A

Project Report

on

HANDWRITTEN TEXT RECOGNITION

Submitted in Partial Fulfillment of

the Requirements for the Degree

of

Bachelor of Engineering

in

Information Technology

to

Kavayitri Bahinabai Chaudhari
North Maharashtra University, Jalgaon

Submitted by

Leena T. Supe
Kanchan A. Prajapati
Rameez M. Jamadar
Nitin B. Rajput

Under the Guidance of

Dr. Dnyaneshwar Kirange

DEPARTMENT OF INFORMATION TECHNOLOGY
SSBT's COLLEGE OF ENGINEERING AND TECHNOLOGY,
BAMBHORI, JALGAON - 425 001 (MS)
2022 - 2023

SSBT's COLLEGE OF ENGINEERING AND TECHNOLOGY, BAMBHORI, JALGAON - 425 001 (MS)
DEPARTMENT OF INFORMATION TECHNOLOGY

CERTIFICATE

This is to certify that the Minor Project entitled *handwritten text recognition* , submitted by

Leena T. Supe
Kanchan A. Prajapati
Rameez M. Jamadar

Nitin B. Rajput

in partial fulfillment of the degree of *Bachelor of Engineering* in *Information Technology* has been satisfactorily carried out under my guidance as per the requirement of Kavayitri Bahinabai Chaudhari North Maharashtra University, Jalgaon.

Date: April 17, 2023
Place: Jalgaon

Dr. Dnyaneshwar Kirange
Guide

Dr. Manoj E. Patil   Prof. Dr. G.K.Patnaik
Head   **Principal**

Acknowledgements

The success of our project depends largely on the encouragement and help from many others. We take this opportunity to express our thankfulness to the people who have contributed in the successful completion of this project and have been the real guiding force in our project.We would like to show our gratitude to our project mentor **Dr. Dnyaneshwar Kirange sir** continuously guiding and tremendously helping us throughout this project. We feel motivated and encouraged every time we attend his meeting. Without his encouragement and guidance the completion of this project on **Convert Handwritten Text into Normal Editable Eext** detection would not have been achieved

<div align="right">

Leena T. Supe
Kanchan A. Prajapati
Rameez M. Jamadar
Nitin B. Rajput

</div>

Contents

Abstract

This system is useful for recognizing all character(English) given as in input image. Once input image of character is given to proposed system, then it will recognize input character which is given in image. Recognition and classification of characters are done by Neural Network.The main aim of this project is to effectively recognize a particular character of type format using the Artificial Neural Network approach. The main object of this system is to provide an easy user interface to input the object image. The User should be able to upload the image and the system should retrieve text present in the image and display them to the user.

Chapter 1

Introduction

The purpose of this project is to take handwritten English characters as input, process the character, train the neural network algorithm, to recognize the pattern and modify the character to a beautified version of the input. This project ts aimed at developing software which will be helpful in recognizing characters of English language. This project is restricted to English characters only. It can be further developed to recognize the characters of different languages. It engulfs the concept of neural network. One of the primary means by which computers are endowed with human like abilities is through the use of a neural network. Neural networks are particularly useful for solving problems that cannot be expressed as a series of steps, such as recognizing patterns, classifying them into groups, series prediction and data mining.

Pattern recognition is perhaps the most common use of neural networks. The neural network is presented with a target vector and also a vector which contains the pattern information, this could be an image and hand written data. The neural network then attempts to determine if the input data matches a pattern that the neural network has memorized. A neural network trained for classification is designed to take input samples and classify them into groups. These groups may be fuzzy, without clearly defined boundaries. This project concerns detecting free handwritten characters.

1.1 Background

This website is useful for recognizing all character(English) given as in input image. Once input image of character is given to proposed system, then it will recognize input character which is given in image.Recognition and classification of characters are done by Neural Network.The main aim of this project isto effectively recognize a particular character of type format using the Artificial Neural Network approach.

1.2 Motivation

The purpose of this project is to take English handwritten documents as input, recognize the text and modify the handwriting such that it is a beautified version of the input. Thus the project comprises of two parts - handwriting recognition and beautification. Lots of work has been done in the field of character recognition but not much for analyzing a complete document. Recognizing the text of a document would be useful in many diverse applications like reading medical prescriptions, bank cheques and other official documents. It will also find uses in detective or police departments in applications like handwriting based person identification, identifying real from forged documents, etc.

1.3 Scope

Applications of offline handwriting recognition are numerous: reading postal addresses, bank check amounts, and forms. Furthermore, OCR plays an important role for digital libraries, allowing the entry of image textual information into computers by digitization, image restoration, and recognition methods.

