)

A CNN-based model, namely Convolutional Universal Text Information Extractor (CUTIE) has been proposed to solve key information extraction challenges [6]. Its input is gridded texts where they are generated by subjecting the input retained from the scanned image to grid positioning mapping method. Texts feature representation is encoded through a word embedding layer. It is better than Bidirectional Encoder Representations from Transformers (BERT) model because it does not require pretraining on a large text data collection.

Each word is tokenized, and each token is classified to its correct class using a predefined dictionary. To generate gridded texts, grid positional mapping is used. It aims to map texts of the scanned image to the target grid where the mapped grid retains the original spatial relationship among texts. Below are two questions required to calculate the mapping position of texts in the grid.

$$c_x^i = c_{gm} \frac{x_{left} + \frac{(x_{right} - x_{left})}{2}}{w}$$

$$r_y^i = r_{gm} \frac{y_{top} + \frac{(y_{bottom} - y_{top})}{2}}{h}$$

Figure 2.3.1.1: Equations to calculate the mapping position of texts in the grid [6].

To improve the ability of CUTIE to better deal with documents with different layouts, two other equations are used to augment gridded texts to shape with different rows and columns. Therefore, tokens in the same row are shifted to place them if there are two tokens occupying the same grid cell.

$$p_c(k) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(k - c_{gt})^2}{2\sigma^2}}$$

$$p_r(k) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(k - r_{gt})^2}{2\sigma^2}}$$

Figure 2.3.1.2: Equations to augment gridded texts to shape with different rows and columns [6].

CHAPTER 2

The CUTIE task is similar to the semantic segmentation task, but it has more sparse data distributions. It should be able to handle multi-scale context because key texts can be either close or distant to each other due to different document layouts. By combining multi-resolution features and atrous convolution, the field of the view of filters can be enlarged. There are two proposed network architectures, namely CUTIE-A which contains multi-resolution features and CUTIE-B which contains the combination of atrous or dilated convolution and Atrous Spatial Pyramid Pooling (ASPP). Both network architectures work with word embedding and dropout layers at the beginning and are trained with the cross-entropy loss.

CUTIE-A

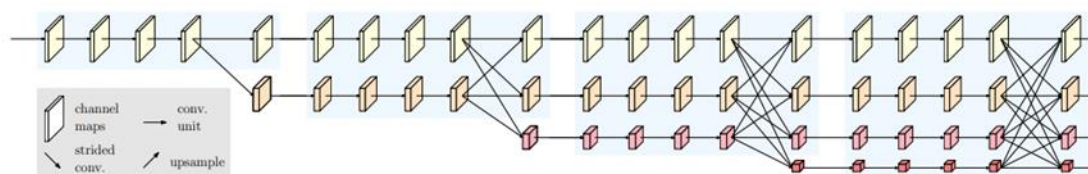


Figure 2.3.1.3: High-Resolution Network (HRNet) architecture [7].

Based on the architecture provided in figure 2.3.1.3, there are four different pathways where a high and consistent resolution for all features will be maintained in each pathway. The strided convolution will reduce the resolution and the size of channel maps. The channel map of the last pathway will be tiny. Left channel maps have low semantic whereas right channel maps have high semantic. Upper channel maps have high resolution which are good at detecting small objects whereas lower channel maps have low resolution which are good at detecting big objects. In other words, high resolution helps in detecting cells which are near to each other whereas low resolution helps in detecting cells which are far apart.

CUTIE-B

layer name	operations	input dimension	output diemnsion	comments
embedding layer	-	20000	128	
conv block	$[3 \times 5] \times 4$	256	256	stride=1
atrous conv block	$[3 \times 5] \times 4$	256	256	stride=1, rate=2
ASPP module	$[3 \times 5] \times 3$, global pooling, concat, 1×1	256	256	stride=1, rate={4,8,16}
shortcut layer	concat, 1×1	256	64	
output layer	1×1	64	9	

Figure 2.3.1.4: Structure of CUTIE-B model [6].

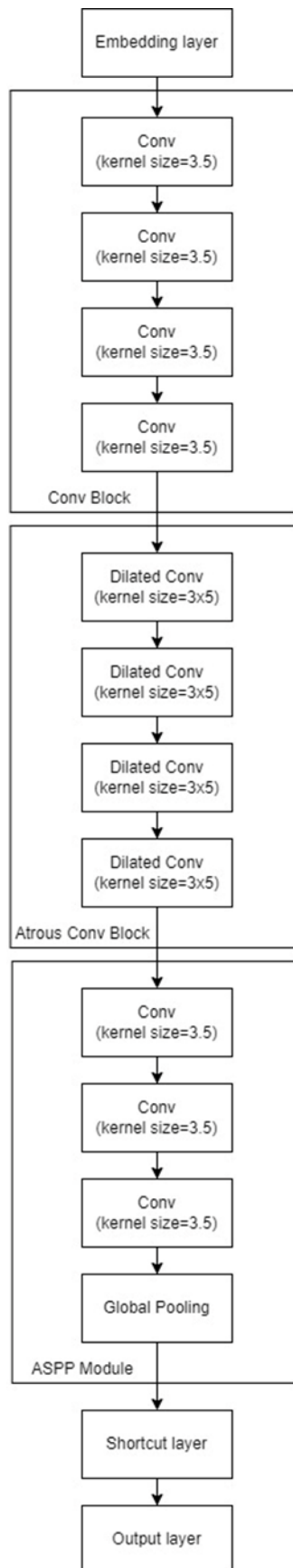


Figure 2.3.1.5: Flow of tokens applying CUTIE-B.

Based on the flow of tokens applying CUTIE-B provided in figure 2.3.1.5, tokens will be embedded into the embedding layer, followed by four consecutive convolution operations in conv block, four consecutive dilated convolution operations in atrous conv block, three consecutive convolution operations and a global pooling in ASPP module, a shortcut layer and ended by an output layer.

Smaller network CUTIE-B outperforms bigger network CUTIE-A because CUTIE-B enlarges the field of view whereas CUTIE-A performs pooling and striding.

2.3.2 Bidirectional Encoder Representations from Transformers (BERT)

Bidirectional Encoder Representations from Transformers (BERT) is a pretrained language model which delivers cutting-edge findings for a wide range of Natural Language Processing (NLP) tasks since it can be fine-tuned for specific NLP task [8]. It learns contextual linkages between words in a text. Since it used to generate a language model, a Transformer encoder mechanism is required to read input texts. To eliminate the difficulty of context learning, it uses two training approaches which are Masked Language Model (MLM) and Next Sentence Prediction (NSP).

For MLM, some input words will be masked randomly, and the masked token will be predicted using its context. The token is predicted by performing Cloze task [9]. Below is an example of a sentence where one of its input words is masked.

My cat is cute -> My cat is [MASK]

However, since the masked token does not appear during the fine-tuning stage, it will result in a mismatch between both pretraining and fine-tuning stages. To eliminate this issue, the chosen word does not have to be replaced by a masked token all of the time. Instead, a masked token is used for 80% of the time, a random token is used for 10% of the time, an original token is used for the remaining 10% of the time. Below is an example to illustrate this concept.

My cat is cute -> My cat is [MASK] (80% of the time)

My cat is cute -> My cat is flying (10% of the time)

My cat is cute -> My cat is cute (10% of the time)

Since the Transformer has no idea which words it will be asked to predict, it is forced to remember the context presentation of each input token.

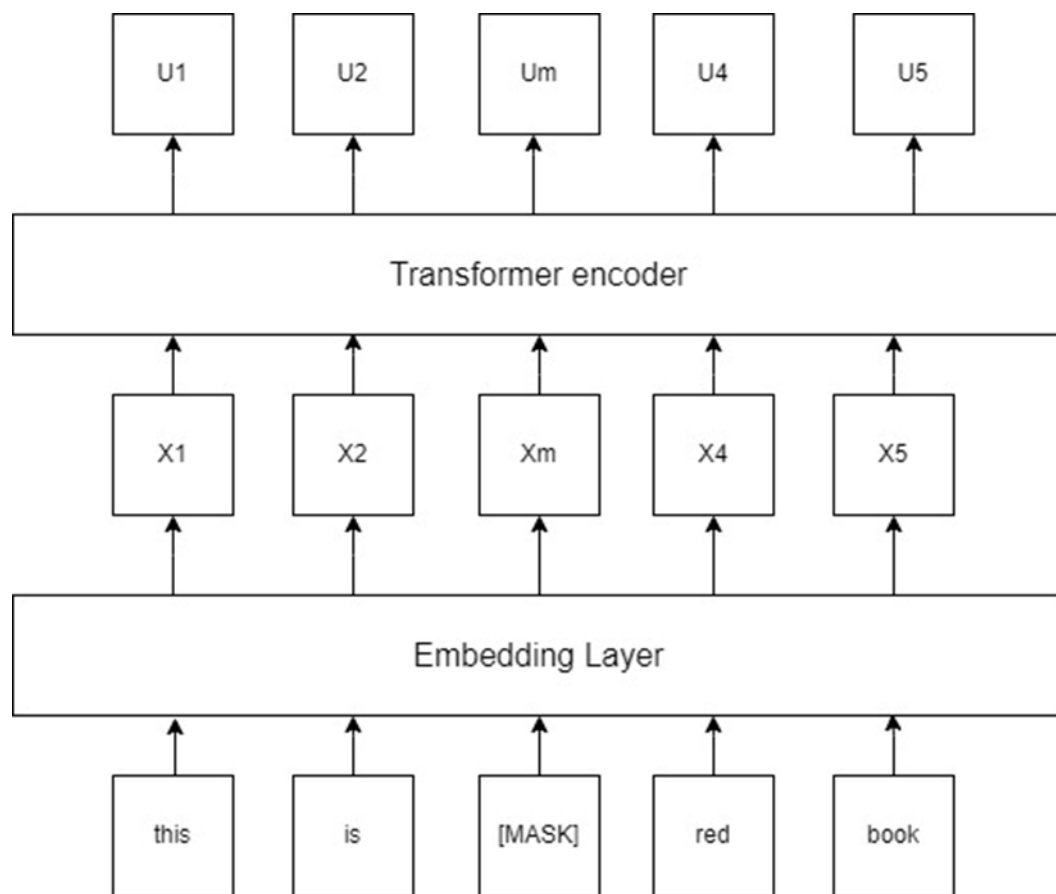


Figure 2.3.2.1: Transformer encoder for Masked Language Model (MLM).

Based on the example provided in figure 2.3.2.1, the original sentence is “this is my red book”. The third word is masked. All inputs are embedded into vectors X and vectors U will be the output of their corresponding inputs. Since Um is not only from Xm but is from all X, although it is the output of the masked word, it is able to know the whole sentence definition. Therefore, Um can be used to predict the original word of the masked word.

For NSP, pairs of sentences will be provided to the model as input for it to learn and predict whether the second sentence is the consecutive sentence from the first sentence. A is the first sentence and B is the second sentence. These two sentences are selected for pretraining. Therefore, it is 50% where B is the actual next sentence of A where another 50% where B is a random sentence from the corpus. Below is an example to illustrate this concept.

“My cat is cute.”

“It has four legs.”

“It was developed by students.”

[CLS] is a token for classification. [SEP] is a token for separating sentences.

Below are the provided two cases and their respective outputs.

- i. [CLS] My cat is cute. [SEP] It has four legs.

Since the second sentence is the actual next sentence of the first sentence, the target will be set as true.

- ii. [CLS] My cat is cute. [SEP] It was developed by students.

Since the second sentence is selected randomly from the corpus, the target will be set as false.

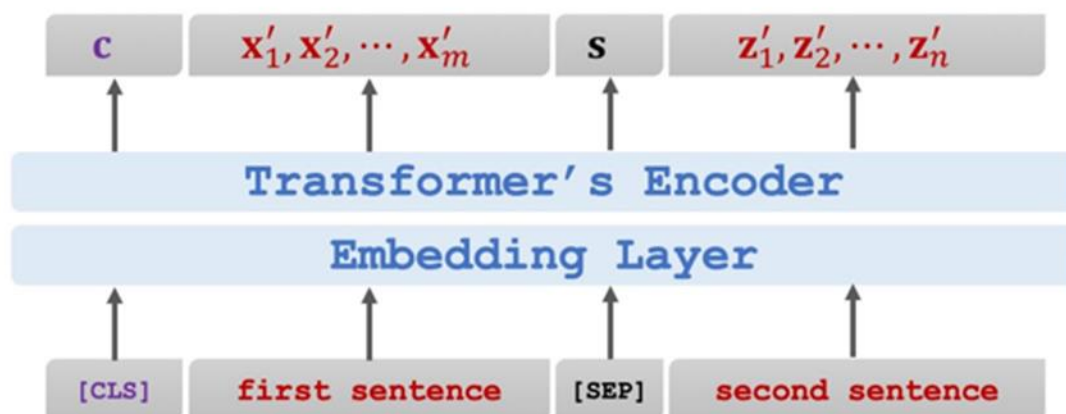


Figure 2.3.2.2: Transformer encoder for Next Sentence Prediction (NSP).

Based on the example provided in figure 2.3.2.2, the output of [CLS] is a vector C where C contains the definition of both sentences. Therefore, C can be used to identify whether both sentences have relationship. When training BERT model, MLM and NSP are trained together to minimize the combined loss function of both strategies where the total loss is calculated by the equation “Total loss = sum of mean masked LM likelihood + mean NSP likelihood” [9].

BERT converts phrases and documents into vectors which capture their meaning [10]. The selected document will be embedded into a vector by turning a chunk of texts into a fixed-size vector which represents semantics of the document. Keywords will be extracted from the

document using simple techniques such as CountVectorizer. Each keyword will be embedded using the same model used previously to embed the document. Cosine similarity between keyword embeddings and document embeddings will be computed to extract the most similar keywords with the highest cosine similarity score. To diverse resulting keywords, Max Sum Similarity (MSS) and Maximal Marginal Relevance (MMR) will be implemented. MSS means computing pairwise similarities between keywords to extract keywords which are the least similar to each other and yet the most similar to the document. MMR starts by selecting keywords which are the most similar to the document. New keywords which are similar to the document and not similar to the selected keywords are selected iteratively. Very diverse results will be obtained if the diversity is set to be high whereas outputs with same keywords which can be repeated will be obtained if the diversity is set to be low.

2.3.3 Character-Aware Neural Language Model

Character-Aware Neural Language model is used to leverage the character-level inputs. The predictions are made at the word-level. It has outperformed majority baseline models which leverage the word embeddings in the input layer [11].

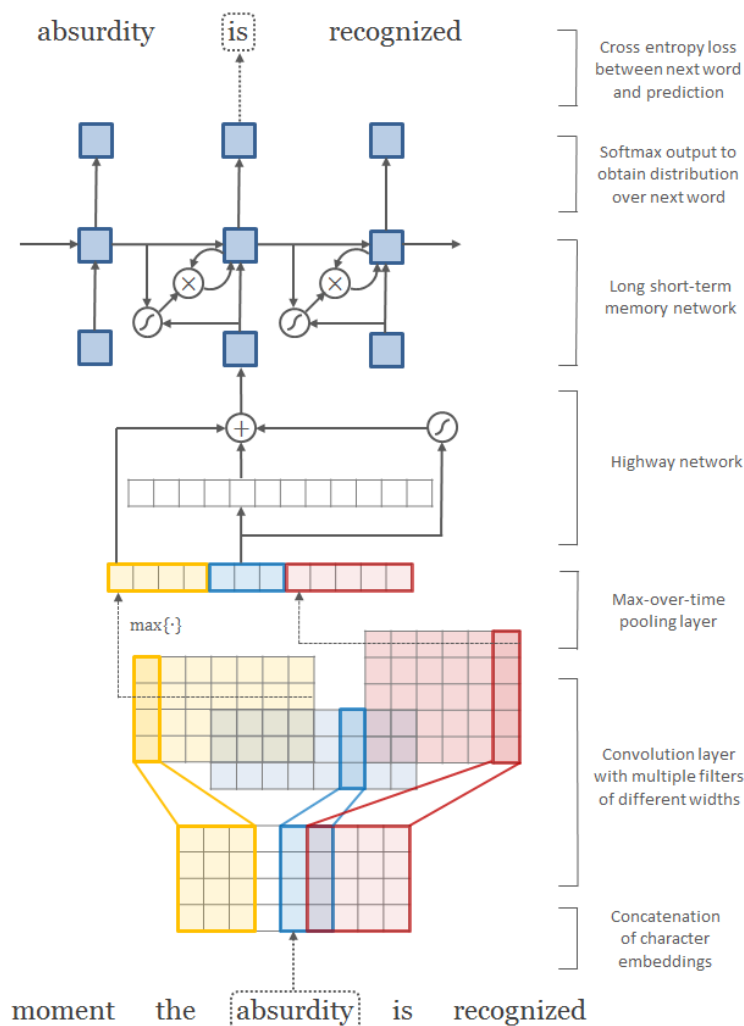


Figure 2.3.3.1: Character-Aware Neural Language model architecture [11].

Based on the architecture provided in figure 2.3.3.1, the input will be accepted by the first layer to perform character embeddings, followed by the convolution operations and max-over-time pooling operation. The fixed-dimensional representation of the word provided by the max-over-time pooling operation is accepted by the highway network. The output of the highway network is accepted by the multi-layer LSTM. An affine transformation followed by a softmax is applied to get the distribution over the next word. Cross entropy loss between the distribution over the next word and actual next word is minimized.

2.4 Limitations of Previous Studies

For OCR, it can be only used to extract the key information without semantic features. Besides, the quality of scanned image can be interrupted when scanning it. The quality of the last image

CHAPTER 2

depends on the quality of the first scanned image. Since it does not guarantee 100% accuracy, all generated outputs should be checked carefully and corrected manually if necessary.

For BERT, a huge text dataset is required to pretrain it. Due to the limited dataset provided, it is not able to perform well.

An independent text localization model, OCR model, or keyword information extraction model is not sufficient to perform keyword information extraction from a receipt. In this project, the text localization, OCR, and keyword information extraction models are proposed to be connected to perform keyword information extraction from a receipt and integrated to develop a web-based receipt key information extraction system.

Chapter 3

System Design

3.1 System Overview

In this web-based receipt key information extraction system, two main modules are included, namely keyword information extraction module and user management module. For keyword information extraction module, text localization, OCR, and key information extraction sub-modules are included. Three models are used as Application Programming Interface (API). CTPN model is used in text localization sub-module, EasyOCR model is used in OCR sub-module, and Character-Aware Neural Language model is used in key information extraction sub-module. For user management module, the user is able to upload receipts to extract key information from them and manage the records.

3.2 System Architecture Diagram

3.2.1 Keyword Information Extraction Module

i. Connectionist Text Proposal Network (CTPN)

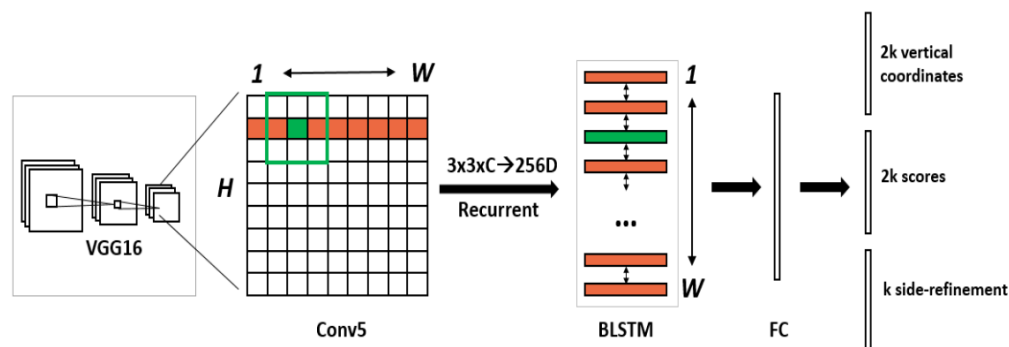


Figure 3.2.1.1: Connectionist Text Proposal Network (CTPN) architecture [12].

In this project, CTPN¹ model is used to perform text localization task [2]. The input of the model is an image whereas its output is a text file which consists of eight coordinates of the detected texts from the image.

¹ Reference code: https://github.com/BlackStar1313/ICDAR-2019-RRR-SROIE/tree/master/text_localization
 Bachelor of Information Systems (Honours) Business Information Systems
 Faculty of Information and Communication Technology (Kampar Campus), UTAR

Based on the architecture provided in figure 3.2.1.1, a 3 x 3 spatial window is slid through the last convolutional maps (Conv5) of the VGG16 model [2]. The consecutive windows in each row are recurrently connected by BLSTM. 3 x 3 x C which is the convolutional feature of each window is the input of 256D BLSTM in which 128D for each LSTM. The RNN layer is connected to 512D fully connected layer. It is followed by the output layer which is used to predict the text or non-text scores, y-axis coordinates and side-refinement offsets of k anchors.

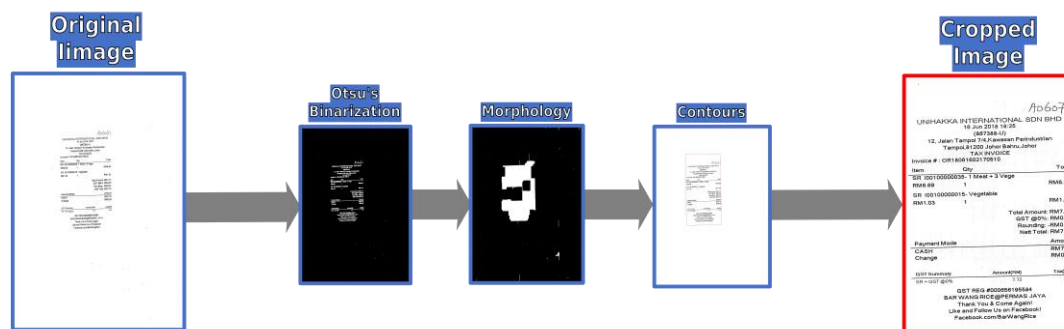


Figure 3.2.1.2: Crop pre-processing flow [13].

Based on the crop pre-processing flow provided in figure 3.2.1.2, the input image consists of a receipt which localizes in the center of the image surrounded by large amount of white space. Otsu's binarization is used to determine the optimal threshold value of the image by going through each possible threshold value from 0 to 255. Morphology is used to extract the useful components of the image which might be representing in a region of shape, boundaries, etc. Contour is an outline which depicts or delimits the form of shape of an object in the image and is useful for object recognition. The receipt has been detected in the center of the image and the original image has been cropped to get the image which consists of the receipt without the white space around it.

ii. EasyOCR

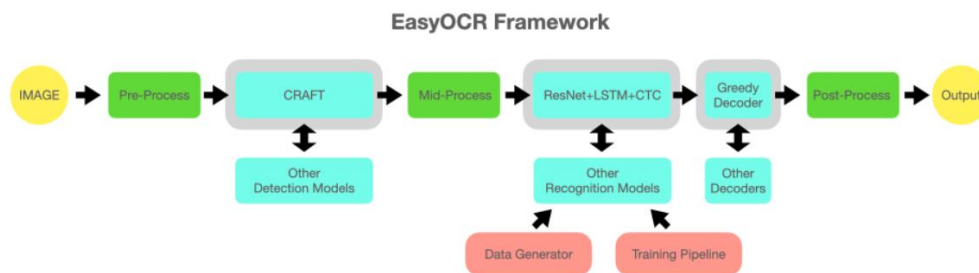


Figure 3.2.1.3: EasyOCR architecture [14].

In this project, EasyOCR² model is used to perform OCR task. Eight coordinates of the detected texts from the input image of CTPN model will be used to obtain the cropped image from the input image in which there are detected texts in the cropped image. The input of EasyOCR model is a cropped image whereas its output is the text extracted from the cropped image.

The input image will be accepted by the EasyOCR model, and the input will be applied multiple processes as mentioned in figure 3.2.1.3. In the end, an output which is the text extracted from the image will be obtained.

² Reference code: <https://github.com/JaidedAI/EasyOCR>
 Bachelor of Information Systems (Honours) Business Information Systems
 Faculty of Information and Communication Technology (Kampar Campus), UTAR

iii. Character-Aware Neural Language Model

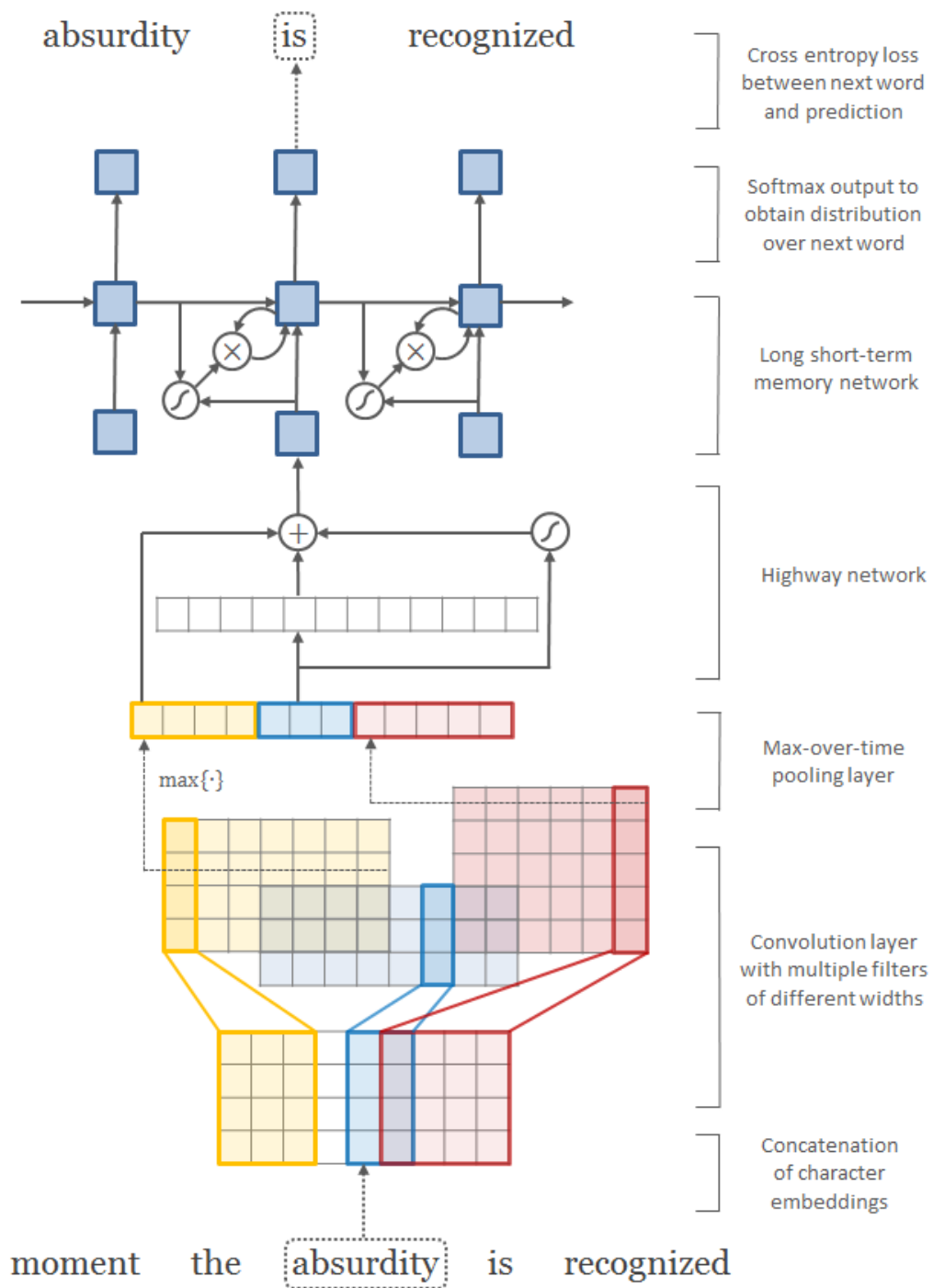


Figure 3.2.1.4: Character-Aware Neural Language model architecture [15].