1.4 Objective

To provide an easy user interface to input the object image.
- User should be able to upload the image.
- System should be able to pre-process the given input to suppress the background.
- System should detect text regions present in the image.
- System should retrieve text present in the image and display them to the user.
- Classifying the news as fake or real
- Deploying the model.

1.5 Selection of Life cycle model

The software development life cycle model selected for this project is the Waterfall Model. Waterfall approach was the first SDLC (Software Development Life Cycle) Model to be widely used in software engineering to ensure success of the project. It was developed by Winston W. Royce in 1970. In "The Waterfall" approach, the whole process of software development is divided into separate phases, typically the outcome of one phase acts as the input for the next phase sequentially. All the phases are cascaded to each other in which progress is seen as flowing steadily downwards (like a waterfall) through the phases. Requirements for this project are well documented and fixed.
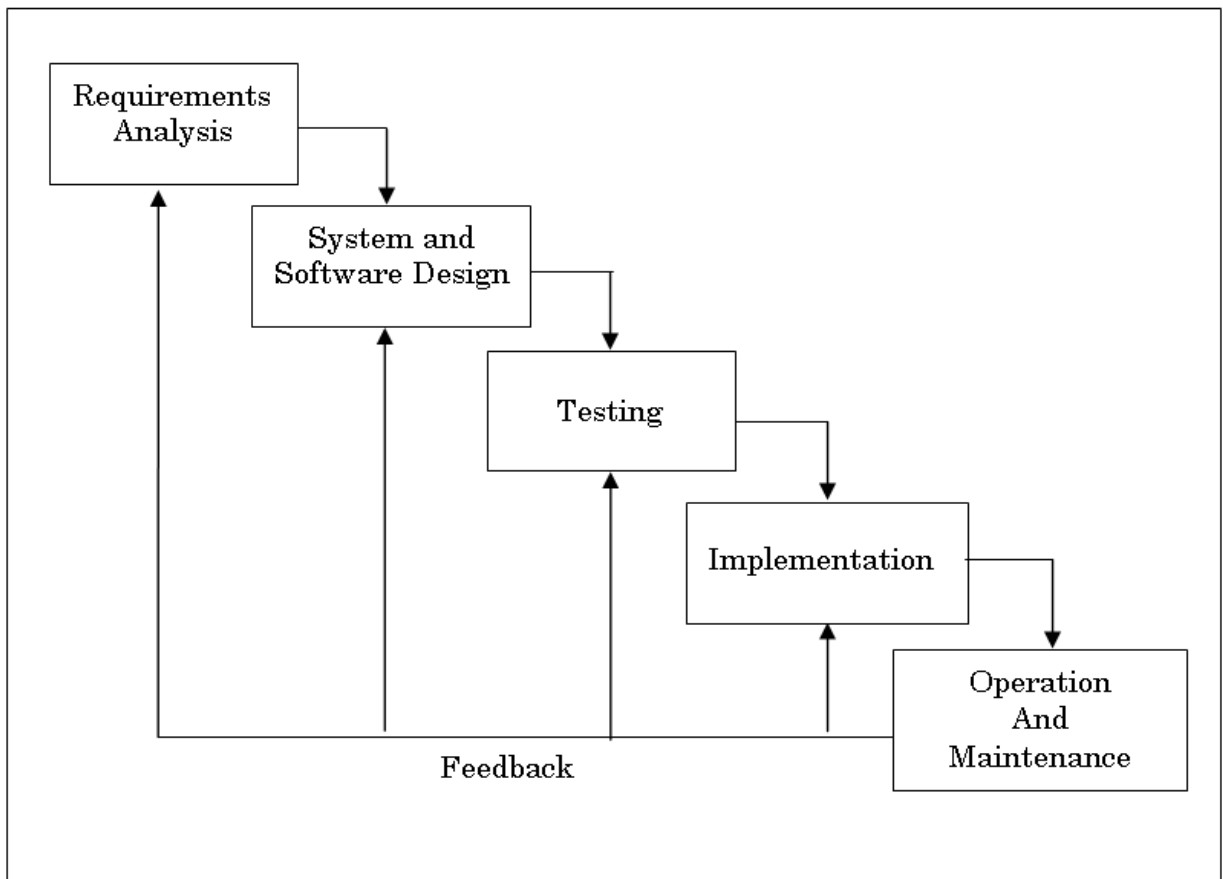
Figure 1.0 Modified Waterfall Approach Model

Waterfall Model is best suited model for this project.

1. Because requirements are easily understandable and defined

2. We can define requirements in early stage of development

3. User involvement in all phases is not necessary

4. Limited user's participation

1.6 Organization Of Report

Chapter 1 entitled as Introduction describes the details about Background, Problem Definition, Scope and Objective of the project, Identification of Software Development Process Model and Organization of report.

Chapter 2 entitled as Project Planning and Management consists of details about the Feasibility Study, Risk Analysis, Project Scheduling, Effort Allocation and Cost Estimation of the project.

Chapter 3 entitled as Analysis describes in detail, the Requirement Collection and Identification, H/w and S/w Requirements, Functional and Non-Functional Requirements and a Software Requirements Specification(SRS).

Chapter 4 includes design about System Architecture, Data Flow Diagram and various UML Diagrams.

Chapter 5 titled conclusion and future work discusses about the overall outcomes and future possibilities of the project.

1.7   Summary

As mentioned in above sections, this project aims at convert handwritten text into normal editable text. The scopes, objective, etc. are as mentioned above. In the next chapter, project planning and management will be discussed.

Chapter 2

Project And Management

Project planning is a procedural step in project management. It is the practice of initiating, planning, executing, controlling and closing the work team to achieve specific goals. Project planning and management is important because it ensures that the right people do the right things, at the right time. It also ensures the proper project lifecycle.

The organization of this chapter is as below. Section 2.1 shows the Feasibility Study of the project. Risk Analysis of the project is represented in Section 2.2 and Project Scheduling is described in Section 2.3. Section 2.4 and 2.5 describe the Effort Allocation and Cost Estimation respectively. The Summary is mentioned in Section 2.6.

2.1 Feasibility Study

A feasibility study is an assessment of the practicality of a proposed project or system. A feasibility study aims to objectively and rationally uncover the strengths and weaknesses of an existing business or proposed venture, opportunities and threats present in the natural environment, the resources required to carry through, and ultimately the prospects for success. In its simplest terms, the two criteria to judge feasibility are cost required and value to be attained. A well-designed feasibility study should provide a historical background of the business or project, a description of the product or service, accounting statements, details of the operations and management, marketing research and policies, financial data, legal requirements and tax obligations. Generally, feasibility studies precede technical development and project implementation.

A feasibility study evaluates the project's potential for success; therefore, perceived objectivity is an important factor in the credibility of the study for potential investors and lending institutions. It must therefore be conducted with an objective, unbiased approach to provide information upon which decisions can be based. Taking into consideration the technical, operational and economic feasibilities as below, the project can be anticipated as feasible overall. There are few types of feasibility that exists. So, developers should take

care of these feasibility and take them into consideration:

2.1.1 Technical Feasibility

This assessment is based on an outline design of system requirements, to determine whether the company has the technical expertise to handle completion of the project. At this level, the concern is whether the proposal is both technically and legally feasible (assuming moderate cost). It is an evaluation of the hardware and software and how it meets the need of the proposed system.

This project is built upon VS code Editor, a simple Web-Application with Python as the programming language and can be easily hosted on cloud server. Also all the other technologies used are capable of building such a project and serve as well as maintain it for longer period of time. All the required hardware and software are easily available in the market. Hence the project is technically feasible.

2.1.2 Operational Feasibility

Operational feasibility is the measure of how well a proposed system solves the problems, and takes advantage of the opportunities identified during scope definition and how it satisfies the requirements identified in the requirements analysis phase of system development.

The operational feasibility assessment focuses on the degree to which the proposed development project fits in with the existing business environment and objectives with regard to development schedule, delivery date, corporate culture and existing business processes. The application is operationally feasible since it is build with the idea for integration with various existing applications and systems.

2.1.3 Economical Feasibility

Describes how much time is available to build the new system, when it can be built, whether it interferes with normal business operations, type and amount of resources required, dependencies, and developmental procedures with company revenue prospectus.

As the necessary hardware and the software are easily available in the market at low cost, the initial investment is the only cost incurred and does not need further enhancement. Hence it is economically feasible.

2.2 Risk Analysis

Risk Analysis and Management is a key project management practice to ensure that the least number of surprises occur while your project is underway. While we can never predict

---

the future with certainty, we can apply a simple and streamlined risk management process to predict the uncertainties in the projects and minimize the occurrence or impact of these uncertainties. This project has a very small window for experiencing failures since the real life interaction is less. This improves the chance of successful project completion and reduces the consequences of those risks.

2.3   Project Scheduling

Generally, project scheduling can be stated as the estimated time required for any project from its time of beginning to the end of the project. In detail, for every task, there is a deadline because all the tasks for the completion of project are planned earlier. So that, each task is scheduled to certain time limit. 15 In short, in project management, listing of projects milestones, activities and all from starting to end date, are considered in the project scheduling. A schedule is generally used in the project planning and management of the project with some kind of attributes as budget, task allocation and duration, resource allocation and all.