In this project, Character-Aware Neural Language³ model is used to perform key information extraction task [11]. The inputs of the model are the input image of CTPN model and its respective text file which consists of eight coordinates of the detected text from the input image and the extracted texts from the cropped images by EasyOCR model. The output of the model is the key information extracted from the input image in which the model is able to extract four details, namely company, address, date, and total.

Based on the architecture provided in figure 3.2.1.4, the input will be accepted by the first layer to perform character embeddings, followed by the convolution operations and max-over-time pooling operation [11]. The fixed-dimensional representation of the word provided by the max-over-time pooling operation is accepted by the highway network. The output of the highway network is accepted by the multi-layer LSTM. An affine transformation followed by a softmax is applied to get the distribution over the next word. Cross entropy loss between the distribution over the next word and actual next word is minimized.

3.2.2 User Management Module

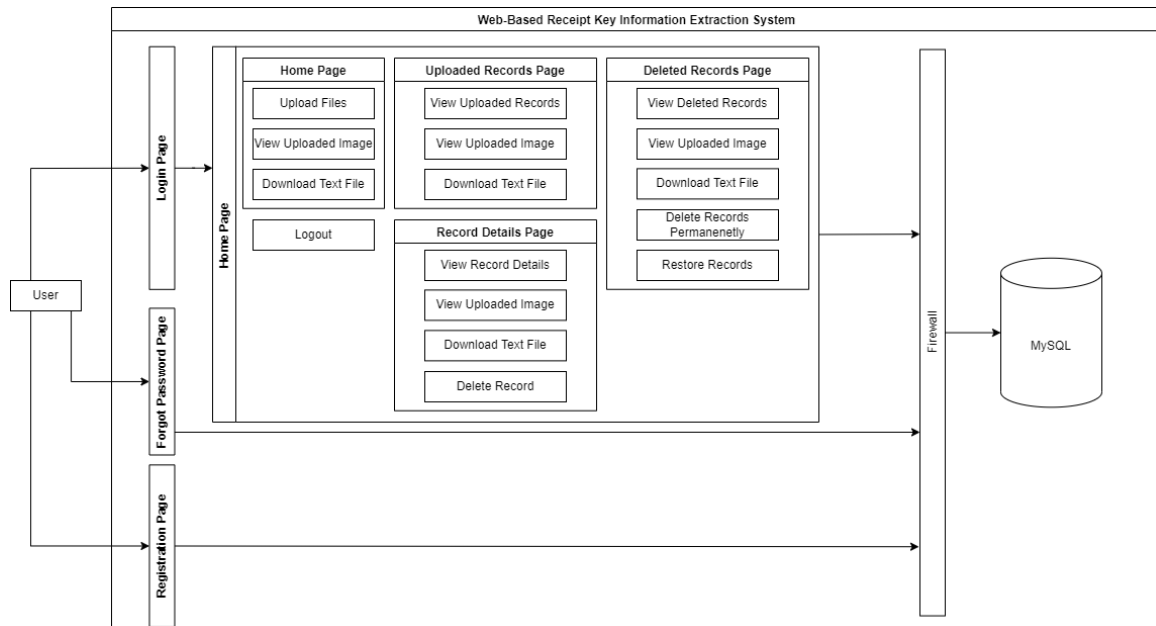


Figure 3.2.2.1: Web-based receipt key information extraction system architecture.

³ Reference code: https://github.com/BlackStar1313/ICDAR-2019-RRC-SROIE/tree/master/keyword_information_extraction

CHAPTER 3

The user is able to login in “login” page, reset password in “forgot password” page, and register in “registration” page. The password stored in MySQL server will be updated after the user resets password. The information provided during the registration will be added into MySQL server. After the user logs in to the system, the user will be navigated from “login” page to home page. In home page, the user is able to upload files, view uploaded image, and download text file. In “uploaded records” page, the user is able to view uploaded records, view uploaded image, and download text file. In “record details” page, the user is able to view record details, view uploaded image, download text file, and delete record. In “deleted records” page, the user is able to view deleted records, view uploaded image, download text file, delete records permanently, and restore records. All information will be added, retrieved, updated, or deleted from MySQL server if the condition has been fulfilled. In the end, there is a “logout” option for the user to logout.

3.3 System Block Diagram

3.3.1 User Management Module

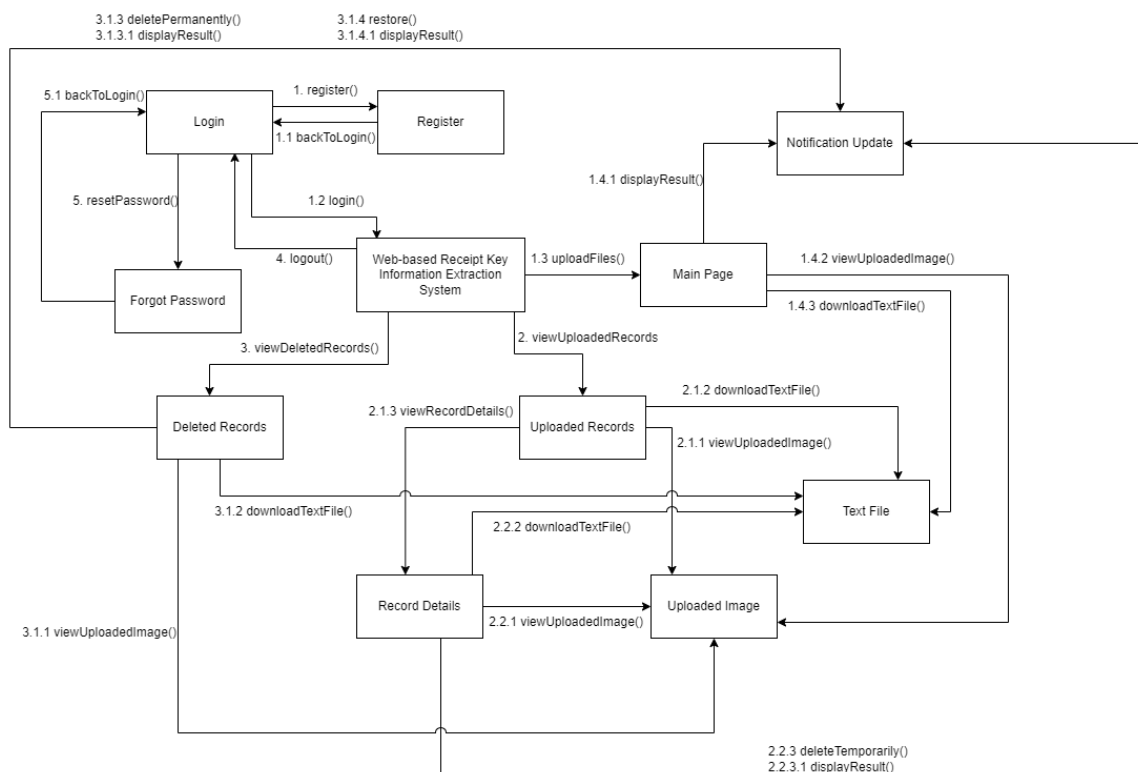


Figure 3.3.1.1: Web-based receipt key information extraction system block diagram.

1. register()

In “login” page, the user will be navigated to “registration” page to register a user account.

1.1 backToLogin()

The user will be navigated to “login” page from “registration” page.

1.2 login()

The user is able to login to the web-based receipt key information extraction system.

1.3 uploadFiles()

The user is able to upload multiple files to extract the keyword information from the uploaded files in the main page.

1.4.1 displayResult ()

The user is able to view the result returned by the system as the notification update will be triggered by the system to display result once the uploaded files have been processed.

1.4.2 viewUploadedImage()

The user is able to view image in a new tab by clicking on the selected image URL.

CHAPTER 3

1.4.3 downloadTextFile()

The user is able to download text file by clicking on the selected text file URL.

1.4.1, 1.4.2, 1.4.3 are the options can be selected by the user in the main page after 1.3.

2. viewUploadedRecords()

The user is able to view the uploaded records in “uploaded records” page.

2.1.1 viewUploadedImage()

The user is able to view image in a new tab by clicking on the selected image URL.

2.1.2 downloadTextFile()

The user is able to download text file by clicking on the selected text file URL.

2.1.3 viewRecordDetails()

The user is able to view the record details of the selected record id.

2.1.1, 2.1.2, 2.1.3 are the options can be selected by the user in “uploaded records” page after 2.

2.2.1 viewUploadedImage()

The user is able to view image in a new tab by clicking on the selected image URL.

2.2.2 downloadTextFile()

The user is able to download text file by clicking on the selected text file URL.

2.2.3 deleteTemporarily()

The user is able to delete the record of the selected record id temporarily.

2.2.3.1 displayResult()

The user is able to view the result returned by the system as the notification update will be triggered by the system to display result once the delete request has been processed.

2.2.4 backToUploadedRecords()

The user will be navigated to “uploaded records” page from “record details” page.

2.2.1, 2.2.2, 2.2.3 are the options can be selected by the user in “record details” page after 2.1.3.

2.2.3.1 can only be carried out after 2.2.3.

3. viewDeletedRecords()

The user is able to view the deleted records in “deleted records” page.

3.1.1 viewUploadedImage()

The user is able to view image in a new tab by clicking on the selected image URL.

3.1.2 downloadTextFile()

The user is able to download text file by clicking on the selected text file URL.

CHAPTER 3

3.1.3 deletePermanently()

The user is able to delete the record of the selected record id permanently.

3.1.3.1 displayResult()

The user is able to view the result returned by the system as the notification update will be triggered by the system to display result once the delete request has been processed.

3.1.4 restore()

The user is able to restore the record of the selected record id.

3.1.4.1 displayResult()

The user is able to view the result returned by the system as the notification update will be triggered by the system to display result once the restore request has been processed.

3.1.1, 3.1.2, 3.1.3, 3.1.4 are the options can be selected by the user in the deleted records page after 3.

3.1.3.1 can only be carried out after 3.1.3.

3.1.4.1 can only be carried out after 3.1.4.

4. logout()

The user is able to logout and will be navigated to “login” page.

5. resetPassword()

In “login” page, the user will be navigated to “forgot password” page to reset password.

5.1 backToLogin()

The user will be navigated to “login” page from “forgot password” page.

3.4 Use Case Diagram for User Management Module

3.4.1 Use Case Diagram for Project I

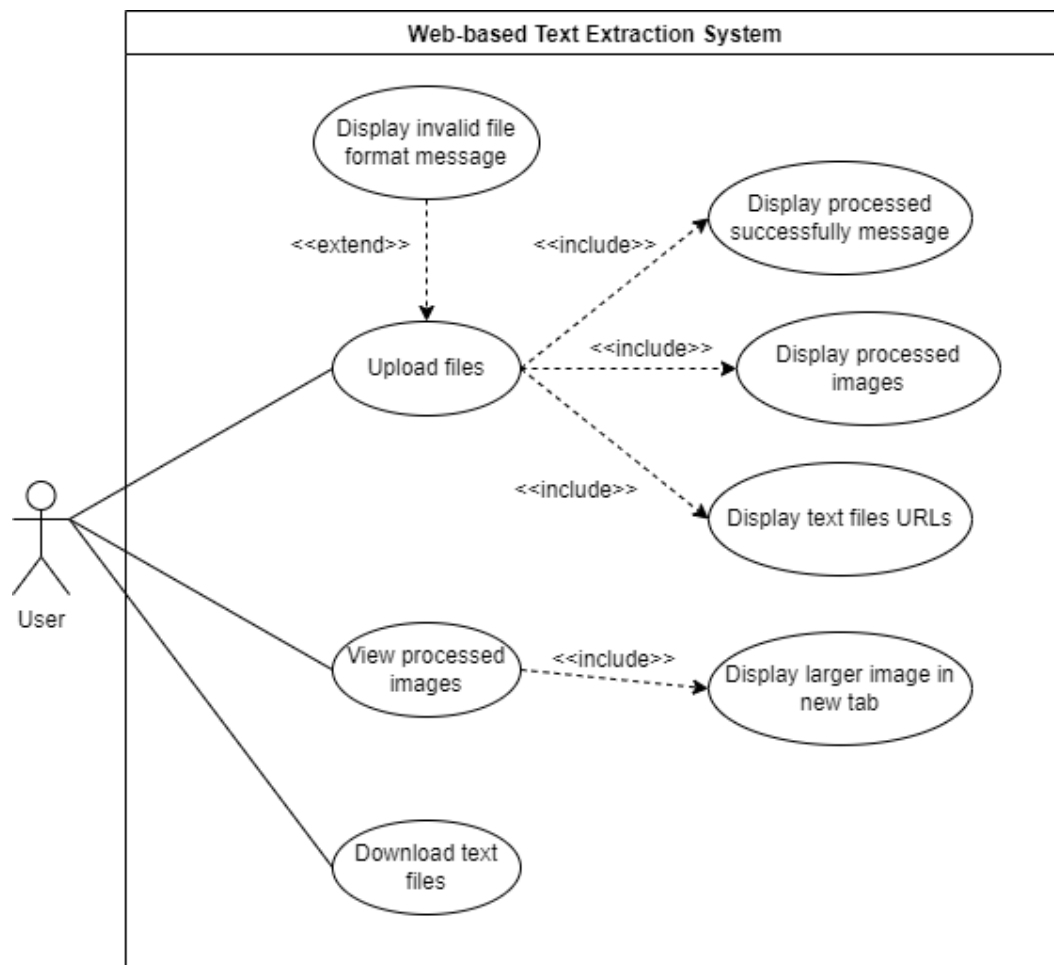


Figure 3.4.1.1: Use case diagram for Project I.

The user is able to upload files, view processed images, and download text files. In normal condition, after uploading the files, successfully processed message, processed images, and text files URLs will be displayed since the uploaded files are in the correct format which are jpg, jpeg, and PNG. If the files with invalid file format have been uploaded, invalid file format message will be displayed. When user selects to view processed images, larger image will be displayed in new tab. When user selects the text file URL to download, the user is able to download the text file successfully.

3.4.2 Use Case Diagram for Project II



Figure 3.4.2.1: Use case diagram for Project II.

The user is able to register, login, logout, reset password, upload files, view uploaded images, download text files, view uploaded records, view record details, view deleted records, delete records, and restore records.

If the user has successfully registered a user account, successfully registered message will be displayed. Otherwise, existing email message will be displayed to notify the user the email used to register has been used.

If the user has logged in successfully, the user will be navigated to “upload” page which is the main page of the system. Otherwise, invalid email or password message will be displayed to notify the user that the input is incorrect.

If the user selects to log out, the user will be navigated to “login” page.

CHAPTER 3

If the user has successfully reset password, successfully reset password message will be displayed. Otherwise, not a registered email message will be displayed to notify the user the email used to reset password is not a registered email.

If the user uploads files successfully, result such as files have been processed message, uploaded images, and the text files URLs will be displayed. If uploaded files with invalid file format have been detected, invalid file format message will be displayed.

If the user selects to view uploaded images, the image will be displayed in new tab.

The user is able to download text files.

If the user selects to view uploaded records, uploaded records will be displayed if records are found. Otherwise, no record found message will be displayed.

If the user selects to view record details, record details will be displayed.

If the user selects to view deleted records, deleted records will be displayed if records are found. Otherwise, no record found message will be displayed.

If the user selects to delete records, delete successfully message will be displayed.

If the user has successfully restored records, successfully restore message will be displayed. Otherwise, failed to restore message will be displayed.

3.5 Activity Diagram for User Management Module

3.5.1 Activity Diagram for Project I

i. “Upload files”

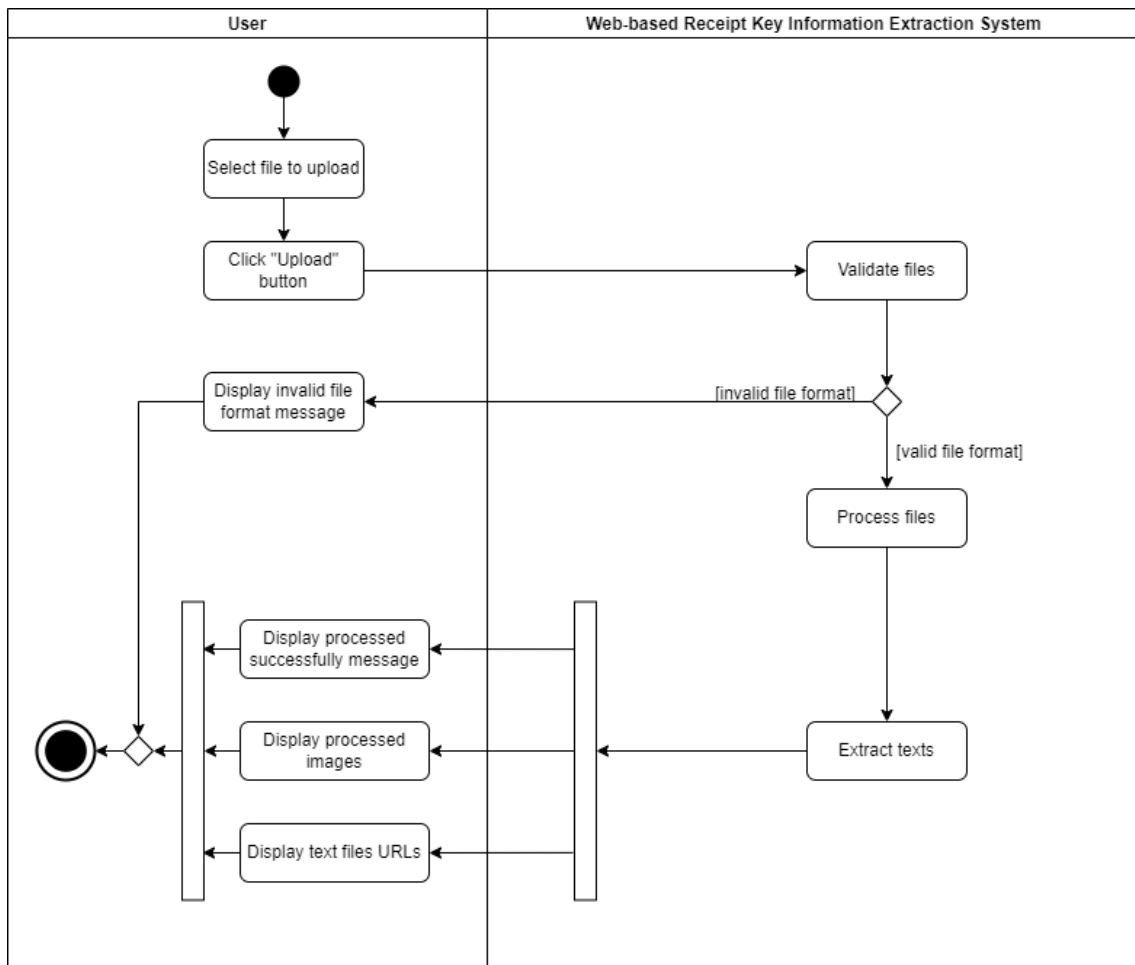


Figure 3.5.1.1: “Upload files” activity diagram for Project I.

When the user has selected files to upload, after clicking “upload” button, the files will be validated by the system. If the files format is invalid, invalid file format message will be displayed. Otherwise, the files will continue to be processed by the system. After extracting the texts, successfully processed message, processed images, and text files URLs will be displayed simultaneously.

ii. “View processed images”

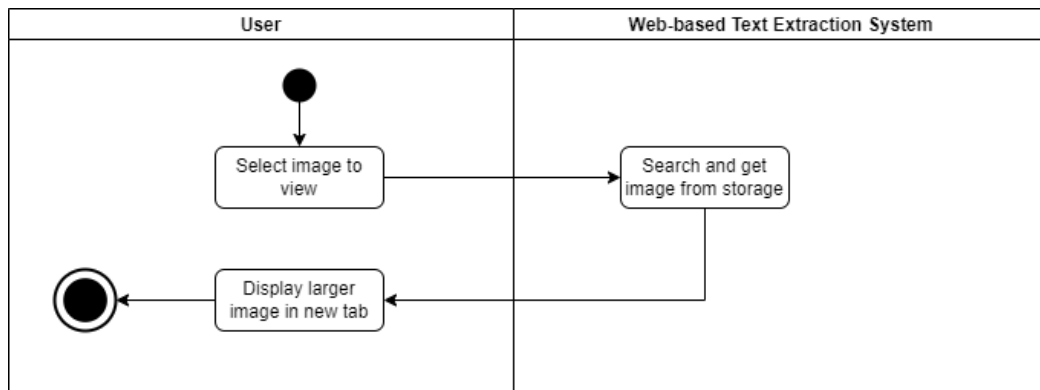


Figure 3.5.1.2: “View processed images” activity diagram for Project I.

When the user has selected image to view, the system will search and get the related image from storage. Larger image will be displayed in new tab.

iii. “Download text files”

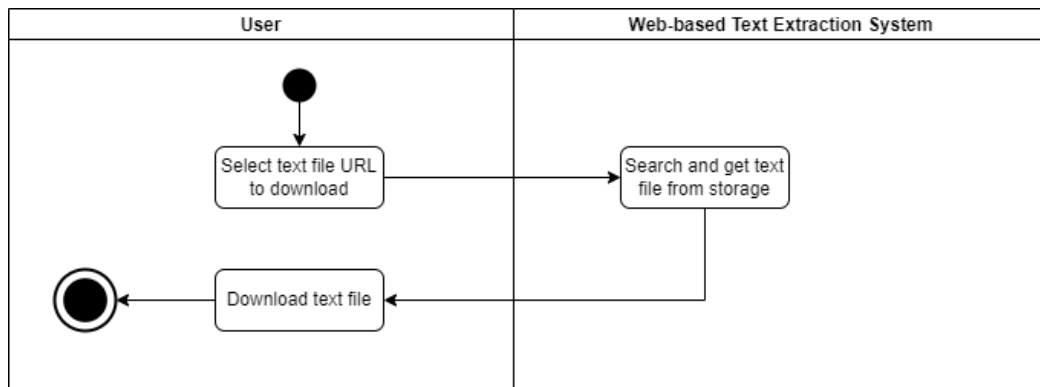


Figure 3.5.1.3: “Download text files” activity diagram for Project I.

When the user has selected text file URL to download, the system will search and get the related text file from storage. The user is able to download the text file.

3.5.2 Activity Diagram for Project II

i. “Register”

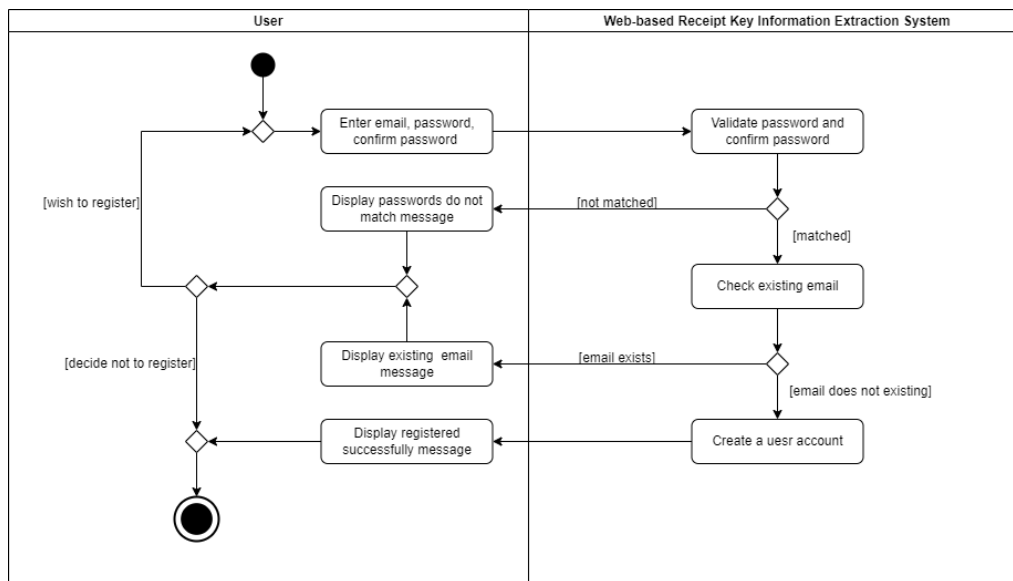


Figure 3.5.2.1: “Register” activity diagram for Project II.