Sr No.	Tasks	Start Date	Completion Date
1	Selection Of Title	14-Sep-22	25-Sep-22
2	Gathering Information	26-Sep-22	10-oct-22
-3	Discussion With Guide	11-Oct-22	16-Oct-22
4	Planning/Requirement Gathering and Analysis	17-Oct-22	31-Oct-22

Table :- Task scheduling for the project
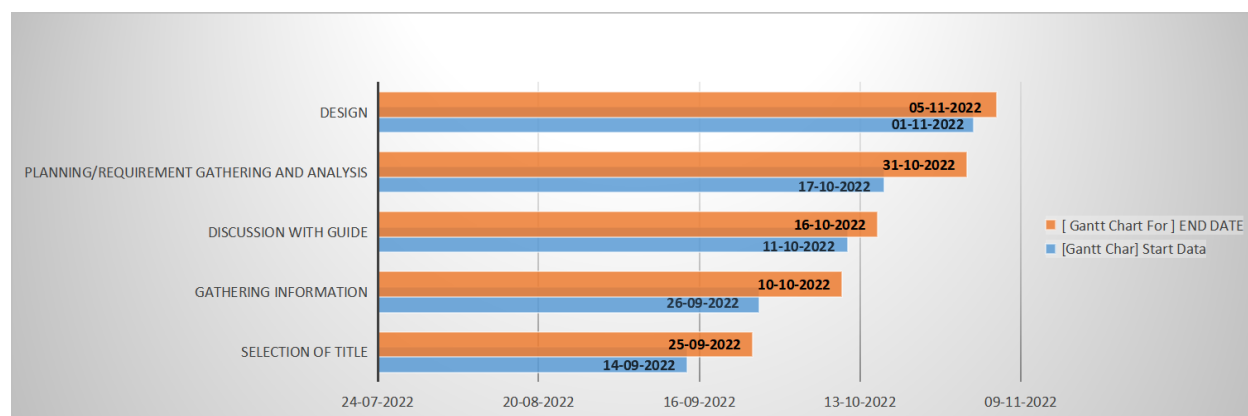
Gantt Chart For Task Scheduling

2.4 Effort Allocation

Effort Allocation is necessary so every team member can give its best to the project. Project was divided into smaller module and task form, for simplification and easy understanding of project overall. Some modules include every team associate's presence to take advantage of team decision taking skills, and some task include some individual member to work on it with precision.

We divided the project into 6 modules.

1. Gathering of Information
2. Planning / Requirement Analysis
3. Study of included Stack and frameworks
4. Selection of Life cycle Model
5. Planning and Management
6. Analysis & Design UML

Sr.no	Modules	Team Associates			
		Leena	Kanchan	Rameez	Nitin
1	Gathering of Information	✓	✓	✓	✓
2	Planning/Requirement Analysis	✓	✓	✓	✓
3	Study Of included stack and framework	✓	✓	✓	✓
4	Selection of life cycle model	✓	✓	✓	
5	Planning and Management	✓	✓		✓
6	Analysis & Design UML			✓	

Table : - Chart of Effort Allocation

2.5 Cost Estimation

Cost Estimation is an important phase for any project. It predicts if the project investment is adequate or there will shortage of capital. It presents the total cost required for development of project. Cost Estimation should be done before initiating the development to prevent loss of efforts and project failure during development. For estimation of cost for this project, we need to consider the server costs for deployment, although the cost is extremely variable since it is dependent on real-time usage. The cost for a machine learning project is generally calculated in three components i.e. data cost, research cost and production cost. Since our project has the required dataset already available, the data cost for the project is zero. The research cost is dependent on the number of people involved in the project for the amount of time required for the project. Assuming the cost for each person who does research to be Rs. 15000 per month, if 4 people work for a timeline of 6 months, the research cost will be

$$\text{Research Cost} = 4 \text{ people x } 6 \text{ months x Rs. } 15000$$
$$= \text{Rs. } 3,60,000$$

According to the Google Cloud Calculator, the cost for deploying Identifier Web Page as a server, 4 vCPU's and Upto 16 GB RAM is Rs. 24,900

$$\text{Production Cost} = 12 \text{ months x Rs. } 3000 = \text{Rs. } 36,000$$

Hence the total cost of the project can be calculated as

$$\text{Total Cost} = \text{Research Cost} + \text{Production Cost}$$
$$= \text{Rs. } 3,60,000 + \text{Rs. } 24,900$$
$$\text{Total Cost} = \text{Rs. } 3,84,900$$