When the user has entered email, password, and confirm password, the system will validate the password and confirm password whether both of them are matched. If they do not match, passwords do not match message will be displayed and the user is required to continue to enter the confirm password which is matched with the password if the user wishes to continue the registration. If both passwords are matched, the system will continue checking whether the email is exists. If it is an existing email, existing email message will be displayed to notify the user the email has been used to register an account and the user is required to continue to enter another email if the user wishes to continue the registration. Otherwise, a user account will be created successfully, and successfully registered message will be displayed.

ii. “Login”

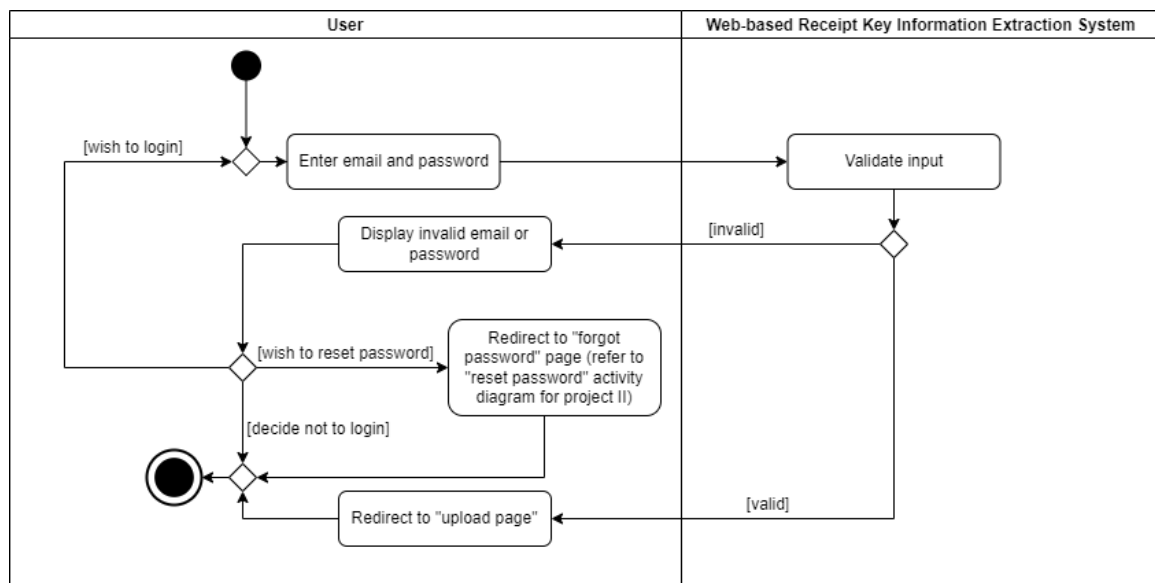


Figure 3.5.2.2: “Login” activity diagram for Project II.

When the user has entered email and password, the system will validate them by checking the existing records in database. If the email and password are valid, the user will be navigated to “upload” page which is the home page of the system. Otherwise, invalid email or password will be displayed. If the user wishes to continue to login, the user can choose to reenter the email and password. If the user wishes to reset password, the user will be navigated to “forgot password” page if the user has selected to reset password and further explanation has been provided in the “reset password” activity diagram for project II.

iii. “Logout”

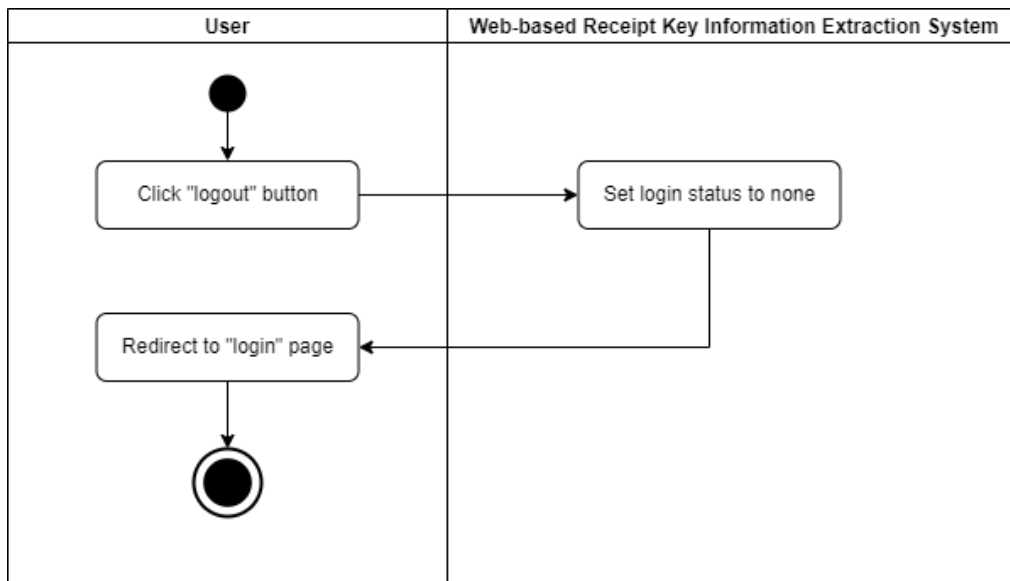


Figure 3.5.2.3: “Logout” activity diagram for Project II.

When the user clicks on “logout” button, the system will set the login status to none which means there is no one has logged in to the system. The user will be navigated to “login” page.

iv. “Reset password”

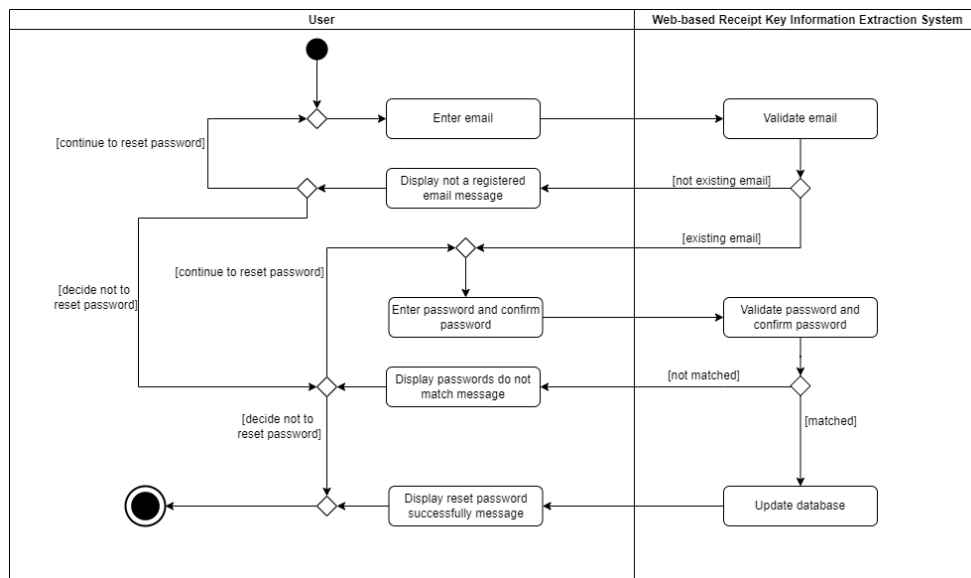


Figure 3.5.2.4: “Reset password” activity diagram for Project II.

When the user has entered the email, the system will validate the email whether it is an existing email. If it is not an existing email, not a registered email message will be displayed to notify the user the entered email is an invalid email which is not able to be used to reset the password. If the password does not match, a message will be displayed. If the password matches, the system will update the database and display a successful message. The user can then decide to continue or stop the process.

password of the user account. The user can choose to continue the reset password process by entering a valid email. If it is an existing email, the user is able to continue to reset the password by entering new password and confirm password. If the passwords are matched, the system will replace the old password stored in database by the new password and successfully reset password message will be displayed. Otherwise, passwords do not match message will be displayed and the user can choose to continue to reset password by entering a pair of valid passwords.

v. “Upload files”

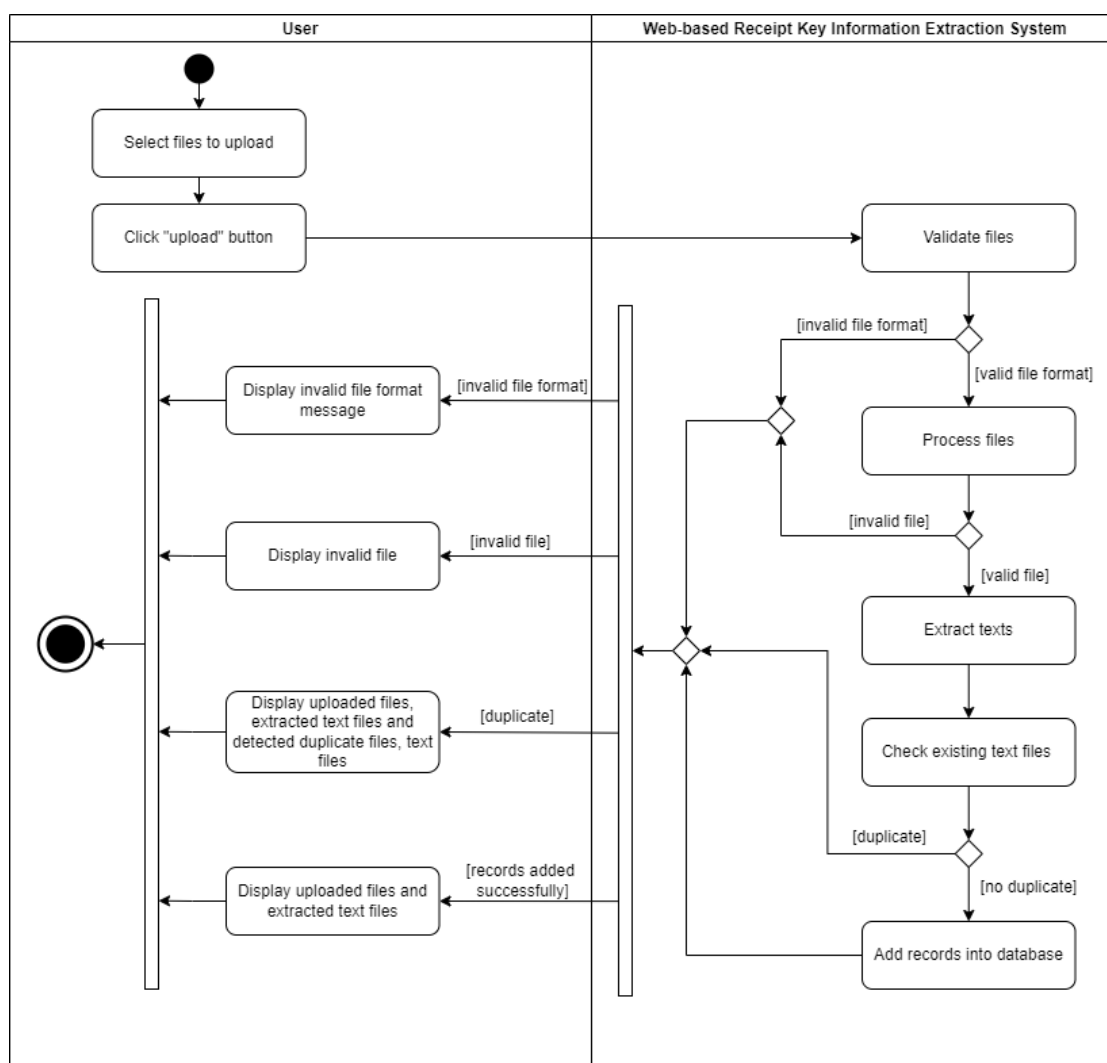


Figure 3.5.2.5: “Upload files” activity diagram for Project II.

When the user has selected files to upload, after clicking on “upload” button, the system will validate the files whether they have valid file format. If the files have invalid file format, invalid file format message will be displayed. Otherwise, the files will continue to be

processed by the system. If the files are invalid, invalid file message will be displayed. Otherwise, the system will extract texts from the files and check and compare them with the existing text files. If duplicate text files have been detected, uploaded files and their corresponding extracted text files, detected duplicated files and their corresponding text files will be displayed. Otherwise, the records will be added into database and uploaded files and extracted text files will be displayed. All messages will be displayed simultaneously if the condition has been fulfilled.

vi. “View uploaded images”

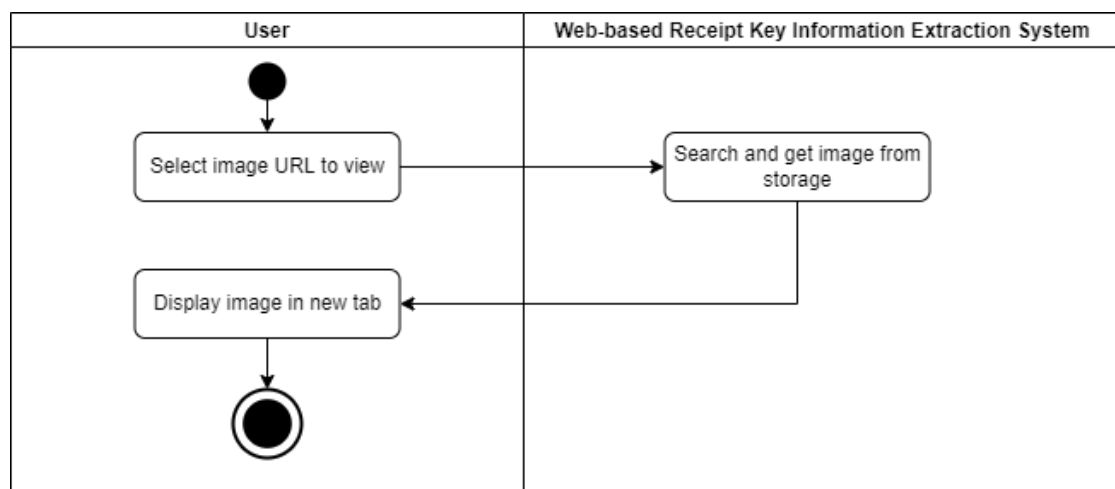


Figure 3.5.2.6: “View uploaded images” activity diagram for Project II.

When the user has selected the image URL to view, the system will search and get the required image from storage. Image will be displayed in new tab.

vii. “Download text files”

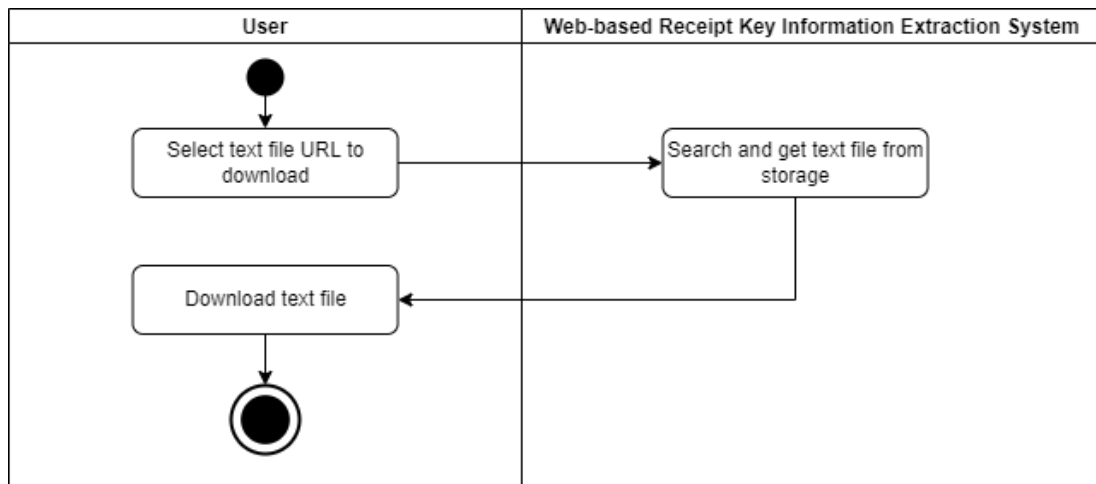


Figure 3.5.2.7: “Download text files” activity diagram for Project II.

When the user has selected text file URL to download, the system will search and get the required text file from storage. The user is able to download the text file.

viii. “View uploaded records”

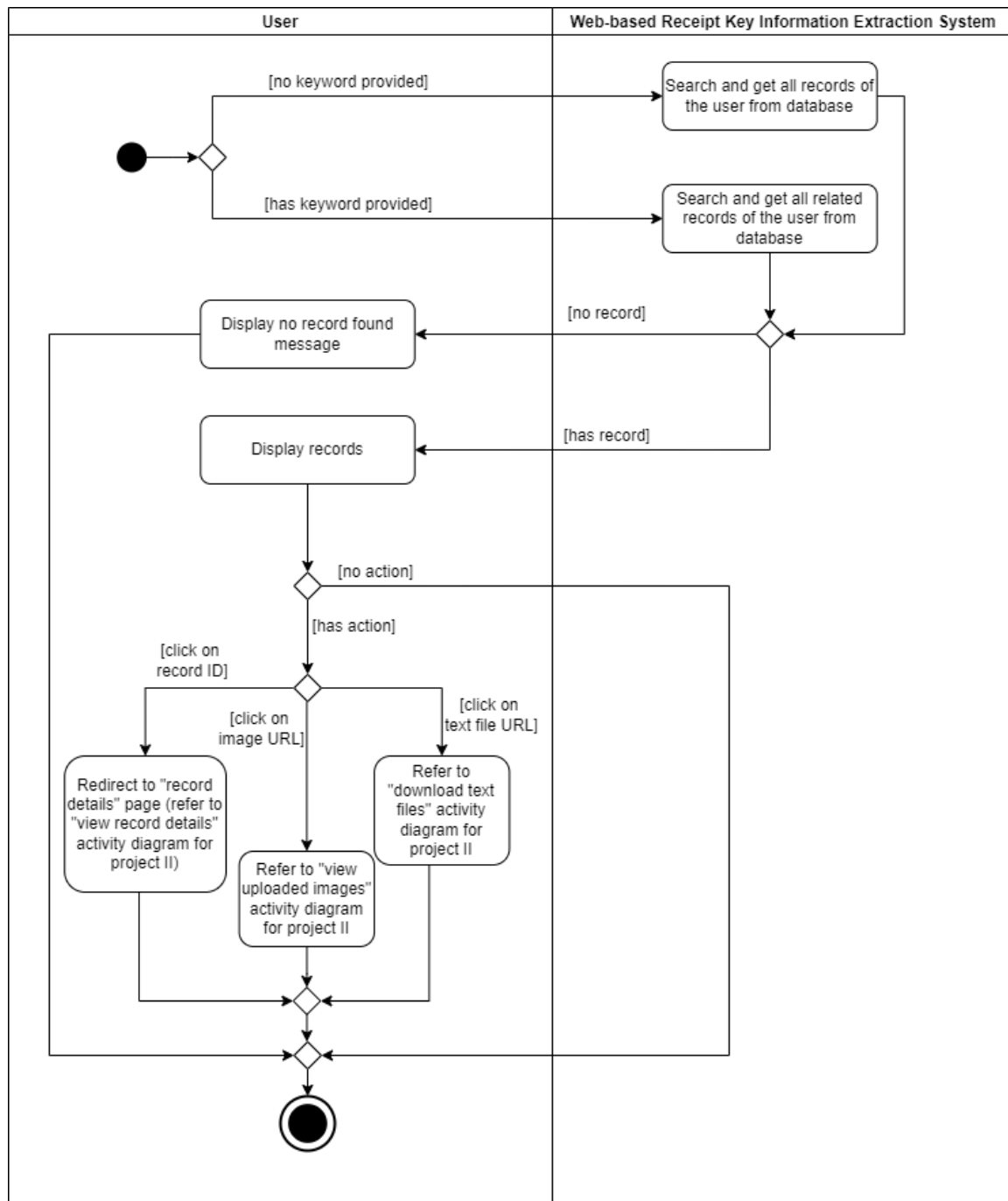


Figure 3.5.2.8: “View uploaded records” activity diagram for Project II.

When the user does not provide keyword, the system will search and get all records of the user from database. Otherwise, the system will search and get all related records of the user from database based on the keyword provided. After retrieving the records, no record found message will be displayed if there is no record found. Otherwise, records will be displayed. If the user has clicked on the record id, the user will be navigated to the corresponding “record

details” page and further explanation has been provided in the “view record details” activity diagram for project II. If the user has clicked on the image URL, further explanation has been provided in the “view uploaded images” activity diagram for project II. If the user has clicked on the text file URL, further explanation has been provided in the “download text files” activity diagram for project II.

ix. “View record details”

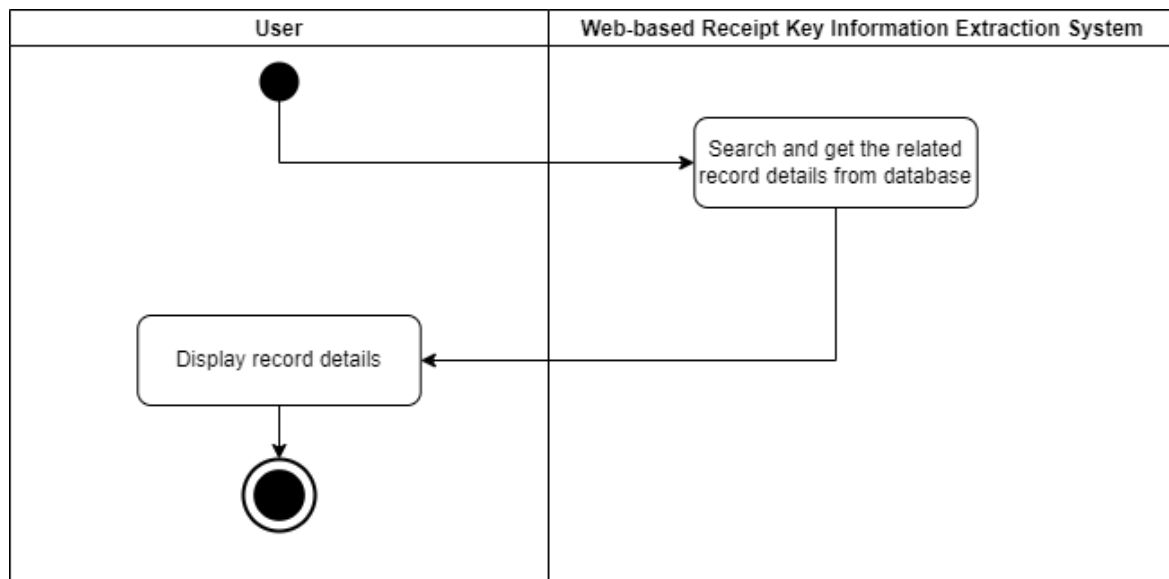


Figure 3.5.2.9: “View record details” activity diagram for Project II.

The system will search and get the related record details from database and display record details to the user.

x. “View deleted records”

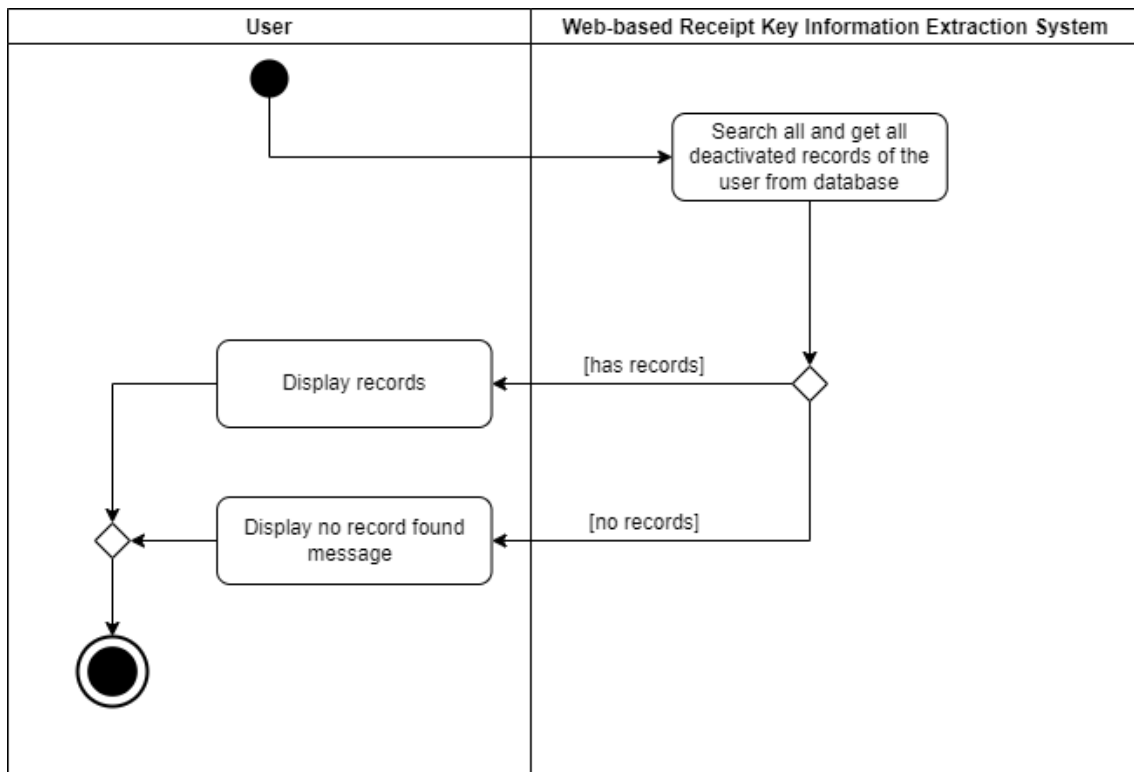


Figure 3.5.2.10: “View deleted records” activity diagram for Project II.

The system will search and get all deactivated records of the user from database. After retrieving the records, no record found message will be displayed if there is no record found. Otherwise, records will be displayed.

xi. “Delete records”

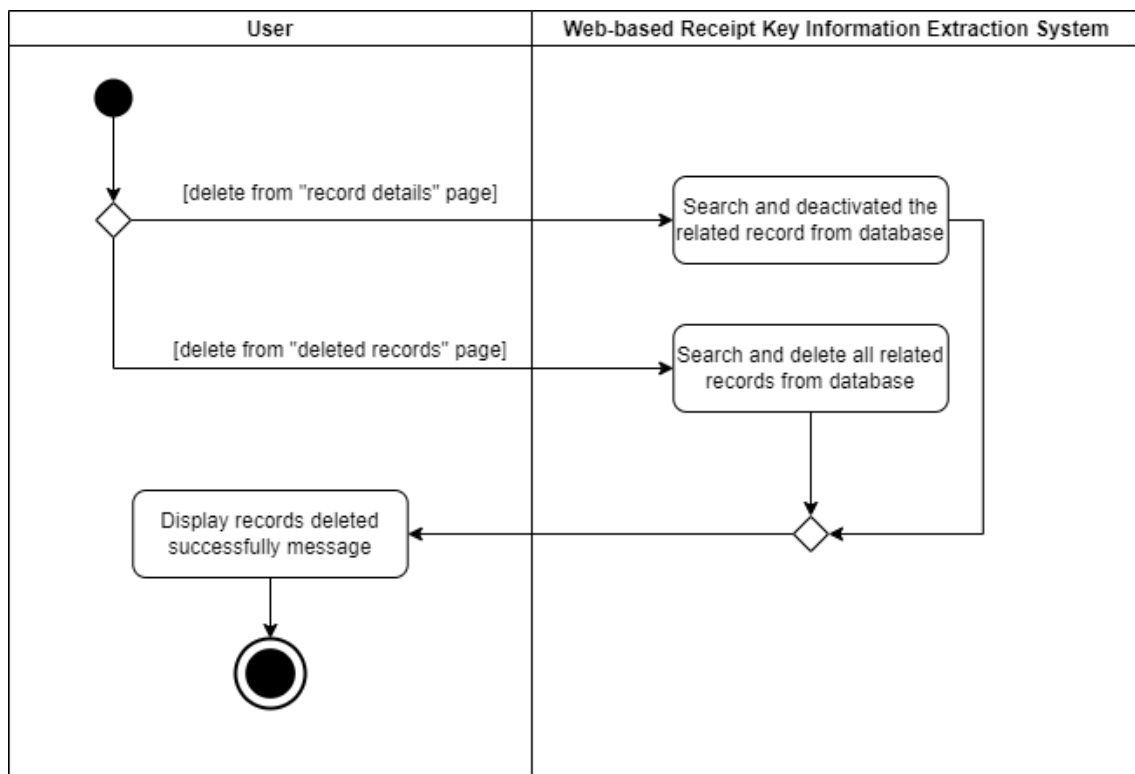


Figure 3.5.2.11: “Delete records” activity diagram for Project II.

There are two “delete” button where one is from “record details” page and another is from “deleted records” page. If the delete request is from “record details” page, the system will search and deactivate the related record from database. In this case, the records are deleted temporarily. If the delete request is from “deleted records” page, the system will search and delete all related records from database. In this situation, the records are deleted permanently.

xii. “Restore records”

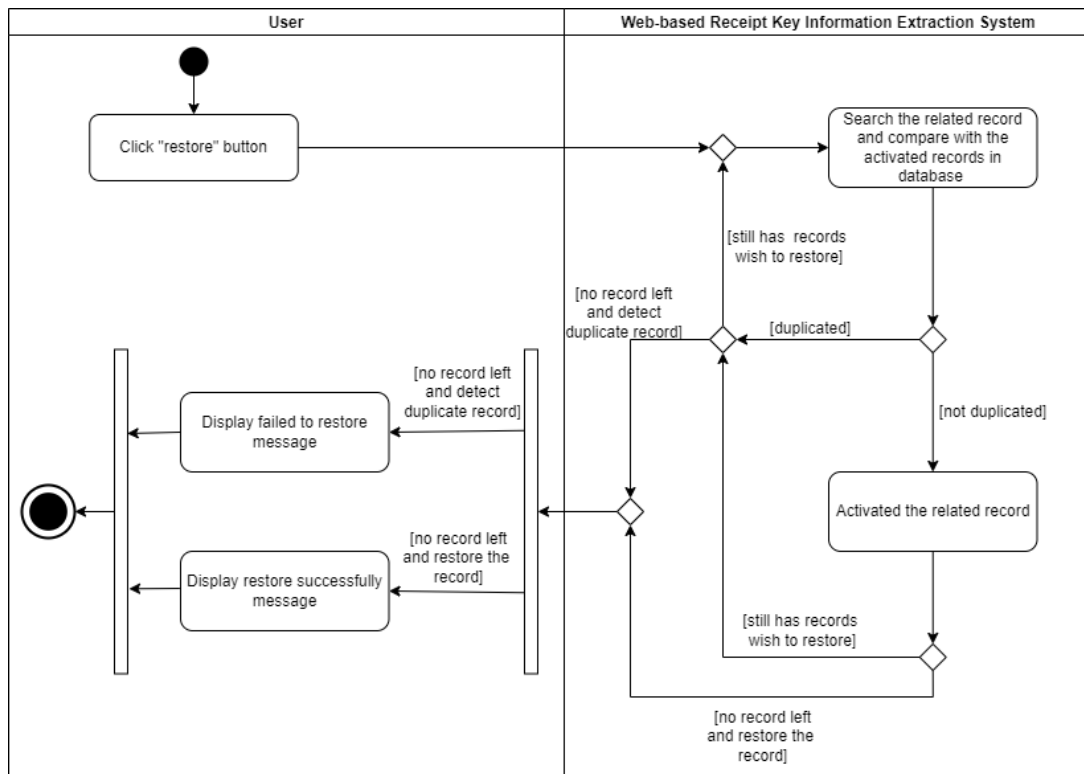


Figure 3.5.2.12: “Restore records” activity diagram for Project II.

When the user has clicked on “restore” button, the system will search the related record and compare with the activated records in database. If duplicate record has been detected, failed to restore message will be displayed. Otherwise, the related record will be activated by the system and successfully restore message will be displayed. If there are still records wish to be restored, the previous process will be repeated by the system. In the end, all messages will be displayed simultaneously if the condition has been fulfilled.

3.6 Database Design

Table	Attribute	Data Type	Description
user	user_id	integer (11)	Auto incremented primary key in table “user” of database “kie”.
	email	varchar (255)	Email of user.
	password	varchar (255)	Password of user.

record	record_id	integer (11)	Auto incremented primary key in table “record” of database “kie”.
	image	varchar (255)	Name of image uploaded by user.
	text	varchar (255)	Name of text file where the text is extracted from the image uploaded by user.
	upload_date	varchar (255)	Upload date by user.
	upload_time	varchar (255)	Upload time by user.
	company	varchar (255)	Company extracted from the image uploaded by user.
	address	varchar (255)	Address extracted from the image uploaded by user.
	date	varchar (255)	Date extracted from the image uploaded by user.
	total	varchar (255)	Total extracted from the image uploaded by user.
	user_id	integer (11)	user_id of user.
	status	integer (11)	Status of the record.

Table 3.6.1: Database structure.

3.7 User Interface Design

Below are the designed user interfaces for the system to be developed.

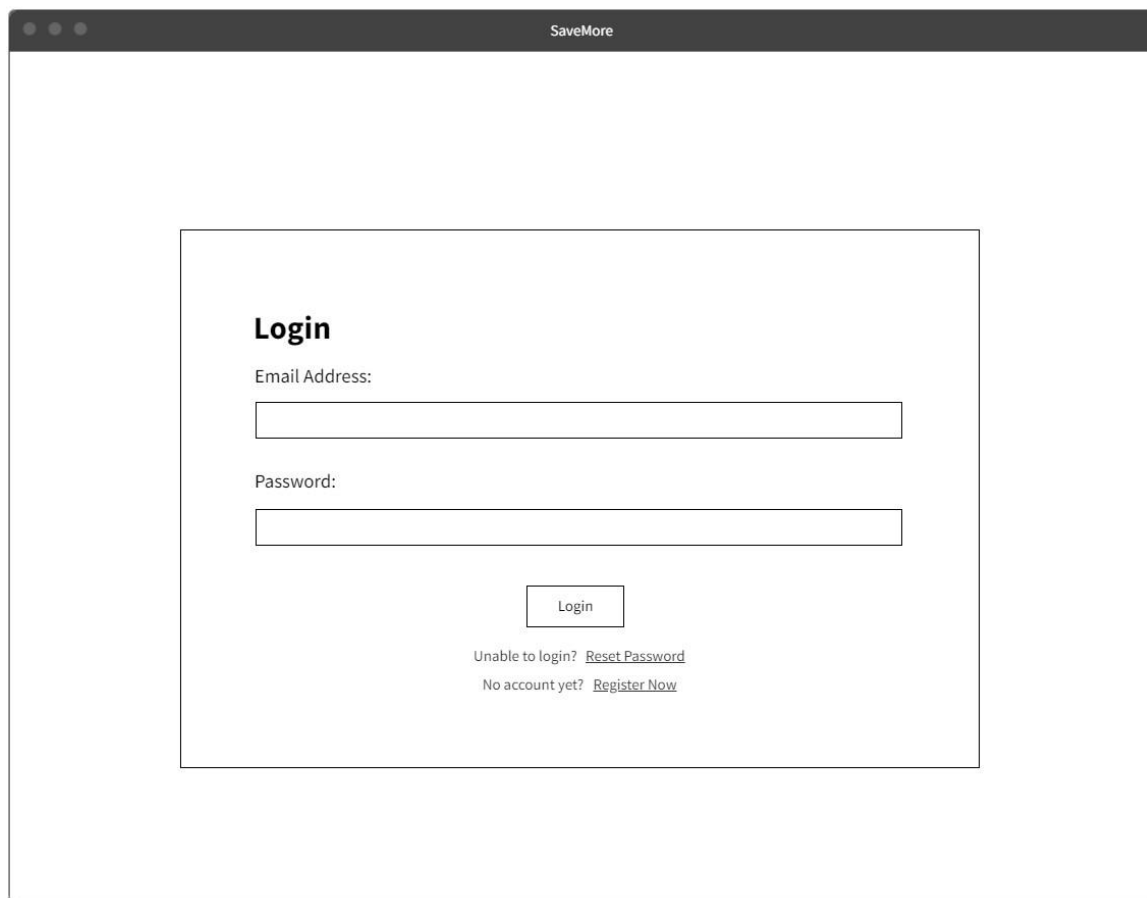


Figure 3.7.1: “login” user interface design.

Login Now'."/>

SaveMore

Register

Email:

Password:

Confirm Password:

Register

Already have an account? [Login Now](#)

Figure 3.7.2: “register” user interface design.

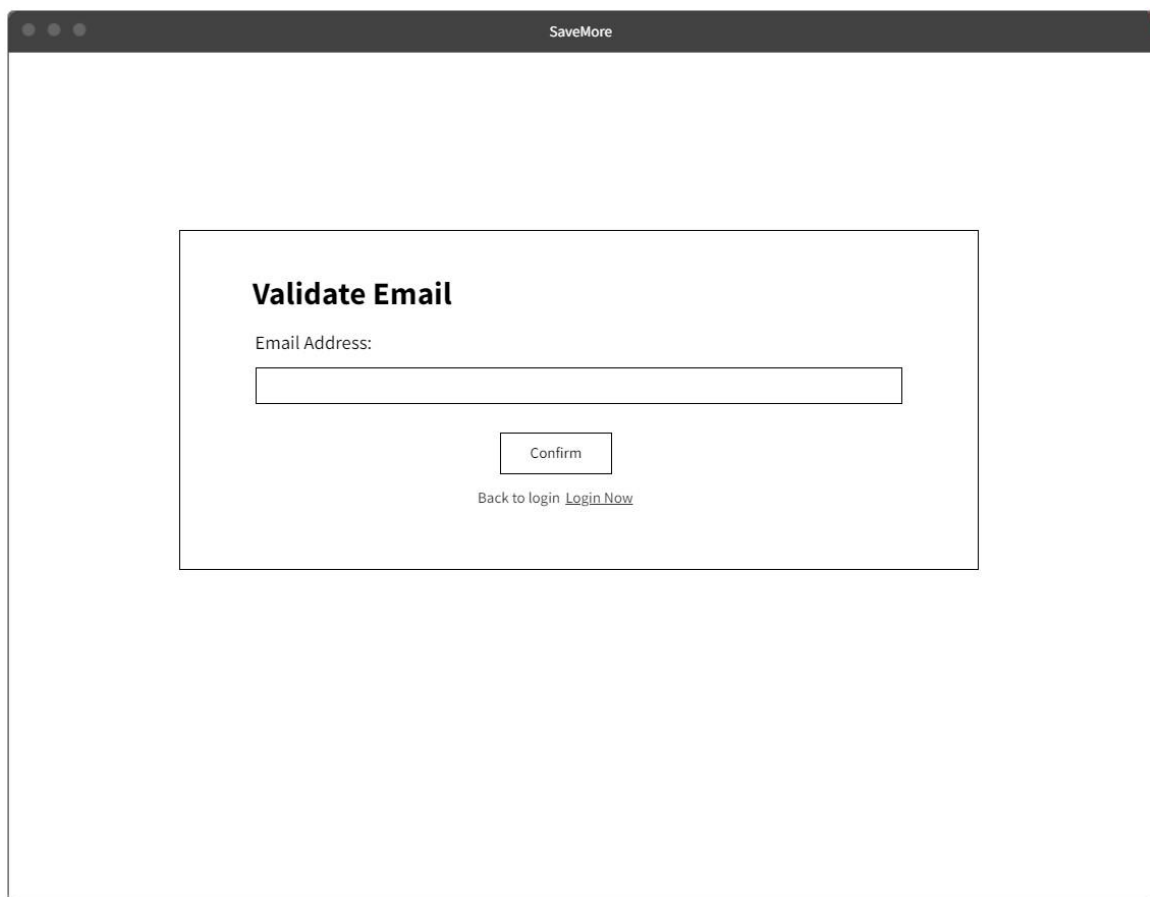


Figure 3.7.3: “validate email” user interface design (first stage to reset password).

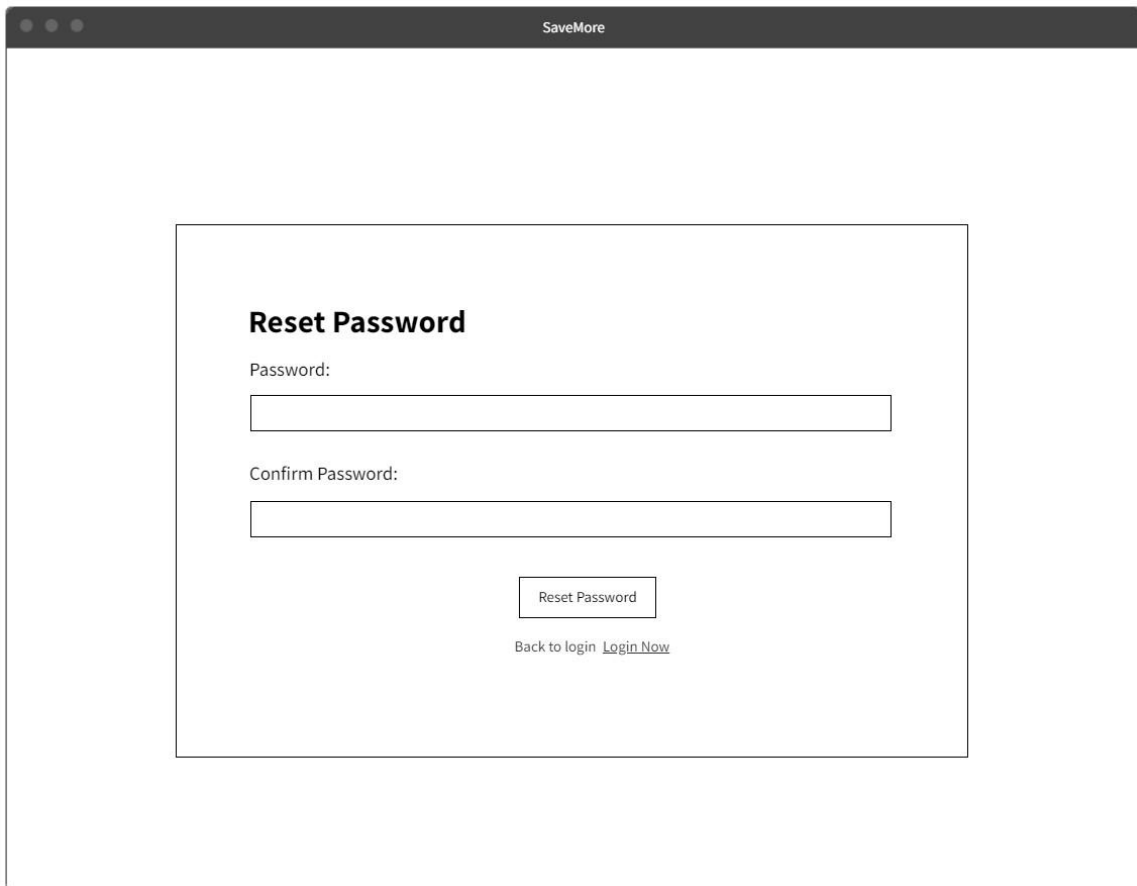


Figure 3.7.4: “reset password” user interface design (second stage to reset password).

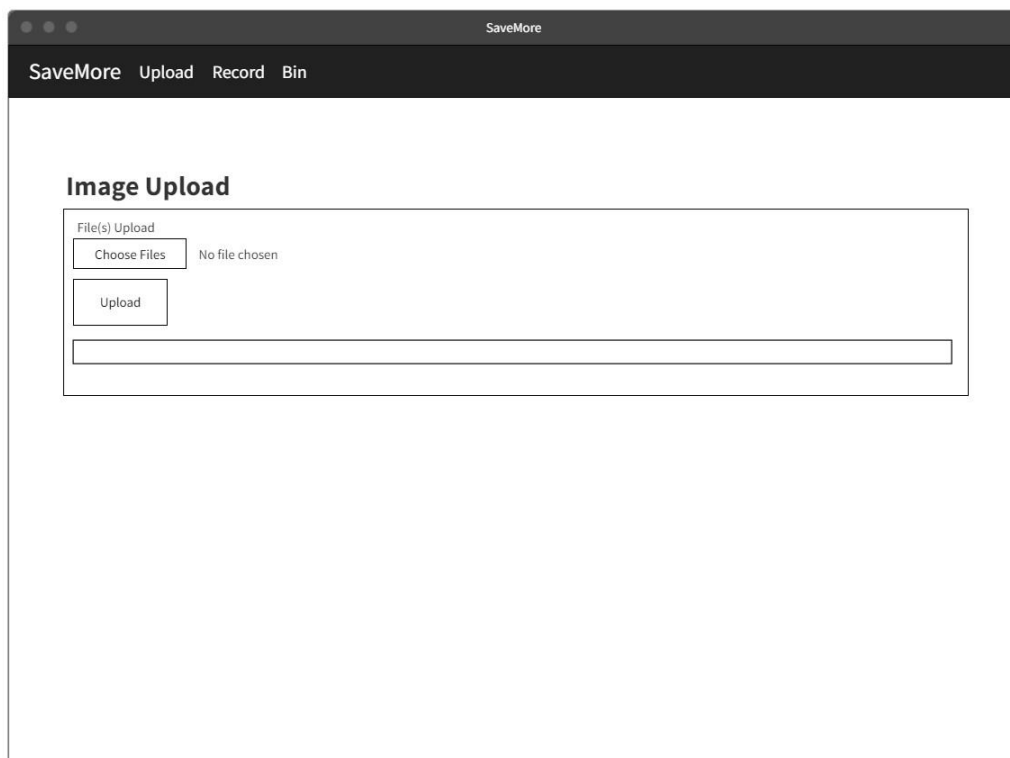


Figure 3.7.5: “upload” user interface design.

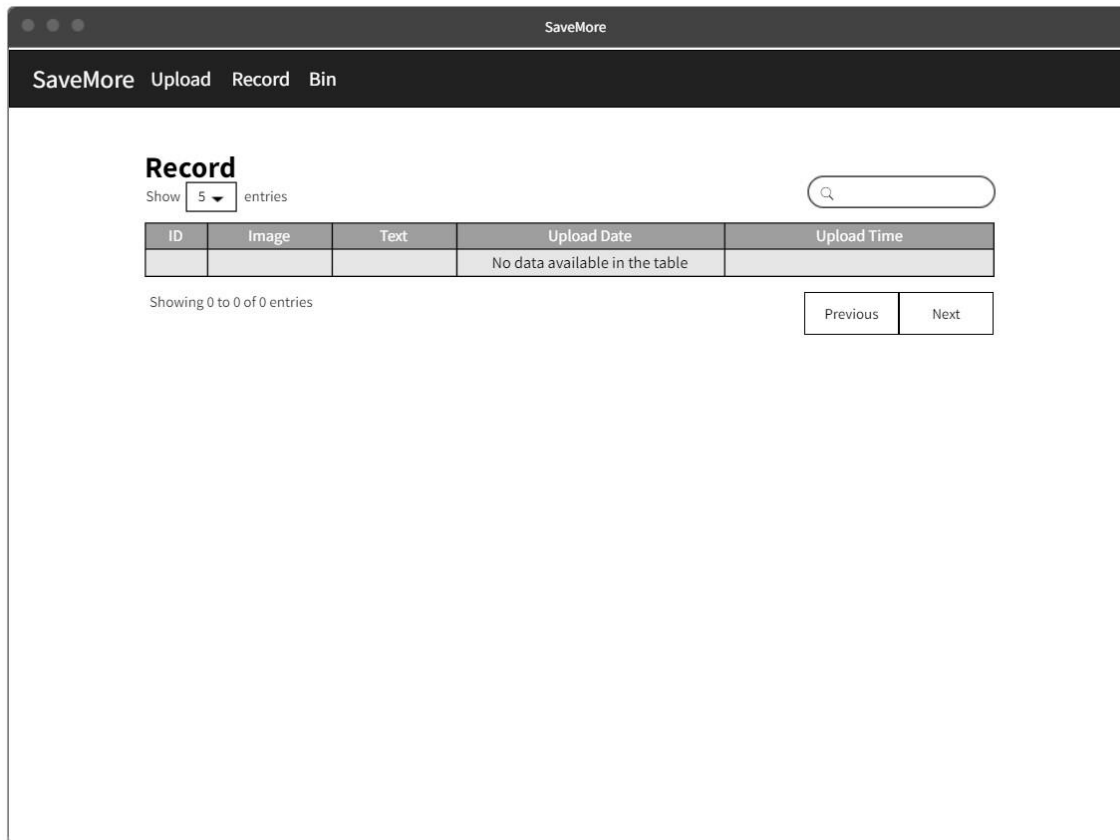


Figure 3.7.6: “uploaded records” user interface design.

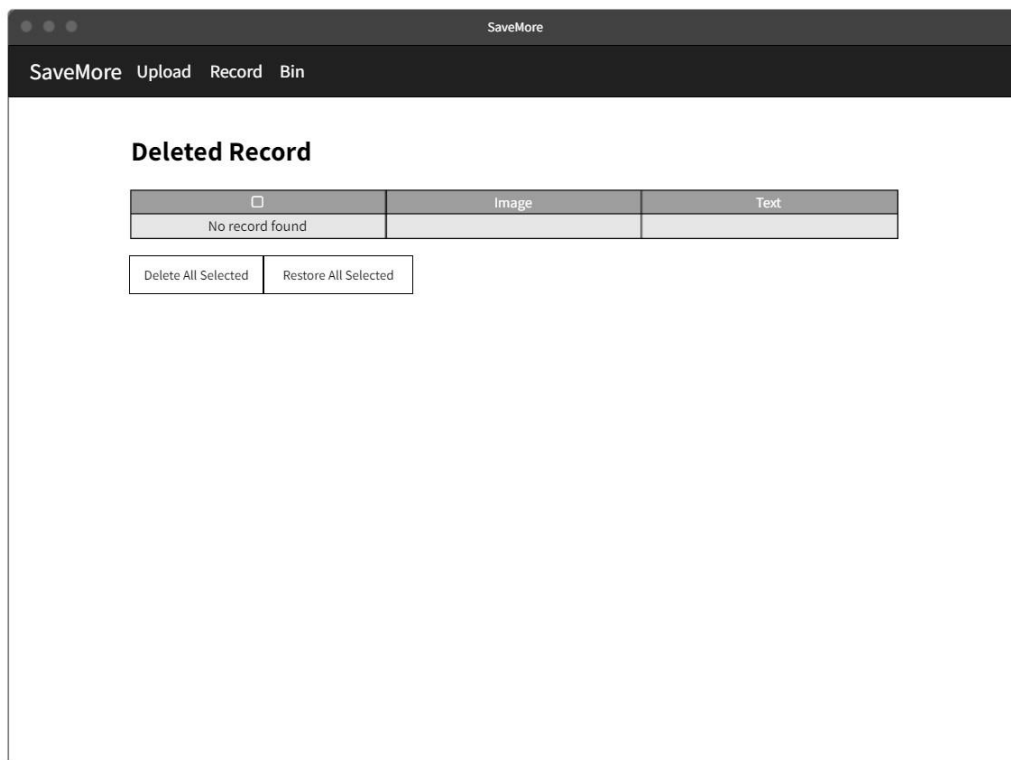


Figure 3.7.7: “deleted records” user interface design.

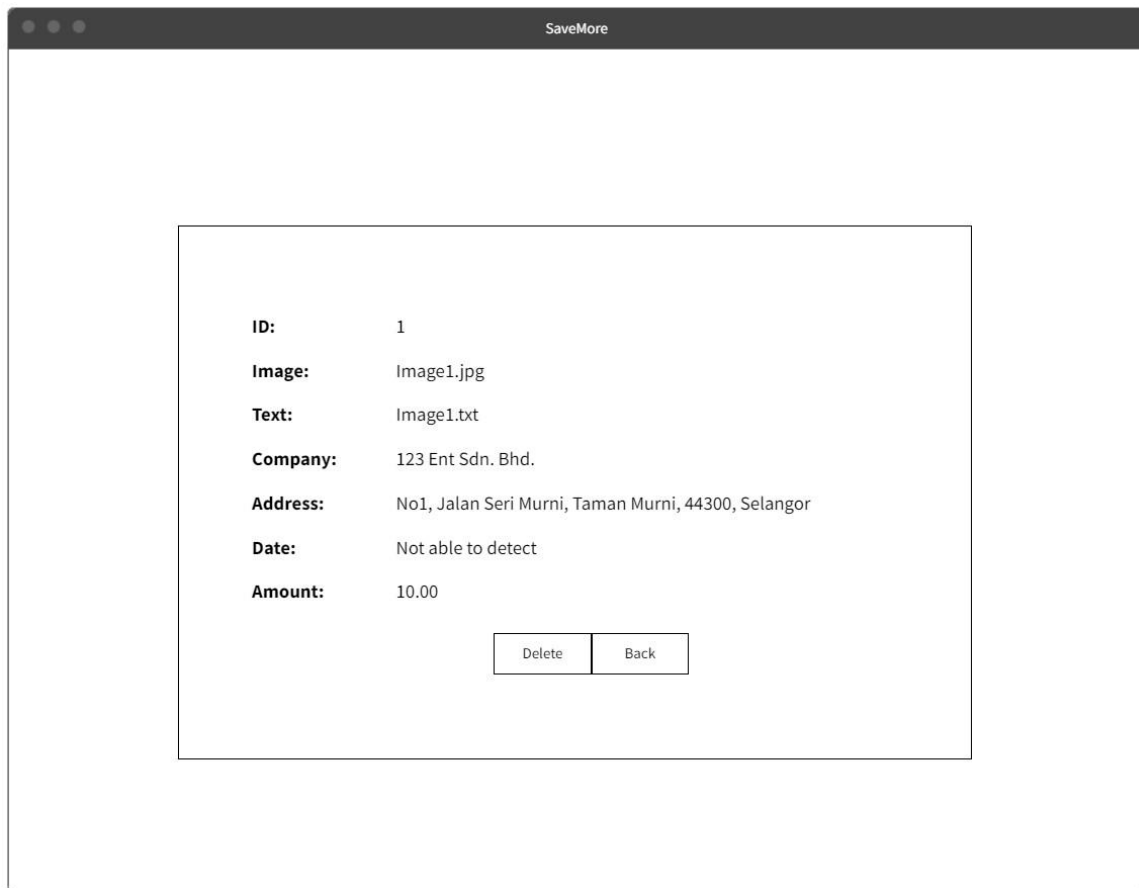


Figure 3.7.8: “record details” user interface design.