2.6 Summary

The project, is hence found to be feasible since there is a balance of resources required and the cost incurred. Also the project reduces burden on existing methods required for analysis and maintenance of roads. The project will be able to easily integrate with other required systems

Chapter 3

Analysis

The development of computer-based information system includes the system analysis phase which produces or enhances the data model which itself is to creating or enhancing a database. There are a number of different approaches to system analysis. The analysis is the process which is used to analyze, refine and scrutinize the gathered information of entities in order to make consistence and unambiguous information. Analysis activity provides a graphical view of the entire System. System Analysis is the process of gathering and interpreting facts, diagnosing problems and using the facts to improve the system. System analysis chapter will show overall system analysis of the concept, description of the system, meaning of the system. System analysis is the study of sets of interacting entities, including computer system analysis. The organization of this Chapter is as follows. Section 3.1 represents Requirement Collection and Identification. Software Requirement and Specification are described in the Section 3.2. Section 3.3 describes summary of the chapter.

3.1 Represent Requirement Collection And Identification

Requirement collection is the process which is used to gather, analyze, and documentation and reviews the requirements. Requirements describe what the system will do in place of how. In practical application, most projects will involve some combination of these various methods in order to collect a full set of useful requirements. Requirements collection is initiated when the project need is first identified and the project "solution" is to be proposed. Requirements refinement continues after the project is "selected" and as the scope is defined, aligned and approved.

3.2 Software Requirement Specification (SRS)

Software Specification will provide a broad understanding of the requirement specification of this system. Also, understand features of this system along with the requirements. Software Requirement Specification documents guide the developers in the development process and it will help to reduce the ambiguity of the requirements provided by the end-user. It's used to provide critical information to multiple teams development, quality assurance, operations, and maintenance. This keeps everyone on the same page.

3.2.1 Production Feature

The product features are high level attributes of a software or product such as software performance, user-friendly interface, security portability, etc. These attributes are defined according to the product, in this case, a software product. They are as follows:

- The user will be able to upload the New Article, Title, Image to be analysed

- The user will be able to view the results of submitted New Article, Title, Image

- The user will be able to integrate this system with any working browser Systems.

3.2.2 Operating Environment

The software will operate within the following environment:

- Operating System: Standard processor with a speed of 1.6 GHz

- Any browser supporting HTML5 and JavaScript

- Any system with at least 2GB RAM

- System with processor Intel Pentium 4 or later

3.2.3 Assumption

- It is assumed that the web portal will load and render correctly and as expected on the operating machine.

- It is assumed that the user will have a working internet connection with sufficient.

- It is assumed that the user is able to upload Text , images through web interface.

- It is assumed that the user will upload Information as Proper manner internet speed.

3.2.4 Functional Requirement

Functional requirements are the functions which are expected from the software or platform. Functional requirements along with requirement analysis help identify missing requirements. They help clearly define the expected system service and behavior.

3.2.5 Non Functional Requirement

Non-functional Requirement is mostly quality requirement. That stipulates how well the portal does, what it has to do. Other than functional requirements in practice, this would entail detail analysis of issues such as availability, security, usability and maintainability. Non-functional requirements are as follows:

3.2.6 External Interfaces

- **User Interfaces**

  The proposed system has several options for users to interact with. Following are the user:

  - Web-application (GUI)
  - Command Line Interface (Terminal based interface)

    The web application will be available so that the users will be able to upload New digital image, Image through a simple web-page will be available so that the user will be able to integrate it with other systems easily.

- **Software Interfaces**

  The only software interface required for this project is the Application Programming Interface with the VS Code Editor which will then process the Data. This software interface will run on local server along with the VS Code extension live server .

3.3 Summary

In the chapter, Analysis was presented which included the hardware and software requirements, functional and non-functional requirements and the software requirements specification (SRS) as well. In the next chapter, Design is described along with various UML diagrams.

Chapter 4

Design

Design is the activity to design and model the various component of software system. The system design provides the understanding and procedural details necessary for implementing the system. Design is helpful for a better understanding of the project. It contains the UML diagrams, data flow diagrams. UML is a modeling language which is used to document the object-oriented analysis and design. The organization of this Chapter is as follows. Section 4.1 describes the system architecture of the project. DFD of the project are represented in Section 4.2. Section 4.3 represents UML Diagrams (Use case, Class, Sequence, Component, Deployment, State chart, Activity diagram, Class Diagram, Component Diagram, etc.) of the project. Finally, the Summary is described in last Section 4.4.