Chapter 4

System Implementation

4.1 Hardware Setup

Below are the specifications of the device used throughout this project to train the models and develop the web-based receipt key information extraction.

Description	Specifications
Model	HP Pavilion Gaming Laptop 15-cx0xxx
Processor	3.5 GHz Quad Core (with Hyper-Threading)
Operating System	Windows 11
Graphic	CUDA-Enabled Nvidia GTX1650Ti
Memory	20 GB
Storage	500 GB SSD Free Space

Table 4.1.1: Laptop specification.

The operating system used in this project must be Windows 11. Initially the operating system of my laptop is Windows 10, there is an error raised when running the imported source code of the models in Visual Studio Code. After upgrading from Windows 10 to Windows 11, the error has been eliminated.

Besides, the memory used in this project must be at least 17 GB and 20 GB listed in the table 4.1.1 is a relatively safe predictive memory. Initially the memory of my laptop is below 16 GB, there is an error raised when training the models and the training process has been terminated. After increasing the memory from below 16 GB to 20 GB, the model training process can be continued and the memory in use when initializing the model training process is between 16 GB and 17 GB.

4.2 Software Setup

4.2.1 Visual Studio Code with WSL

Visual Studio Code is the development tool used throughout this project. In order to train the models, Visual Studio Code with WSL environment is more suitable as the models are developed in a Linux environment.

4.2.2 Python

Bachelor of Information Systems (Honours) Business Information Systems
Faculty of Information and Communication Technology (Kampar Campus), UTAR

CHAPTER 4

Python is the programming language used throughout this project. The Python version used should be above or equal to 3.7 which is recommended by the developers who provide the source code of the developed models and below or equal to 3.9 in which when using Python 3.10 to train the models, some unexpected errors have been raised due to the incompatible Python version. In this case, the Python version is suggested from 3.7 to 3.9.

4.2.3 EasyOCR

EasyOCR is used as an API to connect both CTPN and Character-Aware Neural Language models by extracting the texts in the input image based on the coordinates of the texts extracted by CTPN model to be the input of Character-Aware Neural Language model to extract the keyword information of the input image.

4.2.4 Flask

Flask is a framework used to deploy the trained models and create the web-based receipt key information extraction system.

4.2.5 MySQL Server

A “kie” database with “user” and “record” tables have been created. There are 3 attributes in “user” table, namely user_id, email, and password. There are 11 attributes in “record” table, namely record_id, image, text, upload_date, upload_time, company, address, date, total, user_id, and status. The database is used to store the uploaded records by users.

4.2.6 Planning and Design Tools

i. draw.io

draw.io is used to draw all diagrams included in this project such as use case diagram, activity diagram, system block diagram, and system architecture diagram.

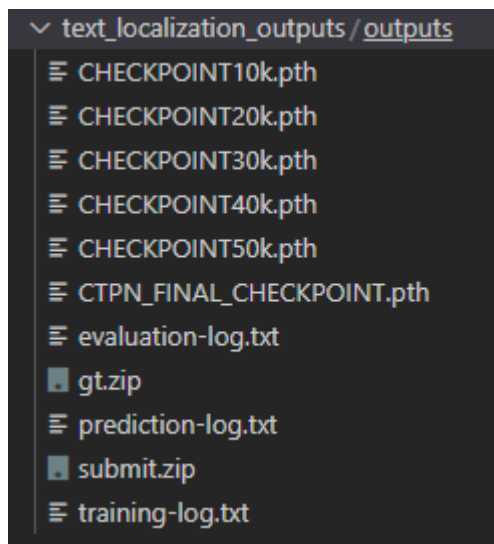
ii. MockFlow

MockFlow is used to design wireframe.

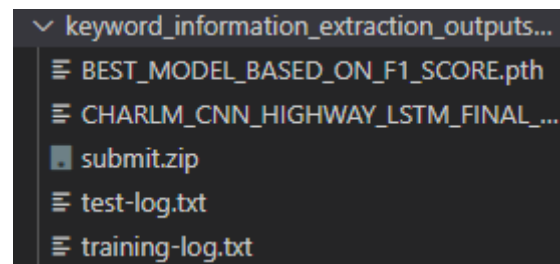
iii. Canva

Canva is used to design poster.

4.3 Models Training



a. Training and evaluation outputs for CTPN model.

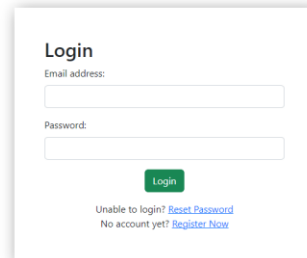


b. Training and evaluation outputs for Character-Aware Neural Language model.

Figure 4.3.1: Training and evaluation outputs for both Connectionist Text Proposal Network (CTPN) and Character-Aware Neural Language models.

After selecting the suitable models to perform text localization and keyword information extraction from a receipt which are CTPN and Character-Aware Neural Language models, the models are required to be trained by Ann Gee and me. After our analysis, Visual Studio Code with WSL environment is more suitable as the models are developed in a Linux environment. In the end, the training and evaluation outputs for both models are stored and displayed in the figure 4.3.1.

4.4 System Operation (with Screenshot)

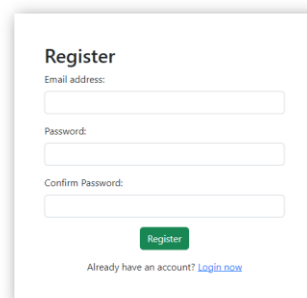


The screenshot shows a 'Login' form with the following elements:

- Title: **Login**
- Label: Email address:
- Input field: A white text box for the email address.
- Label: Password:
- Input field: A white text box for the password.
- Button: A green button labeled 'Login'.
- Links: Two blue links at the bottom: 'Unable to login? [Reset Password](#)' and 'No account yet? [Register now](#)'.

Figure 4.4.1: “login” interface.

When the user interacts with the system, “login” interface is the first interface the user will interact with.



The screenshot shows a 'Register' form with the following elements:

- Title: **Register**
- Label: Email address:
- Input field: A white text box for the email address.
- Label: Password:
- Input field: A white text box for the password.
- Label: Confirm Password:
- Input field: A white text box for the confirm password.
- Button: A green button labeled 'Register'.
- Link: A blue link at the bottom: 'Already have an account? [Login now](#)'.

Figure 4.4.2: “register” interface.

After the user selects to register now, the user is navigated to “register” interface.

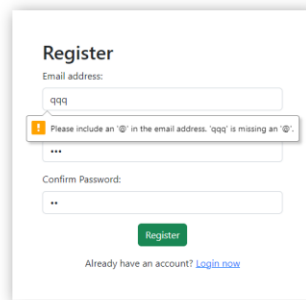


Figure 4.4.3: Invalid email format message.

After providing invalid email format and clicking on register button, invalid email format message is displayed.

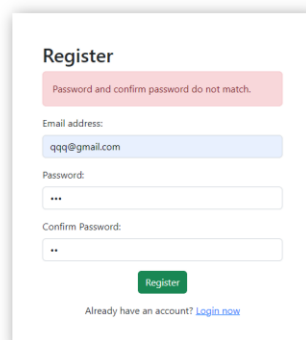


Figure 4.4.4: Password and confirm password do not match message.

After providing unmatched password and confirm password and clicking on register button, password and confirm password do not match message is displayed.

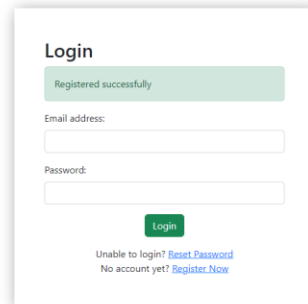


Figure 4.4.5: Successfully registered message.

After providing valid inputs in “register” interface, the user is navigated to “login” page and successfully registered message is displayed.

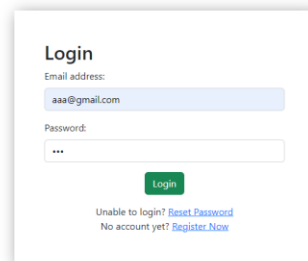


Figure 4.4.6: Login with invalid email or password.

The account has been registered with email “qqq@gmail.com”. The user logs in with invalid email or password.

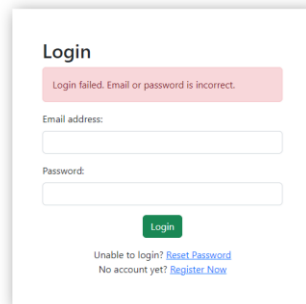


Figure 4.4.7: Invalid email or password message.

After providing invalid inputs and clicking on login button, invalid email or password message is displayed.

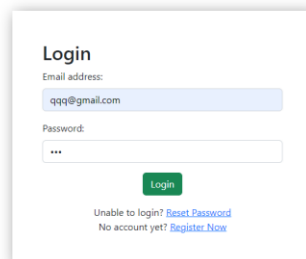


Figure 4.4.8: Login with valid email and password.

The user logs in with valid email and password.

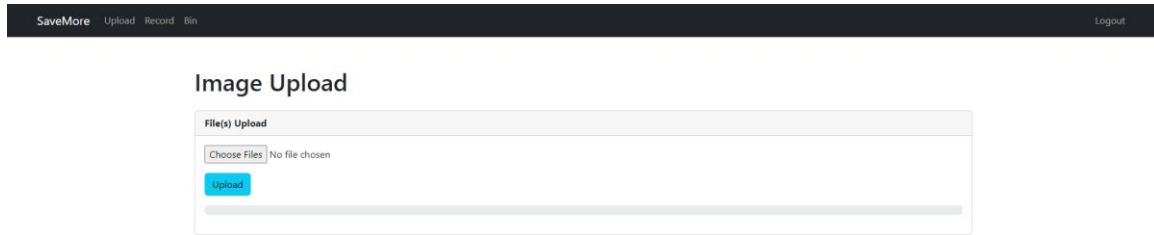


Figure 4.4.9: “upload” interface or home page.

After providing valid inputs in “login” interface, the user is navigated to “upload” interface or home page.

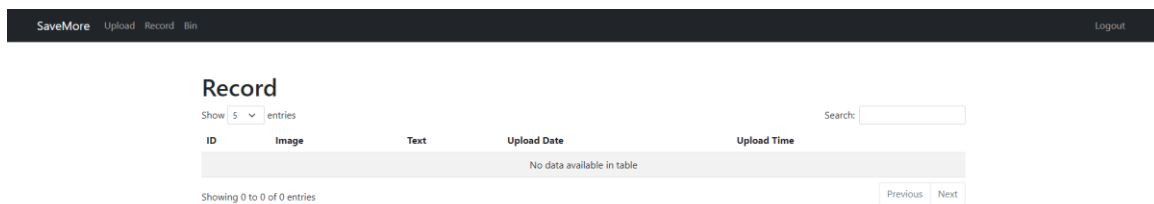


Figure 4.4.10: “uploaded records” page with no record.

The user has no uploaded record. No record found is displayed in “uploaded records” page.

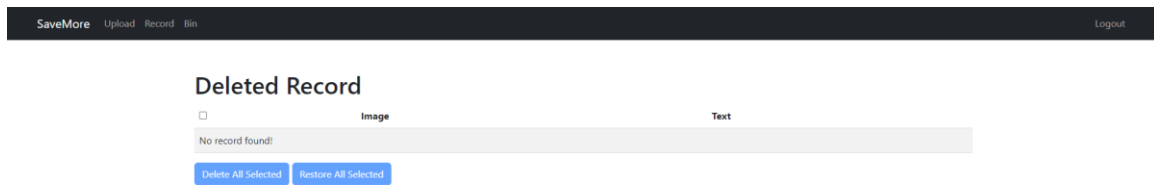


Figure 4.4.11: “deleted records” page with no record.

The user has no deleted record. No record found is displayed in “deleted records” page. The delete and restore buttons have been disabled.

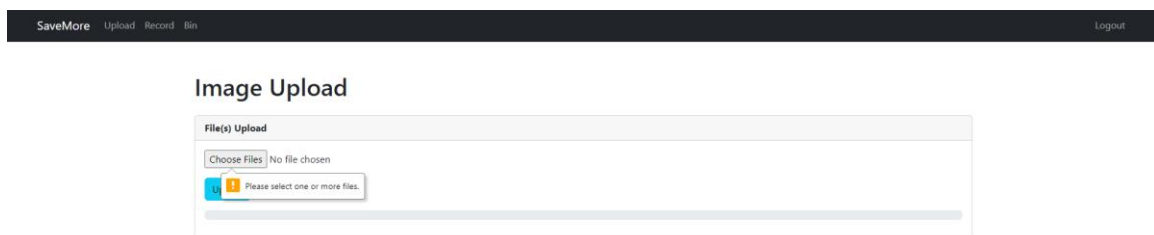


Figure 4.4.12: At least one file required to be uploaded message.

The user clicks on upload button with no file selected. At least one file required to be uploaded message is displayed.

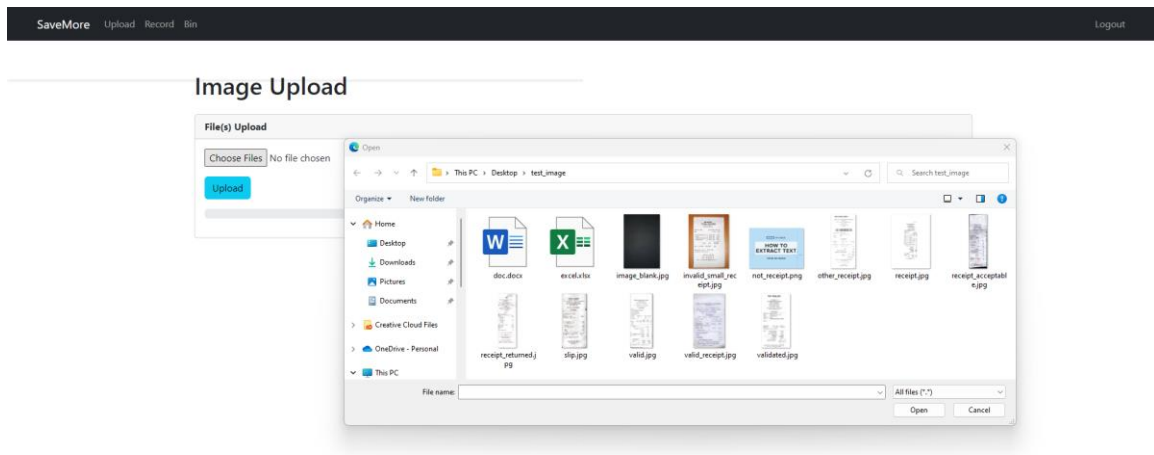


Figure 4.4.13: Files to be uploaded.

All files to be uploaded are displayed in figure 4.4.13.

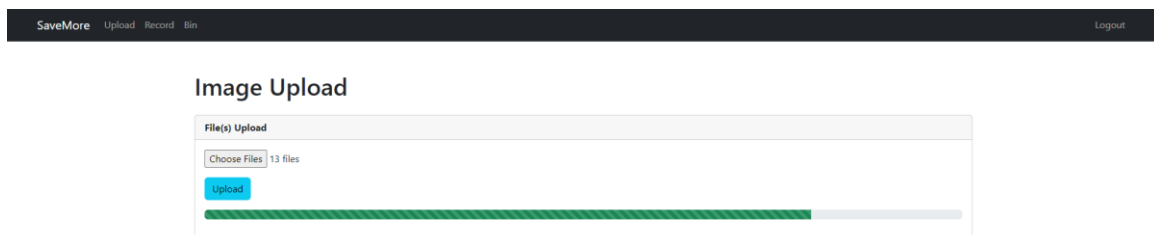


Figure 4.4.14: Progress bar when waiting for files to be processed.

After selecting files and clicking upload button to upload the files, the progress bar notifies the user of files being processed.

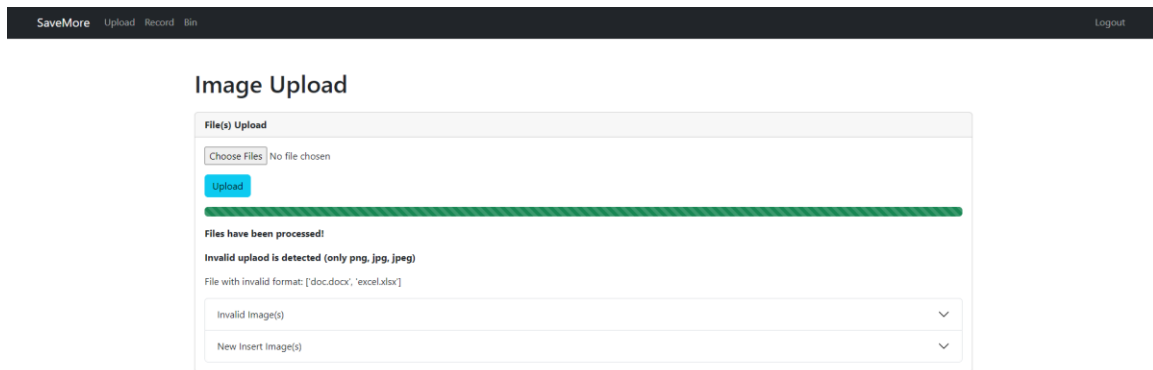


Figure 4.4.15: Results when files have been processed.

When the files have been processed, the progress bar is filled color and results are displayed as shown in figure 4.4.15.

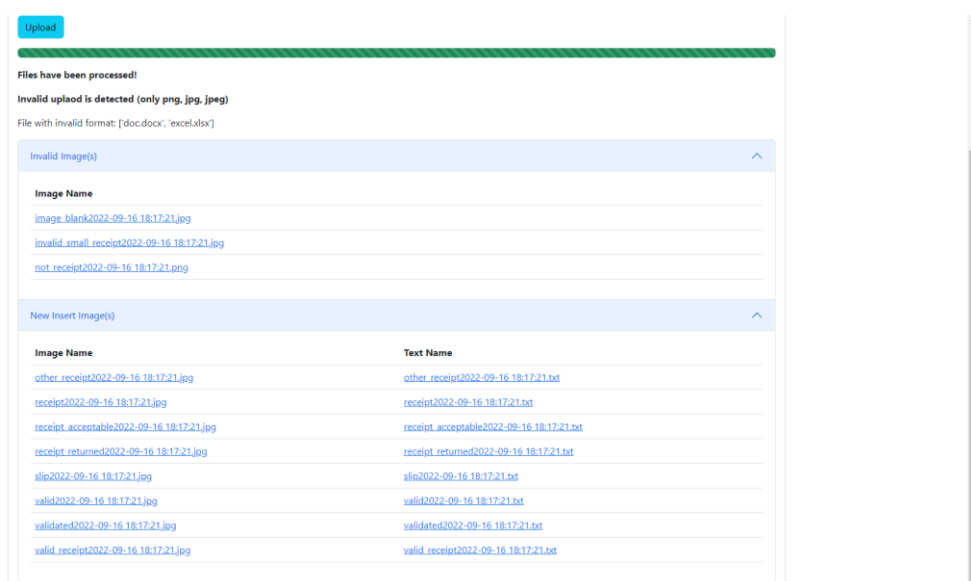


Figure 4.4.16: Expanded results.

The results can be expanded as shown in figure 4.4.16.

ID	Image	Text	Upload Date	Upload Time
1	other_receipt2022-09-16 18:17:21.jpg	other_receipt2022-09-16 18:17:21.txt	2022-09-16	18:17:21
2	receipt2022-09-16 18:17:21.jpg	receipt2022-09-16 18:17:21.txt	2022-09-16	18:17:21
3	receipt_acceptable2022-09-16 18:17:21.jpg	receipt_acceptable2022-09-16 18:17:21.txt	2022-09-16	18:17:21
4	receipt_returned2022-09-16 18:17:21.jpg	receipt_returned2022-09-16 18:17:21.txt	2022-09-16	18:17:21
5	slip2022-09-16 18:17:21.jpg	slip2022-09-16 18:17:21.txt	2022-09-16	18:17:21

Showing 1 to 5 of 8 entries

Figure 4.4.17: “uploaded records” page with records.

Eight records have been inserted successfully and displayed in “uploaded records” page. Option of five entries per page has been selected. Five records are displayed in the first page. The previous button has been disabled.

ID	Image	Text	Upload Date	Upload Time
6	valid2022-09-16 18:17:21.jpg	valid2022-09-16 18:17:21.txt	2022-09-16	18:17:21
7	validated2022-09-16 18:17:21.jpg	validated2022-09-16 18:17:21.txt	2022-09-16	18:17:21
8	valid_receipt2022-09-16 18:17:21.jpg	valid_receipt2022-09-16 18:17:21.txt	2022-09-16	18:17:21

Showing 6 to 8 of 8 entries

Figure 4.4.18: “uploaded records” page with records (cont.).

The remaining three records are displayed in the second page. The next button has been disabled.

ID	Image	Text	Upload Date	Upload Time
1	other_receipt2022-09-16 18:17:21.jpg	other_receipt2022-09-16 18:17:21.txt	2022-09-16	18:17:21
2	receipt2022-09-16 18:17:21.jpg	receipt2022-09-16 18:17:21.txt	2022-09-16	18:17:21
3	receipt_acceptable2022-09-16 18:17:21.jpg	receipt_acceptable2022-09-16 18:17:21.txt	2022-09-16	18:17:21
4	receipt_returned2022-09-16 18:17:21.jpg	receipt_returned2022-09-16 18:17:21.txt	2022-09-16	18:17:21
5	slip2022-09-16 18:17:21.jpg	slip2022-09-16 18:17:21.txt	2022-09-16	18:17:21
6	valid2022-09-16 18:17:21.jpg	valid2022-09-16 18:17:21.txt	2022-09-16	18:17:21
7	validated2022-09-16 18:17:21.jpg	validated2022-09-16 18:17:21.txt	2022-09-16	18:17:21
8	valid_receipt2022-09-16 18:17:21.jpg	valid_receipt2022-09-16 18:17:21.txt	2022-09-16	18:17:21

Showing 1 to 8 of 8 entries

Figure 4.4.19: “uploaded records” page with records (cont.).

Option of ten entries per page has been selected. All of the eight records are displayed in the first page. The previous and next buttons have been disabled.

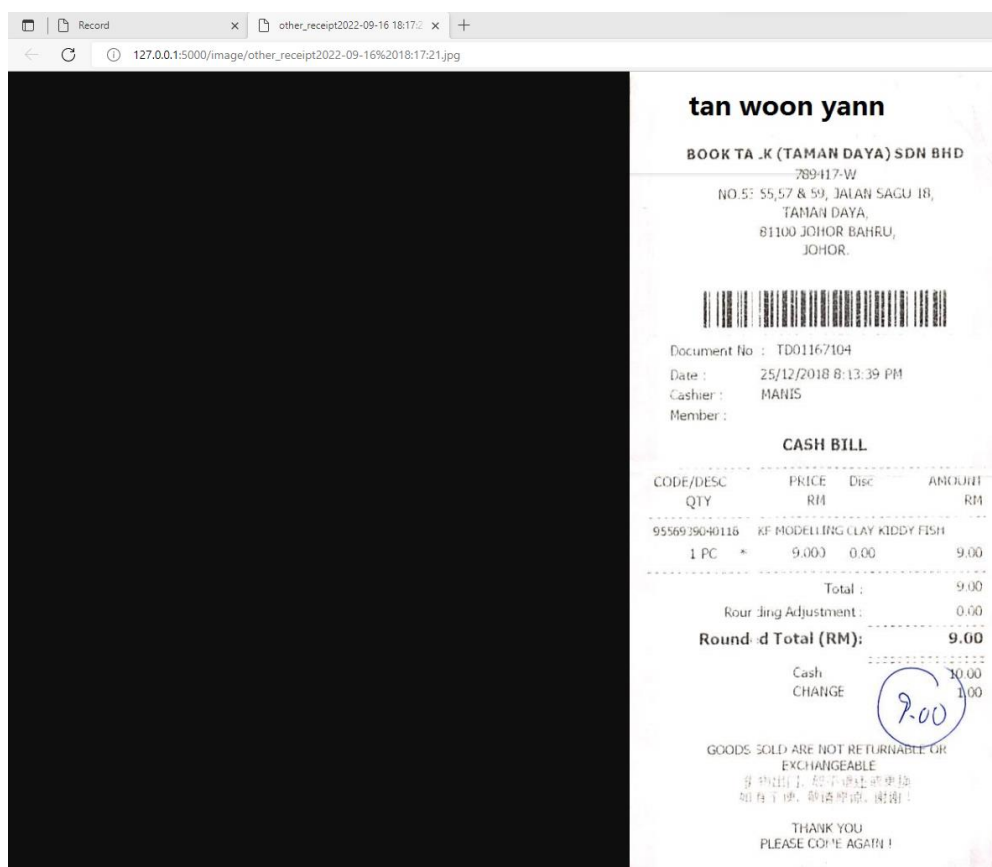


Figure 4.4.20: Uploaded image in new tab.

When clicking on the selected image URL, the corresponding image is displayed in new tab.

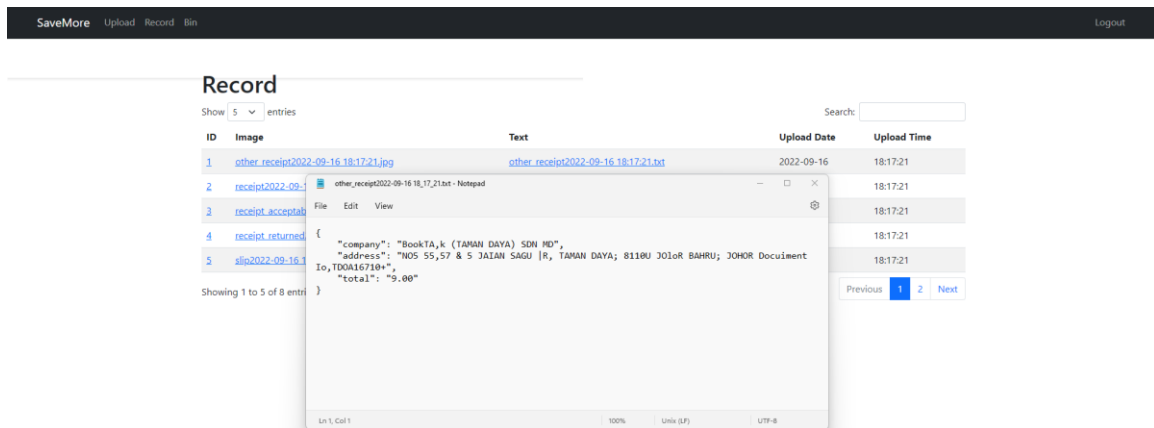


Figure 4.4.21: Extracted key information stored in text file.

The user has downloaded the text file when clicking on the text file URL. The extracted key information is stored in text file.

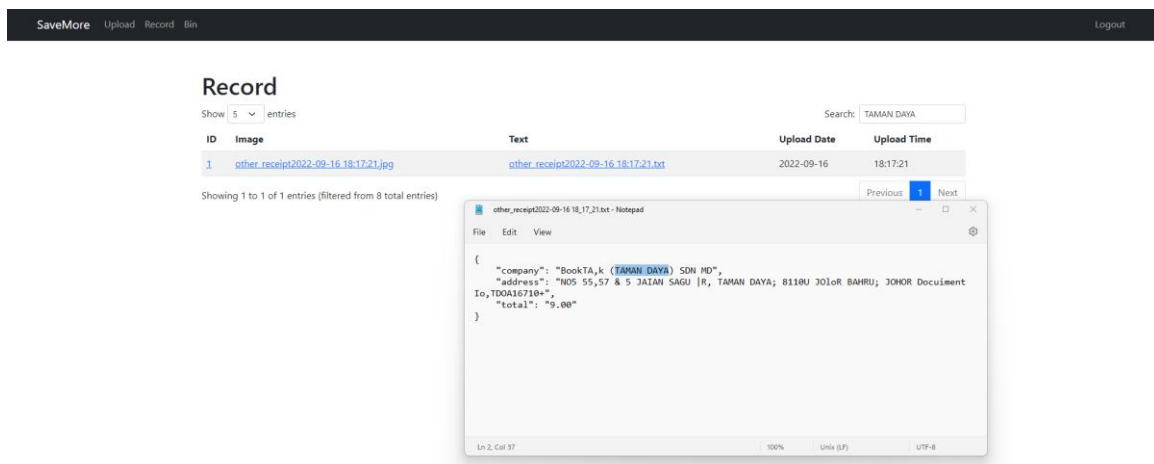


Figure 4.4.22: Search by company details.

The user can search records by company details.

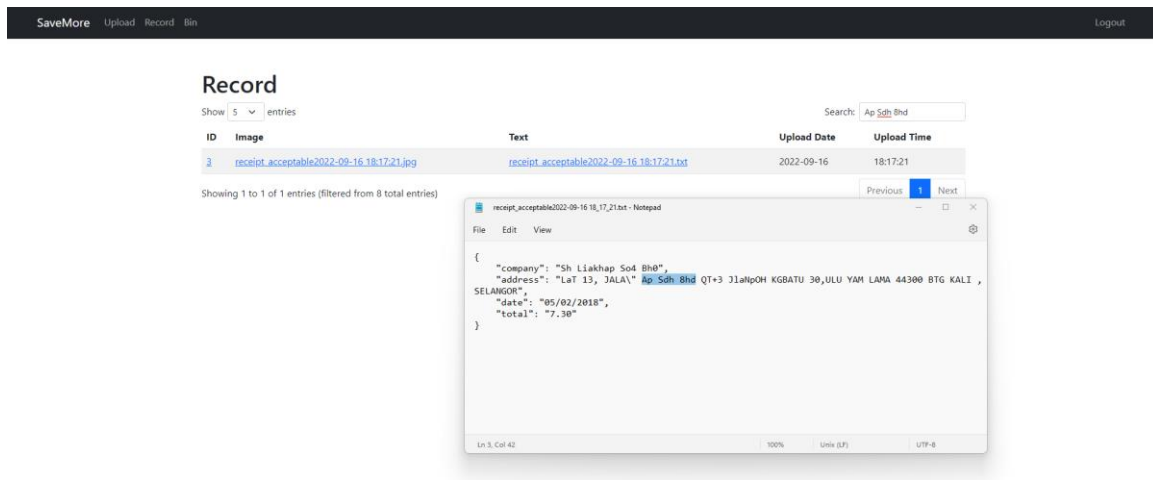


Figure 4.4.23: Search by address details.

The user can search records by address details.

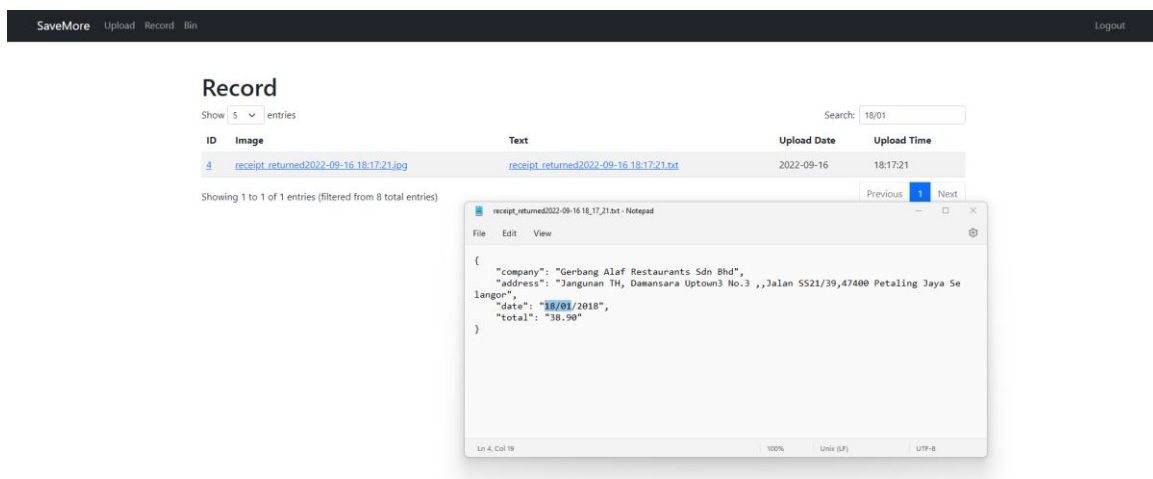


Figure 4.4.24: Search by date details.

The user can search records by date details.

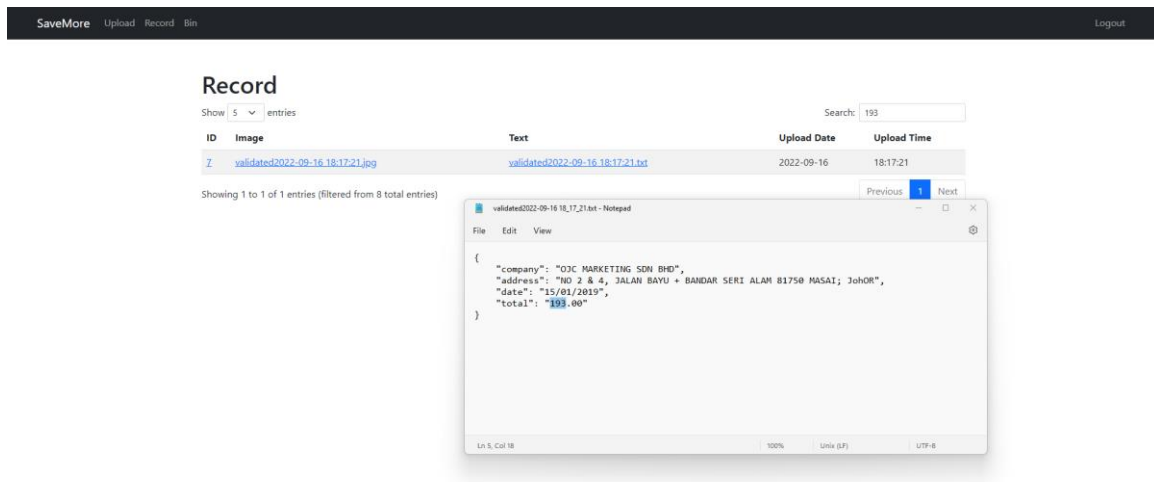


Figure 4.4.25: Search by total details.

The user can search records by total details.

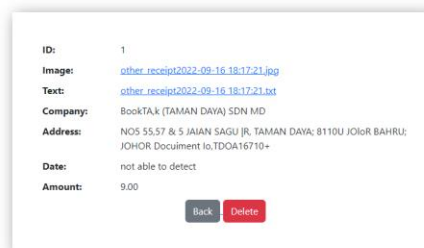


Figure 4.4.26: "record details" page.

The user is navigated to "record details" page of record ID 1 after clicking on the URL of record ID 1.

ID	Image	Text	Upload Date	Upload Time
2	receipt2022-09-16 18:17:21.jpg	receipt2022-09-16 18:17:21.txt	2022-09-16	18:17:21
3	receipt_acceptable2022-09-16 18:17:21.jpg	receipt_acceptable2022-09-16 18:17:21.txt	2022-09-16	18:17:21
4	receipt_returned2022-09-16 18:17:21.jpg	receipt_returned2022-09-16 18:17:21.txt	2022-09-16	18:17:21
5	slip2022-09-16 18:17:21.jpg	slip2022-09-16 18:17:21.txt	2022-09-16	18:17:21
6	valid2022-09-16 18:17:21.jpg	valid2022-09-16 18:17:21.txt	2022-09-16	18:17:21

Showing 1 to 5 of 7 entries

Figure 4.4.27: “uploaded records” page after record ID 1 has been deleted.

After clicking on delete button in “record details” page of record ID 1, the record with record ID 1 has been removed from “uploaded records” page.

<input type="checkbox"/>	Image	Text
<input type="checkbox"/>	other_receipt2022-09-16 18:17:21.jpg	other_receipt2022-09-16 18:17:21.txt

Delete All Selected | Restore All Selected

Figure 4.4.28: “deleted records” page after record ID 1 has been deleted.

After clicking on delete button in “record details” page of record ID 1, the record with record ID 1 is displayed in “deleted records” page.

ID	Image	Text	Upload Date	Upload Time
4	receipt_returned2022-09-16_18:17:21.jpg	receipt_returned2022-09-16_18:17:21.txt	2022-09-16	18:17:21
5	slip2022-09-16_18:17:21.jpg	slip2022-09-16_18:17:21.txt	2022-09-16	18:17:21
6	valid2022-09-16_18:17:21.jpg	valid2022-09-16_18:17:21.txt	2022-09-16	18:17:21
7	validated2022-09-16_18:17:21.jpg	validated2022-09-16_18:17:21.txt	2022-09-16	18:17:21
8	valid_receipt2022-09-16_18:17:21.jpg	valid_receipt2022-09-16_18:17:21.txt	2022-09-16	18:17:21

Figure 4.4.29: “uploaded records” page after deleting other records.

After deleting records with record ID 2 and 3, they have been removed from “uploaded records” page.

Image	Text
<input type="checkbox"/> receipt_acceptable2022-09-16_18:17:21.jpg	receipt_acceptable2022-09-16_18:17:21.txt
<input type="checkbox"/> receipt2022-09-16_18:17:21.jpg	receipt2022-09-16_18:17:21.txt
<input type="checkbox"/> other_receipt2022-09-16_18:17:21.jpg	other_receipt2022-09-16_18:17:21.txt

Figure 4.4.30: “deleted records” page after deleting other records.

After deleting records with record ID 2 and 3, they are displayed in “deleted records” page.

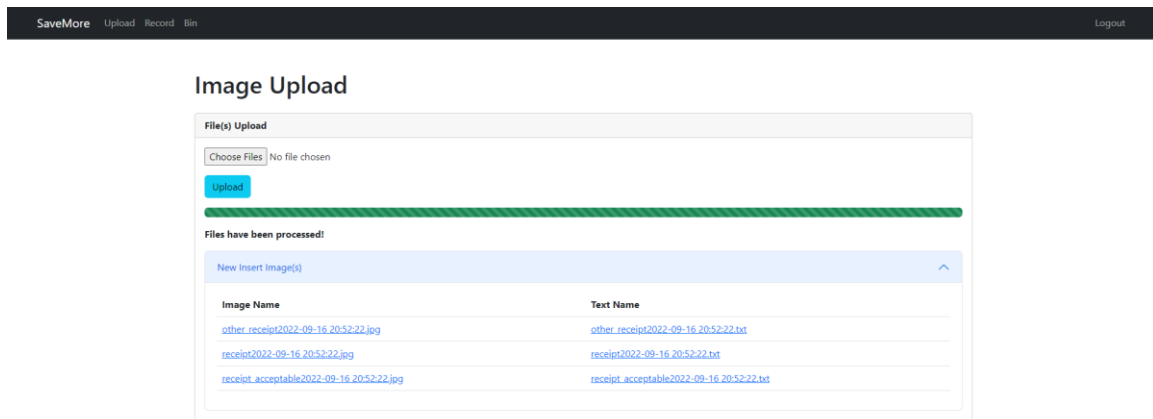


Figure 4.4.31: Upload the records have been deleted from “record details” page.

The records with record ID 1, 2, and 3 deleted previously and displayed in “deleted records” page are re-uploaded by the user. The records have been inserted successfully.

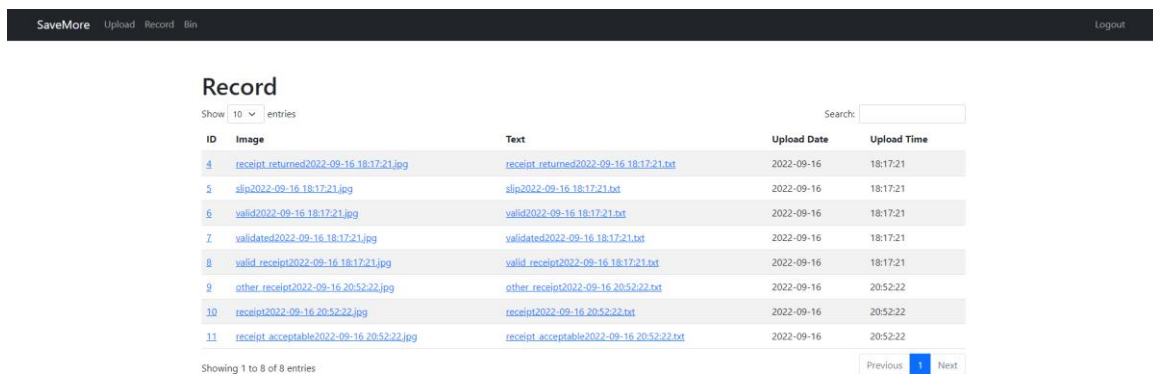


Figure 4.4.32: “uploaded records” page after uploading the records.

The records re-uploaded successfully are displayed in “uploaded records” page with record ID 9, 10, and 11.

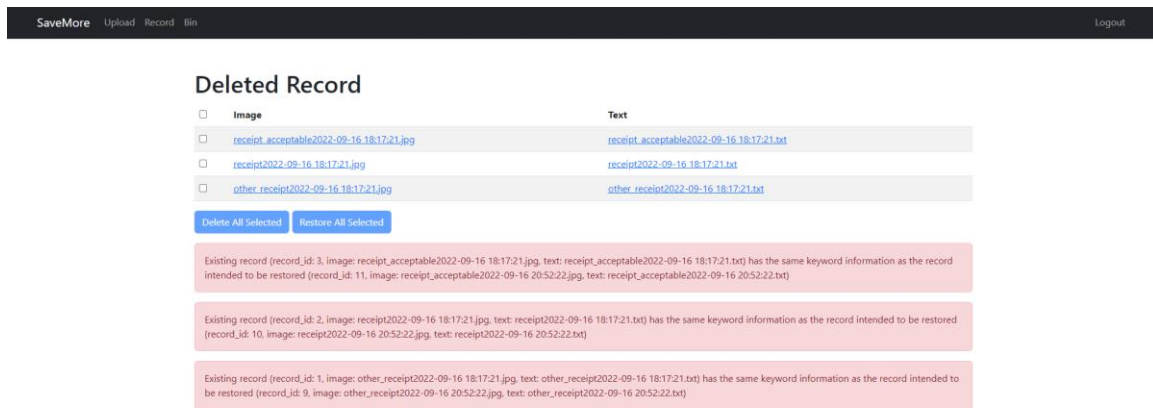


Figure 4.4.33: Record has not been restored message.

Three records displayed in “deleted records” page are the duplicate of records with record ID 9, 10, and 11. They are not restored successfully.

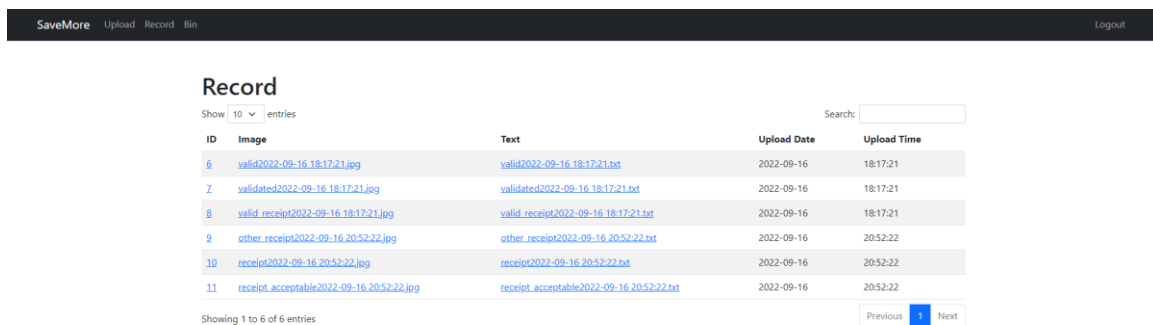


Figure 4.4.34: “uploaded records” page after deleting two records.

The records with record ID 4 and 5 have been deleted and removed from “uploaded records” page.

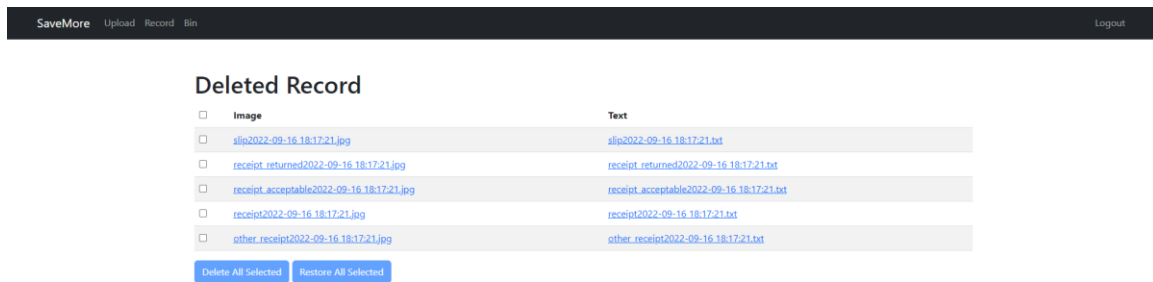


Figure 4.4.35: “deleted records” page after deleting two records.

The records with record ID 4 and 5 deleted previously are displayed in “deleted records” page.

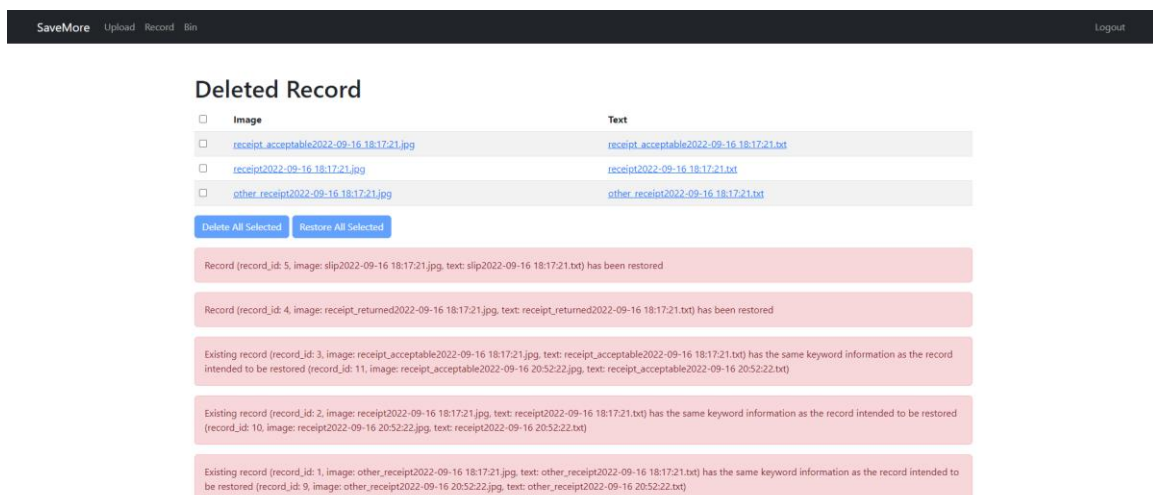


Figure 4.4.36: Successfully restored and record has not been restored messages.

When selecting all records to be restored, records with record ID 4 and 5 can be restored and removed from “deleted records” page whereas records with record ID 1, 2, and 3 are failed to restore.

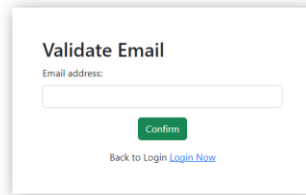
ID	Image	Text	Upload Date	Upload Time
4	receipt_returned2022-09-16 18:17:21.jpg	receipt_returned2022-09-16 18:17:21.txt	2022-09-16	18:17:21
5	slip2022-09-16 18:17:21.jpg	slip2022-09-16 18:17:21.txt	2022-09-16	18:17:21
6	valid2022-09-16 18:17:21.jpg	valid2022-09-16 18:17:21.txt	2022-09-16	18:17:21
7	validated2022-09-16 18:17:21.jpg	validated2022-09-16 18:17:21.txt	2022-09-16	18:17:21
8	valid_receipt2022-09-16 18:17:21.jpg	valid_receipt2022-09-16 18:17:21.txt	2022-09-16	18:17:21
9	other_receipt2022-09-16 20:52:22.jpg	other_receipt2022-09-16 20:52:22.txt	2022-09-16	20:52:22
10	receipt2022-09-16 20:52:22.jpg	receipt2022-09-16 20:52:22.txt	2022-09-16	20:52:22
11	receipt_acceptable2022-09-16 20:52:22.jpg	receipt_acceptable2022-09-16 20:52:22.txt	2022-09-16	20:52:22

Showing 1 to 8 of 8 entries

Figure 4.4.37: “uploaded records” page after two records have been restored.
 After restoring records with record ID 4 and 5, they are displayed in “uploaded records” page.

Image	Text
No record found!	
<input type="button" value="Delete All Selected"/> <input type="button" value="Restore All Selected"/>	
Successfully deleted	

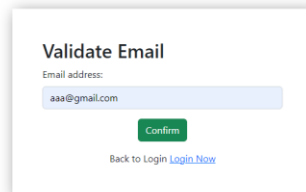
Figure 4.4.38: Successfully deleted message.
 After deleting all records in “deleted records” page, successfully deleted and no record found messages are displayed.



The screenshot shows a web form titled "Validate Email". Below the title, there is a label "Email address:" followed by an empty text input field. Below the input field is a green button with the text "Confirm". At the bottom of the form, there is a link that says "Back to Login [Login Now](#)".

Figure 4.4.39: “validate email” interface (first stage to reset password).

When the user selects to reset password due to the reason that the user forgot password, the user is navigated to “validate email” interface which is the first stage to reset password after clicking on “forgot password” URL.



The screenshot shows the same "Validate Email" form as in Figure 4.4.39. The text input field now contains the email address "aaa@gmail.com". The "Confirm" button and the "Back to Login [Login Now](#)" link are still present.

Figure 4.4.40: Reset password with invalid email.

The user resets password with invalid email.

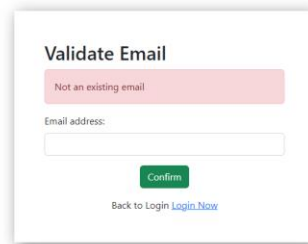


Figure 4.4.41: Not an existing email message.

Not an existing email message is displayed.

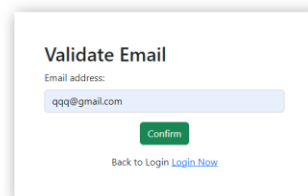
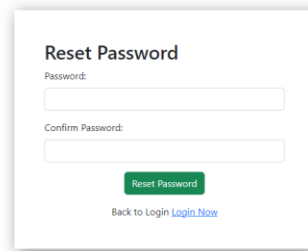


Figure 4.4.42: Reset password with valid email.

The user resets password with valid email.



Reset Password

Password:

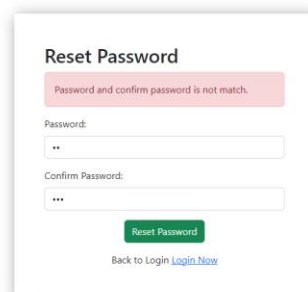
Confirm Password:

Reset Password

Back to Login [Login Now](#)

Figure 4.4.43: “reset password” interface (second stage to reset password).

The user is navigated to “reset password” interface which is the second stage to reset password after clicking on confirm button in “validate email” page and the provided email is valid.



Reset Password

Password and confirm password is not match.

Password:

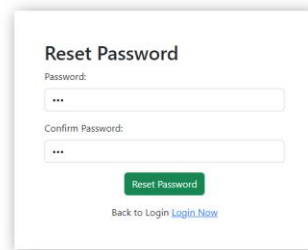
Confirm Password:

Reset Password

Back to Login [Login Now](#)

Figure 4.4.44: Password and confirm password do not match message.

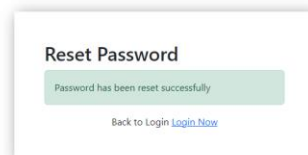
Password and confirm password do not match message is displayed when providing unmatched passwords.



The screenshot shows a 'Reset Password' form. It has a title 'Reset Password' at the top. Below the title, there are two input fields: 'Password:' and 'Confirm Password:'. Both fields contain three asterisks, indicating that the passwords are hidden. Below the input fields, there is a green button labeled 'Reset Password'. At the bottom of the form, there is a link that says 'Back to Login [Login Now](#)'.

Figure 4.4.45: Reset password with matched password and confirm password.

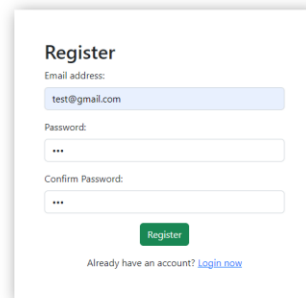
The user resets password with matched passwords.



The screenshot shows the 'Reset Password' form after a successful reset. The title 'Reset Password' is still at the top. Below the title, there is a green message box that says 'Password has been reset successfully'. Below the message box, there is a link that says 'Back to Login [Login Now](#)'.

Figure 4.4.46: Successfully reset password message.

Successfully reset password message is displayed when providing matched passwords.