4.1 System Architecture

Systems Architecture is a generic discipline to handle objects (existing or to be created) called "systems", in a way that supports reasoning about the structural properties of these objects. The system architecture is the conceptual model that defines the structure, behavior and more views of a system. An architecture description is a formal description and representation of a system. It provides broad understanding of the portal. In the system architecture database provide the functionality like get information, select criteria, etc. to users.
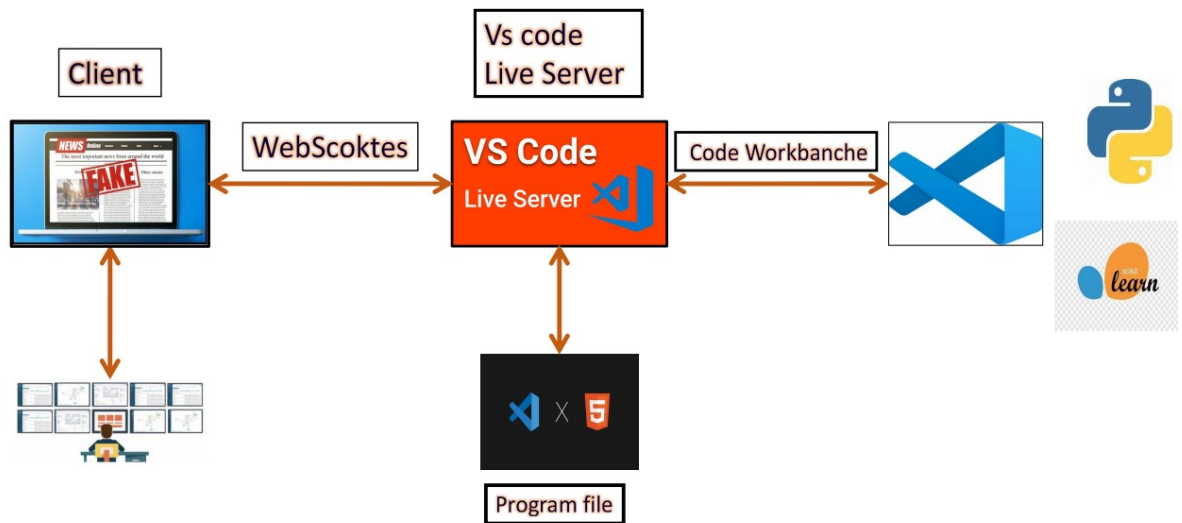
diagram

4.2 Data Flow Diagram

A data flow diagram (DFD) is a graphical representation of the 'flow' of data through an information system, modelling its process aspects. A DFD is often used as a preliminary step to create an overview of the system, which can later be elaborated. DFDs can also be used for the visualization of data processing (structured design). A DFD shows what kind of information will be input to and output from the system, where the data will come from and go to, and where the data will be stored.

4.2.1 Level 0 DFD

Level 0 contains one input and one output. The system provides information to the user means system is input and the user is output. Following Figure shows Level 0 DFD of project.
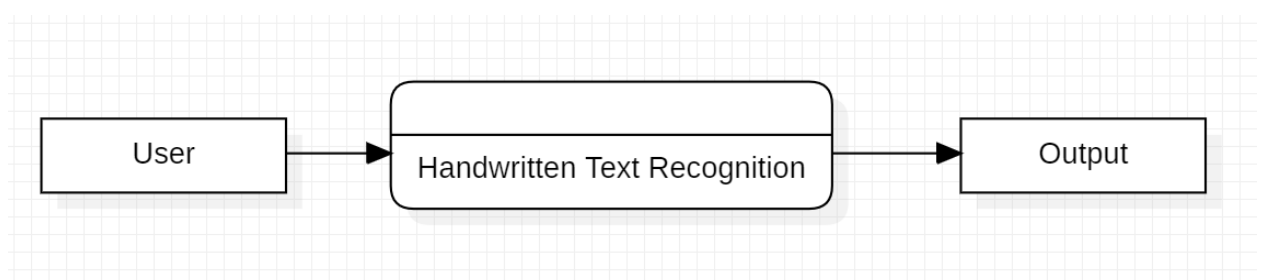
4.2.2 Level 1 DFD