Register

Email address:

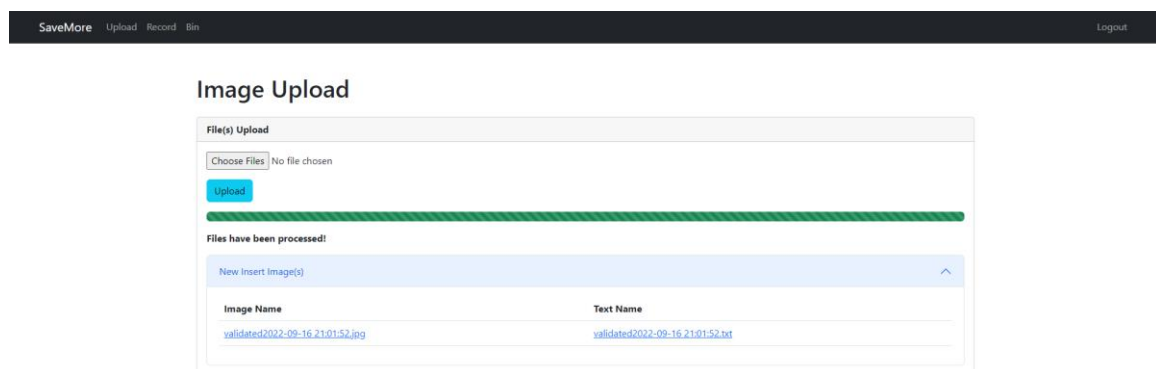
Password:

Confirm Password:

Already have an account? [Login now](#)

Figure 4.4.47: Register another new user account.

Another new user account has been created.



SaveMore Upload Record Bin Logout

Image Upload

File(s) Upload

Choose Files | No file chosen

Files have been processed!

New Insert Image(s)

Image Name	Text Name
validated2022-09-16 21:01:52.jpg	validated2022-09-16 21:01:52.txt

Figure 4.4.48: Expanded results when file has been processed.

The receipt uploaded and stored as a record in previous user account has been uploaded by the new user account. The receipt can be inserted successfully.

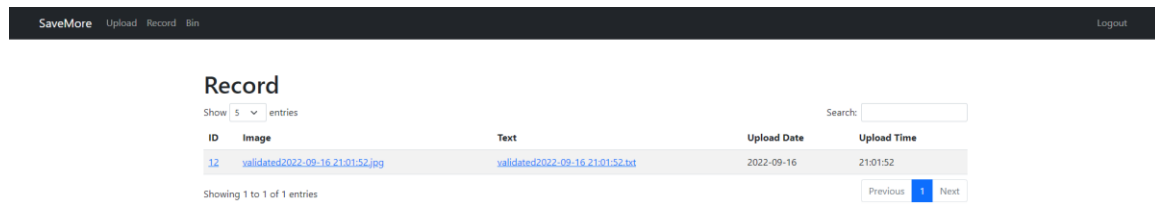


Figure 4.4.49: “uploaded records” page with records.

The “uploaded records” page will only display the record uploaded by the new user account as the new user account is the current logged in account. The records of the previous user account are not displayed in “uploaded records” page of the new user account.

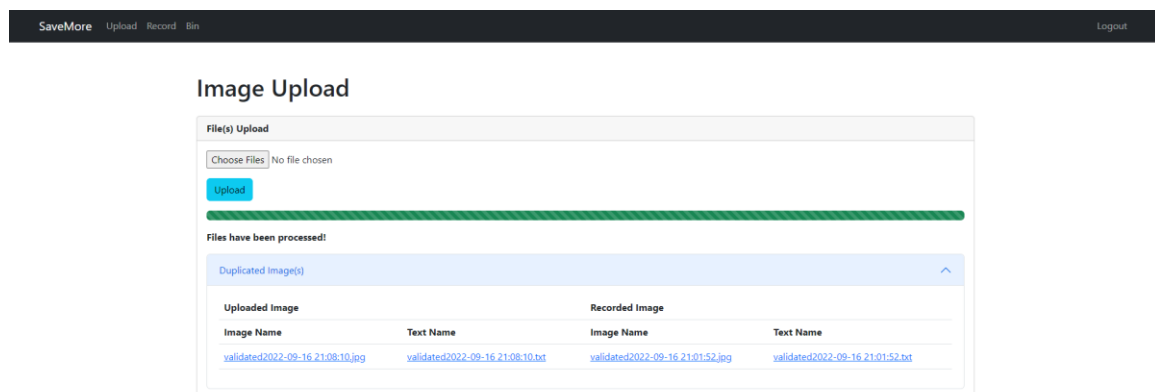


Figure 4.4.50: Expanded results after uploading duplicated image.

The user has uploaded the same receipt which has been stored as a record in “uploaded records” page and duplicate record has been detected. The duplicate record is not able to be stored.

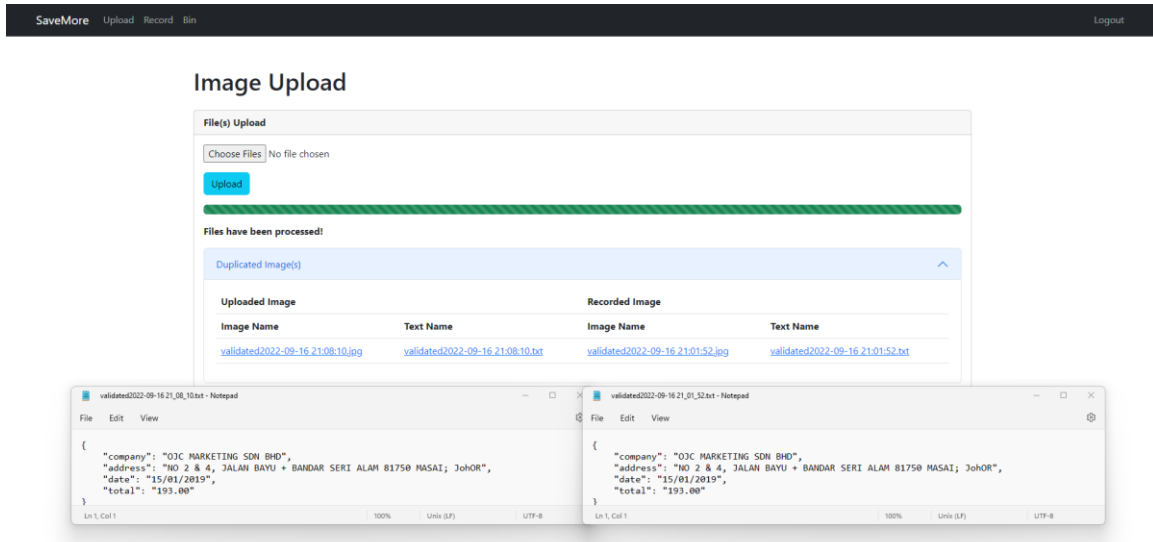


Figure 4.4.51: Text files for uploaded image and recorded image.

Duplicate records have been proved by opening both text files. It can be clearly observed in figure 4.4.51 that both text files have the same content.

user_id	email	password
1	qqq@gmail.com	\$\$\$rounds=535000\$JfcyvqXkEaJPTk0k\$FdvzJ37Vc5NatbGu...
2	test@gmail.com	\$\$\$rounds=535000\$HNR/Y/4qrljgJFSw\$0DkqK1RPDIh10hpq...

Figure 4.4.52: “user” table in “kie” database.

Records stored in “user” table and the passwords have been encrypted.

record_id	image	text	upload_date	upload_time	company	address	date	total	user_id	status
4	receipt_returned2022-09-16 18:17:21.jpg	receipt_returned2022-09-16 18:17:21.txt	2022-09-16	18:17:21	Gerbang Alaf Restaurants Sdn Bhd	Jangunan TH, Damansara Uptown3 No.3 ,Jalan SS21/3...	18/01/2018	38.90	1	1
5	slip2022-09-16 18:17:21.jpg	slip2022-09-16 18:17:21.txt	2022-09-16	18:17:21	PEIROH BKT LANJAN SB ALSERKAM ENTERPRISE	KM 458.4 BKT LANJAN UTARA _LRAYA UTARA SELATAN.SG...	01/02/2018	4.90	1	1
6	valid2022-09-16 18:17:21.jpg	valid2022-09-16 18:17:21.txt	2022-09-16	18:17:21	KEDAI PAPAN YEW CHUAN	LOT 276 JALAN BANTING 43800 DENGKIL , SELANGOR	17/04/2018	190.80	1	1
7	validated2022-09-16 18:17:21.jpg	validated2022-09-16 18:17:21.txt	2022-09-16	18:17:21	OJC MARKETING SDN BHD	NO 2 & 4, JALAN BAYU + BANDAR SERI ALAM 81750 MASA...	15/01/2019	193.00	1	1
8	valid_receipt2022-09-16 18:17:21.jpg	valid_receipt2022-09-16 18:17:21.txt	2022-09-16	18:17:21	SWC ENTERPRISE SDN BHD	NIO . -57, Jalan Mahagoni, 7/1, Sekyzen, 4, Bandar, Ut...	not able to detect	8.00	1	1
9	other_receipt2022-09-16 20:52:22.jpg	other_receipt2022-09-16 20:52:22.txt	2022-09-16	20:52:22	BookTA,k (TAMAN DAYA) SDN MD	NO5 55,57 & 5 JAIAN SAGU JR, TAMAN DAYA, 8110U JOI...	not able to detect	9.00	1	1
10	receipt2022-09-16 20:52:22.jpg	receipt2022-09-16 20:52:22.txt	2022-09-16	20:52:22	PERNIAGAAN ZHENG HUI	NO.59 JALAN PERMAS 915 BANDAR BARU PERMAS JAYA 3AH...	not able to detect	not able to detect	1	1
11	receipt_acceptable2022-09-16 20:52:22.jpg	receipt_acceptable2022-09-16 20:52:22.txt	2022-09-16	20:52:22	Sh Liakhap So4 Bh0	LaT 13, JALA' Ap Sdh 8hd QT+3 ,JlaNpOH KGBATU 30, UL...	05/02/2018	7.30	1	1
12	validated2022-09-16 21:01:52.jpg	validated2022-09-16 21:01:52.txt	2022-09-16	21:01:52	OJC MARKETING SDN BHD	NO 2 & 4, JALAN BAYU + BANDAR SERI ALAM 81750 MASA...	15/01/2019	193.00	2	1

Figure 4.4.53: “record” table in “kie” database.

Records stored in “record” table. Status 1 is for the activated records which will be displayed in “uploaded records” page whereas status 2 is for the deactivated records which will be displayed in “deleted records” page.

4.5 Implementation Issues and Challenges

When importing the source code of the models for training in Visual Studio Code, numerous lines of errors have been raised. Initially, the errors are difficult to be removed. After eliminating one error, new errors are raised continuously. In the end, majority errors have been eliminated after changing the environment of Visual Studio Code to Linux environment.

4.6 Concluding Remark

In the end, the software used in this project are Visual Studio Code with WSL as the development tool, Python with version in between 3.7 and 3.9, EasyOCR as an API, Flask as a framework to deploy the models, MySQL server as the database, and planning and design tools such as draw.io, MockFlow, and Canva.

Chapter 5

System Evaluation and Discussion

5.1 System Testing and Performance Metrics

Black box testing is a testing performed by the testers who are not involved throughout the entire project, have no knowledge on the project details such as the system architecture, database design, etc. In this project, use case testing which is a functional black box testing approach has been adopted to validate the integrations between the actor which is the user and the system which is the web-based receipt key information extraction system.

5.2 Testing Setup and Result

Use Case	Register		
Test Case	Expected Output	Actual Output	Result
User registers with valid inputs.	User account will be created in database. Password will be encrypted. User will be navigated to “login” page. Successfully registered message will be displayed in “login” page.	User account is created in database. Password is encrypted. User is navigated to “login” page. Successfully registered message is displayed in “login” page.	Pass
User registers with empty fields.	Field required to be filled out message will be displayed.	Field required to be filled out message is displayed.	Pass
User registers with invalid email format.	Invalid email format message will be displayed.	Invalid email format message is displayed.	Pass
User registers with unmatched password	Password and confirm password do	Password and confirm password do	Pass

and confirm passwords.	not match message will be displayed.	not match message is displayed.	
User clicks on “login” URL.	User will be navigated to “login” page.	User is navigated to “login” page.	Pass

Table 5.2.1: Register testing.

Use Case	Login		
Test Case	Expected Output	Actual Output	Result
User logs in with valid inputs.	User will be navigated to “upload” page which is the home page of the system.	User is navigated to “upload” page which is the home page of the system.	Pass
User logs in with empty fields.	Field required to be filled out message will be displayed.	Field required to be filled out message is displayed.	Pass
User logs in with invalid email format.	Invalid email format message will be displayed.	Invalid email format message is displayed.	Pass
User logs in with invalid email.	Invalid email or password message will be displayed.	Invalid email or password message is displayed.	Pass
User logs in with invalid password.	Invalid email or password message will be displayed.	Invalid email or password message is displayed.	Pass
User clicks on “register” URL.	User will be navigated to “register” page.	User is navigated to “register” page.	Pass
User clicks on “forgot password” URL.	User will be navigated to “forgot password” page.	User is navigated to “forgot password” page.	Pass

Table 5.2.2: Login testing.

Use Case	Logout		
Test Case	Expected Output	Actual Output	Result
User clicks on “logout” option.	User will be navigated to “login” page.	User is navigated to “login” page.	Pass

Table 5.2.3: Logout testing.

Use Case	Reset password		
Test Case	Expected Output	Actual Output	Result
In the page to validate email:			
User resets password with valid email.	User will be navigated to next stage to reset password.	User is navigated to next stage to reset password.	Pass
User resets password with empty field.	Field required to be filled out message will be displayed.	Field required to be filled out message is displayed.	Pass
User resets password with invalid email format.	Invalid email format message will be displayed.	Invalid email format message is displayed.	Pass
User resets password with invalid email.	Not a registered email message will be displayed.	Not a registered email message will be displayed.	Pass
User clicks on “login” URL.	User will be navigated to “login” page.	User is navigated to “login” page.	Pass
In the page to reset password:			
User resets password with matched password and confirm password.	Password of user will be updated in database. Password will be encrypted. Successfully reset	Password of user is updated in database. Password is encrypted. Successfully reset	Pass

	password message and “login” URL will be displayed.	password message and “login” URL is displayed.	
User resets password with empty fields.	Field required to be filled out message will be displayed.	Field required to be filled out message is displayed.	Pass
User resets password with unmatched password and confirm password.	Password and confirm password do not match message will be displayed.	Password and confirm password do not match message is displayed.	Pass
User clicks on “login” URL.	User will be navigated to “login” page.	User is navigated to “login” page.	Pass
After successfully password reset:			
User clicks on “login” URL.	User will be navigated to “login” page.	User is navigated to “login” page.	Pass

Table 5.2.4: Reset password testing.

Use Case	Upload files		
Test Case	Expected Output	Actual Output	Result
User clicks on “upload” button with no file selected.	At least one file required to be uploaded message will be displayed.	At least one file required to be uploaded message is displayed.	Pass
User uploads one file with invalid file format.	Invalid file format message will be displayed. The file name with invalid file format will be displayed.	Invalid file format message is displayed. The file name with invalid file format is displayed.	Pass

User uploads one English receipt large in size.	Files have been processed message will be displayed. Image and text file URLs will be displayed in new insert image category. Record will be stored in database.	Files have been processed message is displayed. Image and text file URLs are displayed in new insert image category. Record is stored in database.	Pass
User uploads multiple English receipts large in size.	Files have been processed message will be displayed. Image and text file URLs will be displayed in new insert image category. Record will be stored in database.	Files have been processed message is displayed. Image and text file URLs are displayed in new insert image category. Record is stored in database.	Pass
User uploads one image with no text.	Files have been processed message will be displayed. Image URL will be displayed in invalid image category.	Files have been processed message is displayed. Image URL is displayed in invalid image category.	Pass
User uploads one non-receipt image with texts.	Files have been processed message will be displayed. Image URL will be displayed in invalid image category.	Files have been processed message is displayed. Image URL is displayed in invalid image category.	Pass

User uploads one English receipt small in size.	Files have been processed message will be displayed. Image URL will be displayed in invalid image category.	Files have been processed message is displayed. Image URL is displayed in invalid image category.	Pass
User uploads one non-English receipt.	Files have been processed message will be displayed. Image URL will be displayed in invalid image category.	Files have been processed message is displayed. Image URL is displayed in invalid image category.	Pass
User uploads one receipt uploaded and stored in “records” page.	Files have been processed message will be displayed. Image and text file URLs will be displayed in duplicate image category.	Files have been processed message is displayed. Image and text file URLs are displayed in duplicate image category.	Pass
User uploads one receipt deleted temporarily and stored in “deleted records” page.	Files have been processed message will be displayed. Image and text file URLs will be displayed in new insert image category. Record will be stored in database.	Files have been processed message is displayed. Image and text file URLs are displayed in new insert image category. Record is stored in database.	Pass
Images such as image with no text, non-receipt image with texts, English receipt small in size, non-English receipt are considered as unacceptable receipts.			

<p>User uploads multiple combinations of one file with invalid file format, one English receipt large in size, one unacceptable receipt, one receipt uploaded and stored in “records” page, and one receipt deleted temporarily and stored in “deleted records” page.</p>	<p>First possible output: Files have been processed message will be displayed. The results will be displayed in their respective categories. Valid records will be stored in database.</p> <p>Second possible output: Invalid file format message will be displayed. The file name with invalid file format will be displayed. Files have been processed message will be displayed. The results will be displayed in their respective categories. Valid records will be stored in database.</p>	<p>First possible output: Files have been processed message is displayed. The results are displayed in their respective categories. Valid records are stored in database.</p> <p>Second possible output: Invalid file format message is displayed. The file name with invalid file format is displayed. Files have been processed message is displayed. The results are displayed in their respective categories. Valid records are stored in database.</p>	<p>Pass</p>
<p>User clicks on image URL.</p>	<p>Image will be displayed in a new tab.</p>	<p>Image is displayed in a new tab.</p>	<p>Pass</p>
<p>User clicks on text file URL.</p>	<p>User will be able to download the text file.</p>	<p>User is able to download the text file.</p>	<p>Pass</p>

Table 5.2.5: Upload files testing.

Use Case	View uploaded images		
Test Case	Expected Output	Actual Output	Result
User clicks on image URL.	Image will be displayed in a new tab.	Image is displayed in a new tab.	Pass

Table 5.2.6: View uploaded images testing.

Use Case	Download text files		
Test Case	Expected Output	Actual Output	Result
User clicks on text file URL.	User will be able to download the text file.	User is able to download the text file.	Pass

Table 5.2.7: Download text files testing.

Use Case	View uploaded records		
Test Case	Expected Output	Actual Output	Result
User has no uploaded record.	No record found message will be displayed.	No record found message is displayed.	Pass
User has uploaded records.	Records of user will be displayed.	Records of user are displayed.	Pass
User clicks on “record id” URL.	User will be navigated to the corresponding “record details” page.	User is navigated to the corresponding “record details” page.	Pass
User clicks on image URL.	Image will be displayed in a new tab.	Image is displayed in a new tab.	Pass
User clicks on text file URL.	User will be able to download the text file.	User is able to download the text file.	Pass

CHAPTER 5

User provides valid keyword.	Related records of user will be displayed.	Related records of user are displayed.	Pass
User provides invalid keyword.	No record found message will be displayed.	No record found message is displayed.	Pass
User clicks on next button (provided there are records in next page).	Records in next page will be displayed.	Records in next page are displayed.	Pass
User clicks on next page number (provided there are records in next page).	Records in next page will be displayed.	Records in next page are displayed.	Pass
User clicks on previous button (provided there are records in previous page).	Records in previous page will be displayed.	Records in previous page are displayed.	Pass
User clicks on previous page number (provided there are records in previous page).	Records in previous page will be displayed.	Records in previous page are displayed.	Pass
User selects option of 5 entries per page (provided the total number of records is 4).	All of the four records will be displayed in the first page. The previous and next buttons will be disabled.	All of the four records are displayed in the first page. The previous and next buttons are disabled.	Pass
User selects option of 10 entries per	First ten records will be displayed in the	First ten records are displayed in the first	Pass

page (provided the total number of records is 25).	first page. Three pages will be used to display all records. Last five records will be displayed in the third page.	page. Three pages are used to display all records. Last five records are displayed in the third page.	
--	---	---	--

Table 5.2.8: View uploaded records testing.

Use Case	View record details		
Test Case	Expected Output	Actual Output	Result
User is navigated to “record details” page of a record id.	The details of the record id will be displayed.	The details of the record id are displayed.	Pass
User clicks on back button.	User will be navigated to “uploaded records” page.	User is navigated to “uploaded records” page.	Pass
User clicks on image URL.	Image will be displayed in a new tab.	Image is displayed in a new tab.	Pass
User clicks on text file URL.	User will be able to download the text file.	User is able to download the text file.	Pass

Table 5.2.9: View record details testing.

Use Case	View deleted records		
Test Case	Expected Output	Actual Output	Result
User has no deleted record.	No record found message will be displayed. Delete and restore buttons will be disabled.	No record found message is displayed. Delete and restore buttons are disabled.	Pass

CHAPTER 5

User has deleted records.	Deactivated records of user will be displayed.	Deactivated records of user are displayed.	Pass
User clicks on image URL.	Image will be displayed in a new tab.	Image is displayed in a new tab.	Pass
User clicks on text file URL.	User will be able to download the text file.	User is able to download the text file.	Pass
User checks “check all” checkbox (provided there are three deleted records).	All checkboxes will be checked.	All checkboxes are checked.	Pass
User unchecks “check all” checkbox (provided there are three deleted records).	All checkboxes will be unchecked.	All checkboxes are unchecked.	Pass
User checks one checkbox (provided there are three deleted records).	Only the checkbox checked by user will be checked.	Only the checkbox checked by user is checked.	Pass
User checks all checkboxes one by one (provided there are three deleted records)	“Check all” checkbox will be checked automatically.	“Check all” checkbox is checked automatically.	Pass

Table 5.2.10: View deleted records testing.

Use Case	Delete records		
Test Case	Expected Output	Actual Output	Result

User deletes record from “record details” page.	The record will be deactivated in database. User will be navigated to “uploaded records” page. The record will be removed from “uploaded records” page and displayed in “deleted records” page.	The record is deactivated in database. User is navigated to “uploaded records” page. The record is removed from “uploaded records” page and displayed in “deleted records” page.	Pass
User deletes one record from “deleted records” page.	The record will be deleted in database. The record will be removed from “deleted records” page. Successfully delete message will be displayed.	The record is deleted in database. The record is removed from “deleted records” page. Successfully delete message will be displayed.	Pass
User deletes multiple records from “deleted records” page.	The records will be deleted in database. The records will be removed from “deleted records” page. Successfully delete message will be displayed.	The records are deleted in database. The records are removed from “deleted records” page. Successfully delete message will be displayed.	Pass

Table 5.2.11: Delete records testing.

Use Case	Restore records		
Test Case	Expected Output	Actual Output	Result
User restores one record where the record is not a duplicate of any record stored in “uploaded records” page.	The record will be activated in database. The record will be removed from “deleted records” page and displayed in “uploaded records” page. Record has been restored message will be displayed.	The record is activated in database. The record is removed from “deleted records” page and displayed in “uploaded records” page. Record has been restored message is displayed.	Pass
User restores one record where the record is a duplicate of one record stored in “uploaded records” page.	Record has not been restored message followed by the duplicate of the detected record stored in “uploaded records” page will be displayed.	Record has not been restored message followed by the duplicate of the detected record stored in “uploaded records” page is displayed.	Pass
User restores multiple records where all records are not a duplicate of any record stored in “uploaded records” page.	The records will be activated in database. The records will be removed from “deleted records” page and displayed in “uploaded records” page. Record has been	The records are activated in database. The records are removed from “deleted records” page and displayed in “uploaded records” page. Record has been restored	Pass

	restored message will be displayed.	message is displayed.	
User restores multiple records where all records are duplicates of records stored in “uploaded records” page.	Record has not been restored message followed by the duplicate of the detected record stored in “uploaded records” page will be displayed.	Record has not been restored message followed by the duplicate of the detected record stored in “uploaded records” page are displayed.	Pass
User restores multiple records where some records are not a duplicate of any record stored in “uploaded records” page” whereas some records are duplicates of records stored in “uploaded records” page.	The records will be activated in database. The records will be removed from “deleted records” page and displayed in “uploaded records” page. Record has been restored message will be displayed. Record has not been restored message followed by the duplicate of the detected record stored in “uploaded records” page will be displayed.	The records are activated in database. The records are removed from “deleted records” page and displayed in “uploaded records” page. Record has been restored message is displayed. Record has not been restored message followed by the duplicate of the detected record stored in “uploaded records” page are displayed.	Pass

Table 5.2.12: Restore records testing.

5.3 Project Challenges

After training the models successfully, new challenge is raised in which the text localization CTPN model is not able to connect the keyword information extraction Character-Aware Neural Language model successfully as the output of the first model is incompatible with the input received by the second model. In the end, OCR EasyOCR model has been used as an API to connect both text localization and keyword information extraction models.

When developing the web-based receipt key information extraction system, the challenge is to understand the specific syntax used by Flask as it is a new framework which is not been studied previously.

5.4 Objectives Evaluation

The developed system in this project is evaluated according to the project objectives as baseline.

Objective	Evaluation	Conclusion
Text localization – To receive an input image which is a receipt and generate the output which are the coordinates of the detected texts in the format of text file.	CTPN model is able to receive an input image which is a receipt and generate the output which are the coordinates of the detected texts in the format of text file.	Achieved.
Optical Character Recognition (OCR) – To generate the output which are the extracted texts from the receipt based on the output provided by the text localization model.	EasyOCR model is able to generate the output which are the extracted texts from the receipt based on the output provided by CTPN model.	Achieved.
Keyword information extraction – To receive the output from text localization and OCR models and generate the output which is	Character-Aware Neural Language model is able to receive the output from CTPN and EasyOCR models and generate the	Achieved.

the keyword information extracted from the receipt.	output which is the keyword information extracted from the receipt.	
Upload multiple receipts – To develop a web-based system which is able to upload multiple receipts for keyword information extraction from them simultaneously.	The developed system is able to upload multiple receipts for keyword information extraction from them simultaneously.	Achieved.
Store keyword information extraction records – To develop a web-based system which is able to store keyword information extraction records for review purpose.	The developed system is able to store keyword information extraction records for review purpose.	Achieved.

Table 5.4.1: Objectives evaluation.

5.5 Concluding Remark

In the end of this chapter, it can be concluded that the project challenges have been handled properly. Therefore, all tests for each case are passed and the project objectives are claimed to be successfully achieved after completing the objectives evaluation.

Chapter 6

Conclusion and Recommendation

6.1 Conclusion

The developed web-based receipt key information extraction system is able to automate the key information extraction process from multiple receipts simultaneously. It is able to minimize the workload to manage numerous receipts manually. All uploaded records can be stored in MySQL server eases the user to manage the records.