A level 1 DFD notates each of the main sub-processes that together form the complete system. We can think of a level 1 DFD as an "exploded view" of the context diagram. Following Figure shows Level 1 DFD of project.
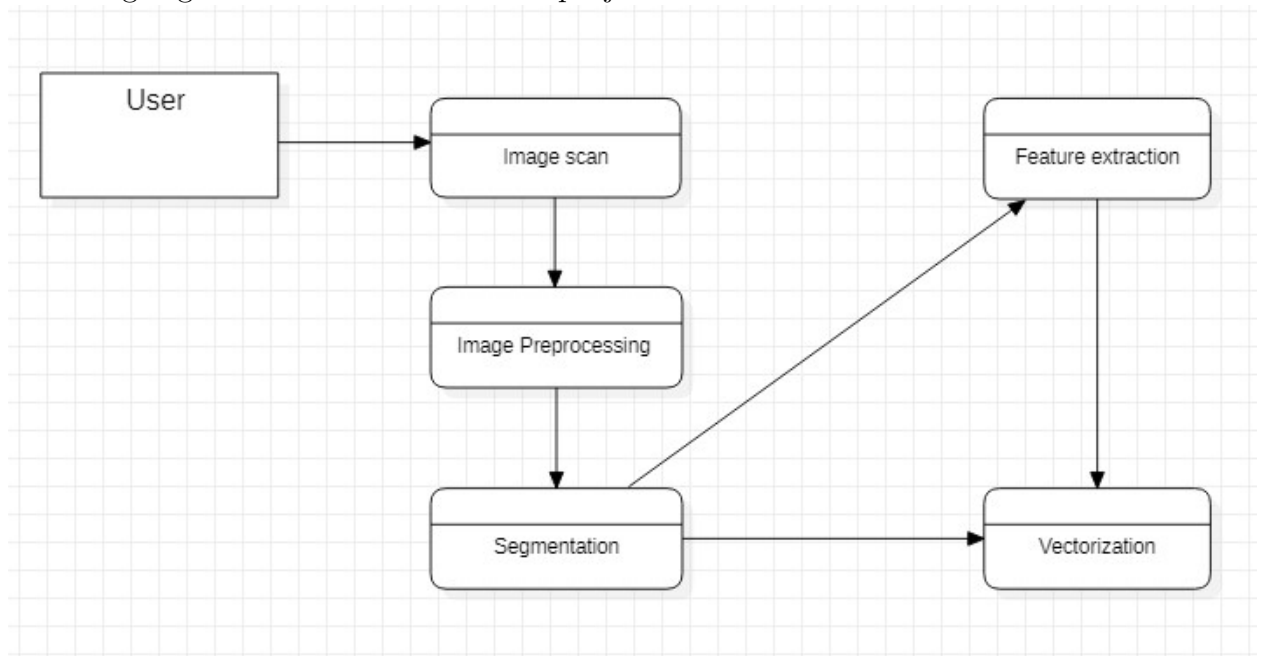
4.2.3 Level 2 DFD

A level 2 data flow diagram offers a more detailed look at the processes that make up an information system than a level 1 DFD does. It can be used to plan or record the specific makeup of a system. Following Figure shows Level 2 DFD of project.
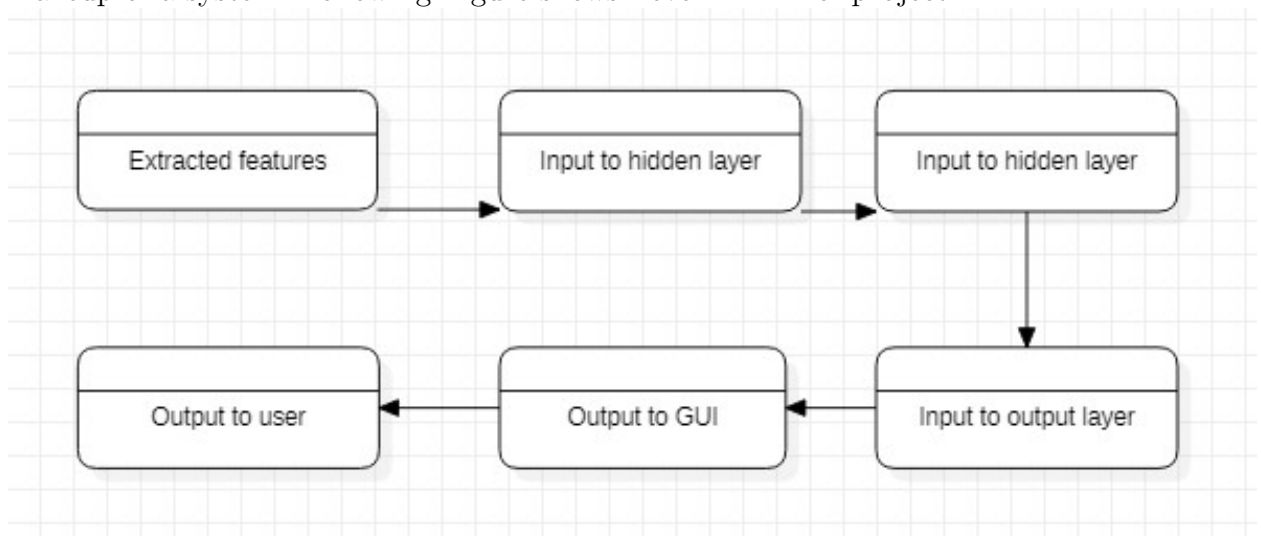
4.3　UML Diagrams

A UML diagram is a diagram based on the UML (Unified Modeling Language) with the purpose of visually representing a system along with its main actors, roles, actions, artifacts or classes, in order to better understand, alter, maintain, or document information about the system.