To summarize, the system has been developed in Python in Visual Studio Code with WSL environment and Flask framework is used, CTPN, EasyOCR, and Character-Aware Neural Language models are used as API to perform their respective tasks and MySQL server is used as database. This project is successfully carried out as the main project objective and sub-objectives mentioned in the previous chapter have been completed.

6.2 Recommendation

Even the project has been completed successfully, there are some recommendations to further improve the developed system. From the extracted text obtained by EasyOCR model, it has been clearly observed that the text extraction accuracy of the model is low compared with human even the text detected in the image is clear. In order to improve the text detection accuracy of the model, model retrain can be taken into consideration. Besides, the option to generate the report to obtain the summary of key information extracted from the receipts should be provided. The generated report is able to help the user have a clearer view of the details extracted from the receipts instead of only viewing one record details each time after clicking on the URL of the selected record ID.

REFERENCES

- [1] D. S. says, “Optical Character Recognition (OCR) in 2021: In-depth Guide,” research.aimultiple.com, Nov. 23, 2019. <https://research.aimultiple.com/ocr/> (accessed Apr. 14, 2022).
- [2] Z. Tian, W. Huang, T. He, P. He, and Y. Qiao, “Detecting Text in Natural Image with Connectionist Text Proposal Network.” [Online]. Available: <https://arxiv.org/pdf/1609.03605.pdf> (accessed Aug. 30, 2022).
- [3] I. Sydorenko, “Automation with an OCR Algorithm,” labelyourdata.com, Dec. 02, 2020. https://labelyourdata.com/articles/automation-with-ocr-algorithm#the_process_of_building_labeling_and_training_text_recognition_algorithms (accessed Apr. 14, 2022).
- [4] S. Reddy, “Pre-Processing in OCR!!!,” Medium, Jul. 24, 2020. <https://towardsdatascience.com/pre-processing-in-ocr-fc231c6035a7> (accessed Apr. 14, 2022).
- [5] S. Reddy, “Segmentation in OCR !!,” Medium, Aug. 22, 2020. <https://towardsdatascience.com/segmentation-in-ocr-10de176cf373> (accessed Apr. 14, 2022).
- [6] X. Zhao, E. Niu, Z. Wu, and X. Wang, “CUTIE: Learning to Understand Documents with Convolutional Universal Text Information Extractor.” [Online]. Available: <https://arxiv.org/pdf/1903.12363.pdf> (accessed Apr. 14, 2022).
- [7] J. Wang et al., “IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, MARCH 2020 1 Deep High-Resolution Representation Learning for Visual Recognition.” [Online]. Available: <https://arxiv.org/pdf/1908.07919v2.pdf> (accessed Apr. 14, 2022).
- [8] Rani Horev, “BERT Explained: State of the art language model for NLP,” Medium, Nov. 10, 2018. <https://towardsdatascience.com/bert-explained-state-of-the-art-language-model-for-nlp-f8b21a9b6270> (accessed Apr. 14, 2022).
- [9] SG, “BERT,” Medium, Jul. 22, 2019. <https://medium.com/@shoray.goel/bert-f6d23b06069f> (accessed Apr. 14, 2022).
- [10] M. Grootendorst, “Keyword Extraction with BERT,” Medium, Oct. 29, 2020. <https://towardsdatascience.com/keyword-extraction-with-bert-724efca412ea> (accessed Apr. 14, 2022).

- [11] Y. Kim, Y. Jernite, D. Sontag, and A. M. Rush, "Character-Aware Neural Language Models." [Online]. Available: <https://arxiv.org/pdf/1508.06615.pdf> (accessed Aug. 30, 2022).
- [12] BlackStar1313, "ICDAR-2019-RRC-SROIE," Oct. 27, 2021.
https://github.com/BlackStar1313/ICDAR-2019-RRC-SROIE/blob/master/text_localization/ctpn/figures/ctpn_architecture.png (accessed Aug. 30, 2022).
- [13] BlackStar1313, "ICDAR-2019-RRC-SROIE," Oct. 27, 2021.
https://github.com/BlackStar1313/ICDAR-2019-RRC-SROIE/blob/master/text_localization/figures/crop_preprocessing.png (accessed Aug. 30, 2022).
- [14] rkcosmos, "EasyOCR," Jul. 16, 2020.
https://github.com/JaidedAI/EasyOCR/blob/master/examples/easyocr_framework.jpeg (accessed Aug. 30, 2022).
- [15] BlackStar1313, "ICDAR-2019-RRC-SROIE," Oct. 27, 2021.
https://github.com/BlackStar1313/ICDAR-2019-RRC-SROIE/blob/master/keyword_information_extraction/figures/charlm_architecture.png (accessed Aug. 30, 2022).

FINAL YEAR PROJECT WEEKLY REPORT

(Project II)

Trimester, Year: Y3S1	Study week no.: 1
Student Name & ID: Fong Jia Yee, 19ACB04133	
Supervisor: Ts Dr Tan Hung Khoon	
Project Title: Web-based Receipt Key Information Extraction System	

1. WORK DONE

[Please write the details of the work done in the last fortnight.]

The models for text localization, Optical Character Recognition (OCR), and keyword information extraction have been studied and analyzed.

2. WORK TO BE DONE

Start doing the refinement of chapter 1 (introduction) and chapter 2 (literature review) of the report.


Select the suitable text localization, OCR, and keyword information extraction models.

3. PROBLEMS ENCOUNTERED

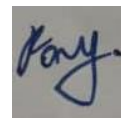
So far so good.

4. SELF EVALUATION OF THE PROGRESS

So far so good.



Supervisor's signature



Student's signature

FINAL YEAR PROJECT WEEKLY REPORT

(Project II)

Trimester, Year: Y3S1	Study week no.: 2
Student Name & ID: Fong Jia Yee, 19ACB04133	
Supervisor: Ts Dr Tan Hung Khoon	
Project Title: Web-based Receipt Key Information Extraction System	

1. WORK DONE

[Please write the details of the work done in the last fortnight.]

The refinement of chapter 1 (introduction) and chapter 2 (literature review) of the project report.

The suitable text localization, OCR, and keyword information extraction models have been selected.

2. WORK TO BE DONE

Start doing chapter 3 (system design), chapter 4 (system implementation) and chapter 5 (system evaluation discussion) of the project report simultaneously.

Start doing prototype and knowledge building for the project.

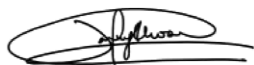
Import the code of the models into Visual Studio Code.

3. PROBLEMS ENCOUNTERED

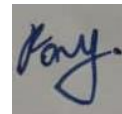
So far so good.

4. SELF EVALUATION OF THE PROGRESS

So far so good.



Supervisor's signature



Student's signature

FINAL YEAR PROJECT WEEKLY REPORT

(Project II)

Trimester, Year: Y3S1	Study week no.: 3
Student Name & ID: Fong Jia Yee, 19ACB04133	
Supervisor: Ts Dr Tan Hung Khoon	
Project Title: Web-based Receipt Key Information Extraction System	

1. WORK DONE

[Please write the details of the work done in the last fortnight.]

The code of the models has been imported into Visual Studio Code without any error.

2. WORK TO BE DONE

Continue doing chapter 3 (system design), chapter 4 (system implementation) and chapter 5 (system evaluation discussion) of the project report simultaneously.

Continue doing prototype and knowledge building for the project.


Study and analyze the code of the models.

3. PROBLEMS ENCOUNTERED

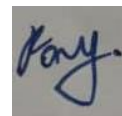
So far so good.

4. SELF EVALUATION OF THE PROGRESS

So far so good.



Supervisor's signature



Student's signature

FINAL YEAR PROJECT WEEKLY REPORT

(Project II)

Trimester, Year: Y3S1	Study week no.: 4
Student Name & ID: Fong Jia Yee, 19ACB04133	
Supervisor: Ts Dr Tan Hung Khoon	
Project Title: Web-based Receipt Key Information Extraction System	

1. WORK DONE

[Please write the details of the work done in the last fortnight.]

The code of the models has been studied and analyzed.

2. WORK TO BE DONE

Continue doing chapter 3 (system design), chapter 4 (system implementation) and chapter 5 (system evaluation discussion) of the project report simultaneously.

Continue doing prototype and knowledge building for the project.


Train the text localization and keyword information extraction models.

3. PROBLEMS ENCOUNTERED

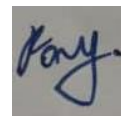
So far so good.

4. SELF EVALUATION OF THE PROGRESS

So far so good.



Supervisor's signature



Student's signature

FINAL YEAR PROJECT WEEKLY REPORT

(Project II)

Trimester, Year: Y3S1	Study week no.: 5
Student Name & ID: Fong Jia Yee, 19ACB04133	
Supervisor: Ts Dr Tan Hung Khoon	
Project Title: Web-based Receipt Key Information Extraction System	

1. WORK DONE

[Please write the details of the work done in the last fortnight.]

The text localization and keyword information extraction models have been trained.

2. WORK TO BE DONE

Continue doing chapter 3 (system design), chapter 4 (system implementation) and chapter 5 (system evaluation discussion) of the project report simultaneously.

Continue doing prototype and knowledge building for the project.


Connect the text localization, OCR, and keyword information extraction models.

3. PROBLEMS ENCOUNTERED

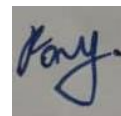
So far so good.

4. SELF EVALUATION OF THE PROGRESS

So far so good.



Supervisor's signature



Student's signature

FINAL YEAR PROJECT WEEKLY REPORT

(Project II)

Trimester, Year: Y3S1	Study week no.: 6
Student Name & ID: Fong Jia Yee, 19ACB04133	
Supervisor: Ts Dr Tan Hung Khoon	
Project Title: Web-based Receipt Key Information Extraction System	

1. WORK DONE

[Please write the details of the work done in the last fortnight.]

The text localization, OCR, and keyword information extraction models have been connected.

The text localization model is able to receive multiple images as input. The keyword information extraction model is able to generate key information extracted from each image.

2. WORK TO BE DONE

Continue doing chapter 3 (system design), chapter 4 (system implementation) and chapter 5 (system evaluation discussion) of the project report simultaneously.

Continue doing prototype and knowledge building for the project.

Register and login of the web-based receipt key information extraction system.

3. PROBLEMS ENCOUNTERED

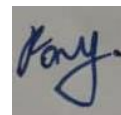
So far so good.

4. SELF EVALUATION OF THE PROGRESS

So far so good.



Supervisor's signature



Student's signature

FINAL YEAR PROJECT WEEKLY REPORT

(Project II)

Trimester, Year: Y3S1	Study week no.: 7
Student Name & ID: Fong Jia Yee, 19ACB04133	
Supervisor: Ts Dr Tan Hung Khoon	
Project Title: Web-based Receipt Key Information Extraction System	

1. WORK DONE

[Please write the details of the work done in the last fortnight.]

Register and login of the web-based receipt key information extraction system.

2. WORK TO BE DONE

Continue doing chapter 3 (system design), chapter 4 (system implementation) and chapter 5 (system evaluation discussion) of the project report simultaneously.

Continue doing prototype and knowledge building for the project.

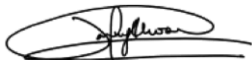
Function of the web-based receipt key information extraction system to upload multiple images.

3. PROBLEMS ENCOUNTERED

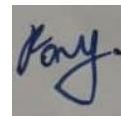
So far so good.

4. SELF EVALUATION OF THE PROGRESS

So far so good.



Supervisor's signature



Student's signature

FINAL YEAR PROJECT WEEKLY REPORT

(Project II)

Trimester, Year: Y3S1	Study week no.: 8
Student Name & ID: Fong Jia Yee, 19ACB04133	
Supervisor: Ts Dr Tan Hung Khoon	
Project Title: Web-based Receipt Key Information Extraction System	

1. WORK DONE

[Please write the details of the work done in the last fortnight.]

Function of the web-based receipt key information extraction system to upload multiple images.

2. WORK TO BE DONE

Continue doing chapter 3 (system design), chapter 4 (system implementation) and chapter 5 (system evaluation discussion) of the project report simultaneously.

Continue doing prototype and knowledge building for the project.

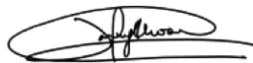
Function of the web-based receipt key information extraction system to detect duplicate receipt.

3. PROBLEMS ENCOUNTERED

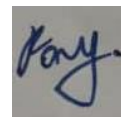
So far so good.

4. SELF EVALUATION OF THE PROGRESS

So far so good.



Supervisor's signature



Student's signature

FINAL YEAR PROJECT WEEKLY REPORT

(Project II)

Trimester, Year: Y3S1	Study week no.: 9
Student Name & ID: Fong Jia Yee, 19ACB04133	
Supervisor: Ts Dr Tan Hung Khoon	
Project Title: Web-based Receipt Key Information Extraction System	

1. WORK DONE

[Please write the details of the work done in the last fortnight.]

Function of the web-based receipt key information extraction system to detect duplicate receipt.

2. WORK TO BE DONE

Continue doing chapter 3 (system design), chapter 4 (system implementation) and chapter 5 (system evaluation discussion) of the project report simultaneously.

Continue doing prototype of the project.

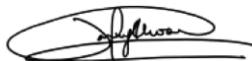
Function of the web-based receipt key information extraction system to display the records owned by the user.

3. PROBLEMS ENCOUNTERED

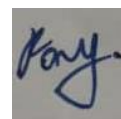
So far so good.

4. SELF EVALUATION OF THE PROGRESS

So far so good.



Supervisor's signature



Student's signature

FINAL YEAR PROJECT WEEKLY REPORT

(Project II)

Trimester, Year: Y3S1	Study week no.: 10
Student Name & ID: Fong Jia Yee, 19ACB04133	
Supervisor: Ts Dr Tan Hung Khoon	
Project Title: Web-based Receipt Key Information Extraction System	

1. WORK DONE

[Please write the details of the work done in the last fortnight.]

Function of the web-based receipt key information extraction system to display the records owned by the user.

2. WORK TO BE DONE

Continue doing chapter 3 (system design), chapter 4 (system implementation) and chapter 5 (system evaluation discussion) of the project report simultaneously.

Continue doing prototype of the project.

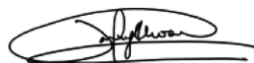
Function of the web-based receipt key information extraction system to display the record details and delete the record.

3. PROBLEMS ENCOUNTERED

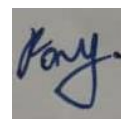
So far so good.

4. SELF EVALUATION OF THE PROGRESS

So far so good.



Supervisor's signature



Student's signature

FINAL YEAR PROJECT WEEKLY REPORT

(Project II)

Trimester, Year: Y3S1	Study week no.: 11
Student Name & ID: Fong Jia Yee, 19ACB04133	
Supervisor: Ts Dr Tan Hung Khoon	
Project Title: Web-based Receipt Key Information Extraction System	

1. WORK DONE

[Please write the details of the work done in the last fortnight.]

Function of the web-based receipt key information extraction system to display the record details and delete the record.

2. WORK TO BE DONE

Continue doing chapter 3 (system design), chapter 4 (system implementation) and chapter 5 (system evaluation discussion) of the project report simultaneously.

Continue doing prototype of the project.

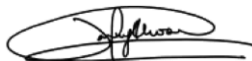
Function of the web-based receipt key information extraction system to display the deleted records.

3. PROBLEMS ENCOUNTERED

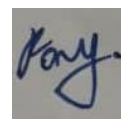
So far so good.

4. SELF EVALUATION OF THE PROGRESS

So far so good.



Supervisor's signature



Student's signature

FINAL YEAR PROJECT WEEKLY REPORT

(Project II)

Trimester, Year: Y3S1	Study week no.: 12
Student Name & ID: Fong Jia Yee, 19ACB04133	
Supervisor: Ts Dr Tan Hung Khoon	
Project Title: Web-based Receipt Key Information Extraction System	

1. WORK DONE

[Please write the details of the work done in the last fortnight.]

Function of the web-based receipt key information extraction system to display the deleted records.

2. WORK TO BE DONE

Continue doing chapter 3 (system design), chapter 4 (system implementation) and chapter 5 (system evaluation discussion) of the project report simultaneously.

Continue doing prototype of the project.

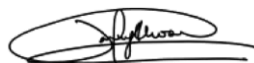
Function of the web-based receipt key information extraction system to delete the deleted record permanently, restore the deleted record, and reset password.

3. PROBLEMS ENCOUNTERED

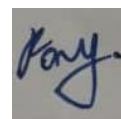
So far so good.

4. SELF EVALUATION OF THE PROGRESS

So far so good.



Supervisor's signature



Student's signature

FINAL YEAR PROJECT WEEKLY REPORT

(Project II)

Trimester, Year: Y3S1	Study week no.: 13
Student Name & ID: Fong Jia Yee, 19ACB04133	
Supervisor: Ts Dr Tan Hung Khoon	
Project Title: Web-based Receipt Key Information Extraction System	

1. WORK DONE

[Please write the details of the work done in the last fortnight.]

Function of the web-based receipt key information extraction system to delete the deleted record permanently, restore the deleted record, and reset password.

Chapter 3 (system design), chapter 4 (system implementation), chapter 5 (system evaluation discussion) of the project report have been completed.

Prototype of the project has been completed.

2. WORK TO BE DONE

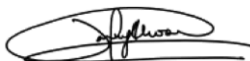
Start doing chapter 6 (conclusion and recommendation) of the project report.

3. PROBLEMS ENCOUNTERED

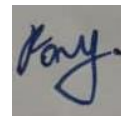
So far so good.

4. SELF EVALUATION OF THE PROGRESS

So far so good.



Supervisor's signature



Student's signature



FACULTY OF INFORMATION AND COMMUNICATION TECHNOLOGY

SaveMore
Web-based Receipt Key Information Extraction System

INTRODUCTION

Difficult to manage numerous receipts manually? SaveMore helps you manage all receipt images automatically.

KEY INFORMATION EXTRACTION

A process of extracting texts from a scanned document where extracted texts have both spatial and semantic features. The input of key information extraction is a scanned document, and its output is extracted texts from the document.

PROBLEM STATEMENT

Receipts may come from different layouts, angles. Keys and values might have no association between them due to their arrangement.

METHODS

- Developed by Python on Visual Studio Code with WSL environment
- Flask framework
- Database stored in MySQL server
- CTPN, EasyOCR, and Character-Aware Neural Language models as API

PROJECT OBJECTIVES

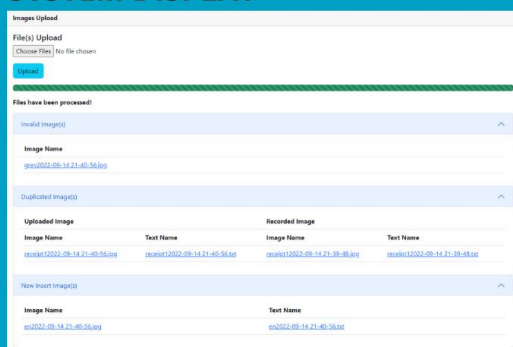
Before developing web-based receipt key information extraction system:

- To connect text localization, Optical Character Recognition, and keyword information extraction models to receive an input receipt image and generate an output text file consists of key information extracted from the receipt.

After connecting the models and obtaining the desired outputs:

- To develop a system which is able upload multiple receipts for keyword information extraction from them simultaneously.
- To develop a system which is able to store keyword information extraction records for review purpose.

SYSTEM DISPLAY



DEVELOPED BY FONG JIA YEE | COLLABORATE WITH TAN ANN GEE |
SUPERVISED BY TS DR TAN HUNG KHOON | MODERATED BY PN NORAINI BINTI IBRAHIM

PLAGIARISM CHECK RESULT

19ACB04133_FYP2

ORIGINALITY REPORT

9%

SIMILARITY INDEX

5%

INTERNET SOURCES

4%

PUBLICATIONS

5%

STUDENT PAPERS

PRIMARY SOURCES

1

link.springer.com

Internet Source

1%

2

Jingxue Liu, Fanrong Meng, Yong Zhou, Bing Liu. "Character-Level neural networks for short text classification", 2017 International Smart Cities Conference (ISC2), 2017

Publication

1%

3

eprints.utar.edu.my

Internet Source

<1%

4

Submitted to Asia Pacific University College of Technology and Innovation (UCTI)

Student Paper

<1%

5

Submitted to University of Cincinnati

Student Paper

<1%

6

Submitted to University of Greenwich

Student Paper

<1%

7

Submitted to Whitireia Community Polytechnic

Student Paper

<1%

8	Rohitash Chandra, Venkatesh Kulkarni. "Semantic and Sentiment Analysis of Selected Bhagavad Gita Translations Using BERT-Based Language Framework", IEEE Access, 2022 Publication	<1 %
9	www.ncbi.nlm.nih.gov Internet Source	<1 %
10	Submitted to Universiti Tunku Abdul Rahman Student Paper	<1 %
11	towardsdatascience.com Internet Source	<1 %
12	Submitted to Associatie K.U.Leuven Student Paper	<1 %
13	Qingheng Zhuang, Li Zhu, Sen Lin. "A GAN-Bert Based Fault Diagnosis Model for CBTC Data Communication Systems Using Edge-to-edge Collaboration Training", ICC 2022 - IEEE International Conference on Communications, 2022 Publication	<1 %
14	docplayer.net Internet Source	<1 %
15	Zhi Tian, Weilin Huang, Tong He, Pan He, Yu Qiao. "Chapter 4 Detecting Text in Natural Image with Connectionist Text Proposal Network", Springer Nature, 2016	<1 %

16	Submitted to Liverpool John Moores University Student Paper	<1 %
17	www.gabormelli.com Internet Source	<1 %
18	Submitted to University of Westminster Student Paper	<1 %
19	S-space.snu.ac.kr Internet Source	<1 %
20	worldwidescience.org Internet Source	<1 %
21	ikee.lib.auth.gr Internet Source	<1 %
22	deepai.org Internet Source	<1 %
23	patentimages.storage.googleapis.com Internet Source	<1 %
24	Submitted to Isik University Student Paper	<1 %
25	"Computer Vision – ECCV 2016", Springer Nature, 2016 Publication	<1 %
26	Submitted to University of East London Student Paper	<1 %

27	Submitted to North East Surrey College of Technology, Surrey Student Paper	<1 %
28	Submitted to Panipat Institute of Engineering & Technology Student Paper	<1 %
29	Submitted to University of Sunderland Student Paper	<1 %
30	Submitted to Asia Pacific Institute of Information Technology Student Paper	<1 %
31	Submitted to INTI International University Student Paper	<1 %
32	Submitted to TAR University College Student Paper	<1 %
33	Submitted to Taylor's Education Group Student Paper	<1 %
34	kc.umn.ac.id Internet Source	<1 %
35	Jingjing Zhang, Yuanyuan Li, Teng Li, Qing Yan, Yi Xia, Caifeng Shan. "License Plate Localization in Unconstrained Scenes Using a Two-Stage CNN-RNN", IEEE Sensors Journal, 2019 Publication	<1 %

www.hindawi.com

36	Internet Source	<1 %
37	"Handbook of Print Media", Springer Science and Business Media LLC, 2001 Publication	<1 %
38	Rongyan Li, Zeyu Xu, Cheng Su, Rong Yang. "Automatic identification of semi-tracks on apatite and mica using a deep learning method", Computers & Geosciences, 2022 Publication	<1 %
39	assets.researchsquare.com Internet Source	<1 %
40	vtechworks.lib.vt.edu Internet Source	<1 %
41	Dongyao Jia, Chuanwang Zhang, Bing Zhang. "Crowd density classification method based on pixels and texture features", Machine Vision and Applications, 2021 Publication	<1 %
42	etheses.whiterose.ac.uk Internet Source	<1 %
43	txsun1997.github.io Internet Source	<1 %

Universiti Tunku Abdul Rahman			
Form Title : Supervisor's Comments on Originality Report Generated by Turnitin for Submission of Final Year Project Report (for Undergraduate Programmes)			
Form Number: FM-IAD-005	Rev No.: 0	Effective Date: 01/10/2013	Page No.: 1 of 1



FACULTY OF INFORMATION AND COMMUNICATION TECHNOLOGY

Full Name(s) of Candidate(s)	Fong Jia Yee
ID Number(s)	19ACB04133
Programme / Course	Bachelor of Information Systems (Honours) Business Information Systems / Project II
Title of Final Year Project	Web-based Receipt Key Information Extraction System

Similarity	Supervisor's Comments (Compulsory if parameters of originality exceed the limits approved by UTAR)
Overall similarity index: 9 % Similarity by source Internet Sources: 5 % Publications: 4 % Student Papers: 5 %	
Number of individual sources listed of more than 3% similarity: 0	
Parameters of originality required and limits approved by UTAR are as Follows: (i) Overall similarity index is 20% and below, and (ii) Matching of individual sources listed must be less than 3% each, and (iii) Matching texts in continuous block must not exceed 8 words <i>Note: Parameters (i) – (ii) shall exclude quotes, bibliography and text matches which are less than 8 words.</i>	

Note Supervisor/Candidate(s) is/are required to provide softcopy of full set of the originality report to Faculty/Institute

Based on the above results, I hereby declare that I am satisfied with the originality of the Final Year Project Report submitted by my student(s) as named above.

Signature of Supervisor

Name: Tan Hung Khoon

Date: 17/9/2022

-

Signature of Co-Supervisor

Name: -

Date: -



UNIVERSITI TUNKU ABDUL RAHMAN

FACULTY OF INFORMATION & COMMUNICATION TECHNOLOGY (KAMPAR CAMPUS)

CHECKLIST FOR FYP2 THESIS SUBMISSION

Student Id	19ACB04133
Student Name	Fong Jia Yee
Supervisor Name	Ts Dr Tan Hung Khoon

TICK (√)	DOCUMENT ITEMS
	Your report must include all the items below. Put a tick on the left column after you have checked your report with respect to the corresponding item.
	Front Plastic Cover (for hardcopy)
√	Title Page
√	Signed Report Status Declaration Form
√	Signed FYP Thesis Submission Form
√	Signed form of the Declaration of Originality
√	Acknowledgement
√	Abstract
√	Table of Contents
√	List of Figures (if applicable)
√	List of Tables (if applicable)
	List of Symbols (if applicable)
√	List of Abbreviations (if applicable)
√	Chapters / Content
√	Bibliography (or References)
√	All references in bibliography are cited in the thesis, especially in the chapter of literature review
	Appendices (if applicable)
√	Weekly Log
√	Poster
√	Signed Turnitin Report (Plagiarism Check Result - Form Number: FM-IAD-005)
√	I agree 5 marks will be deducted due to incorrect format, declare wrongly the ticked of these items, and/or any dispute happening for these items in this report.

*Include this form (checklist) in the thesis (Bind together as the last page)

I, the author, have checked and confirmed all the items listed in the table are included in my report.

(Signature of Student)

Date: 9/9/2022