4.3.1　UseCase Diagram

Use case diagram shows the interaction between Use case which represents system functionality and actor which represent the people or system.
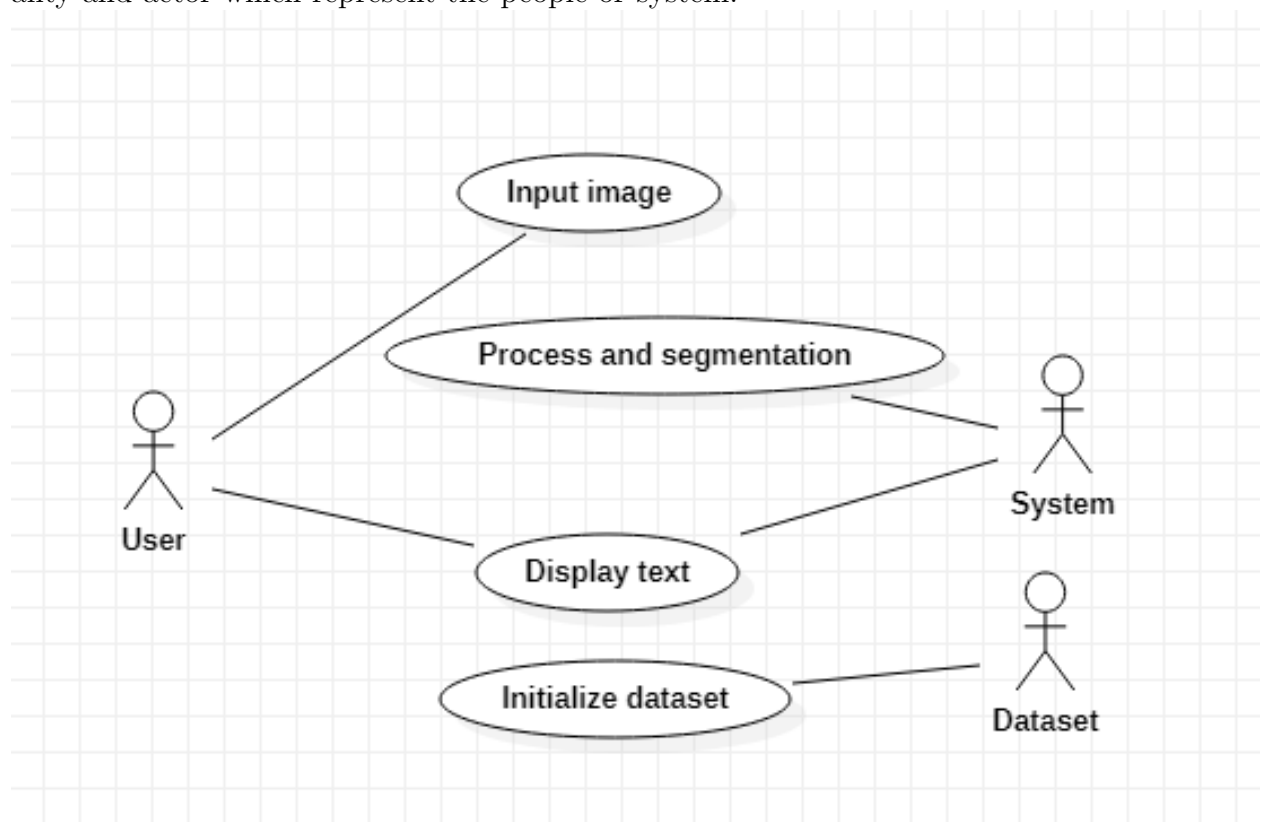
4.3.2　Sequence Diagram

The sequence diagram shows the flow of functionality through Use case. A sequence diagram is a type of interaction diagram because it describes how—and in what order—a group of objects works together. These diagrams are used by software developers and business professionals to understand requirements for a new system or to document an existing process.

interaction SequenceDiagram1

: User   System   Text

1 : Select image
2 : Upload image
3 : Preprocessing
4 : Segmentation
5 : Recognize text
6 : recognising
7 : Display text
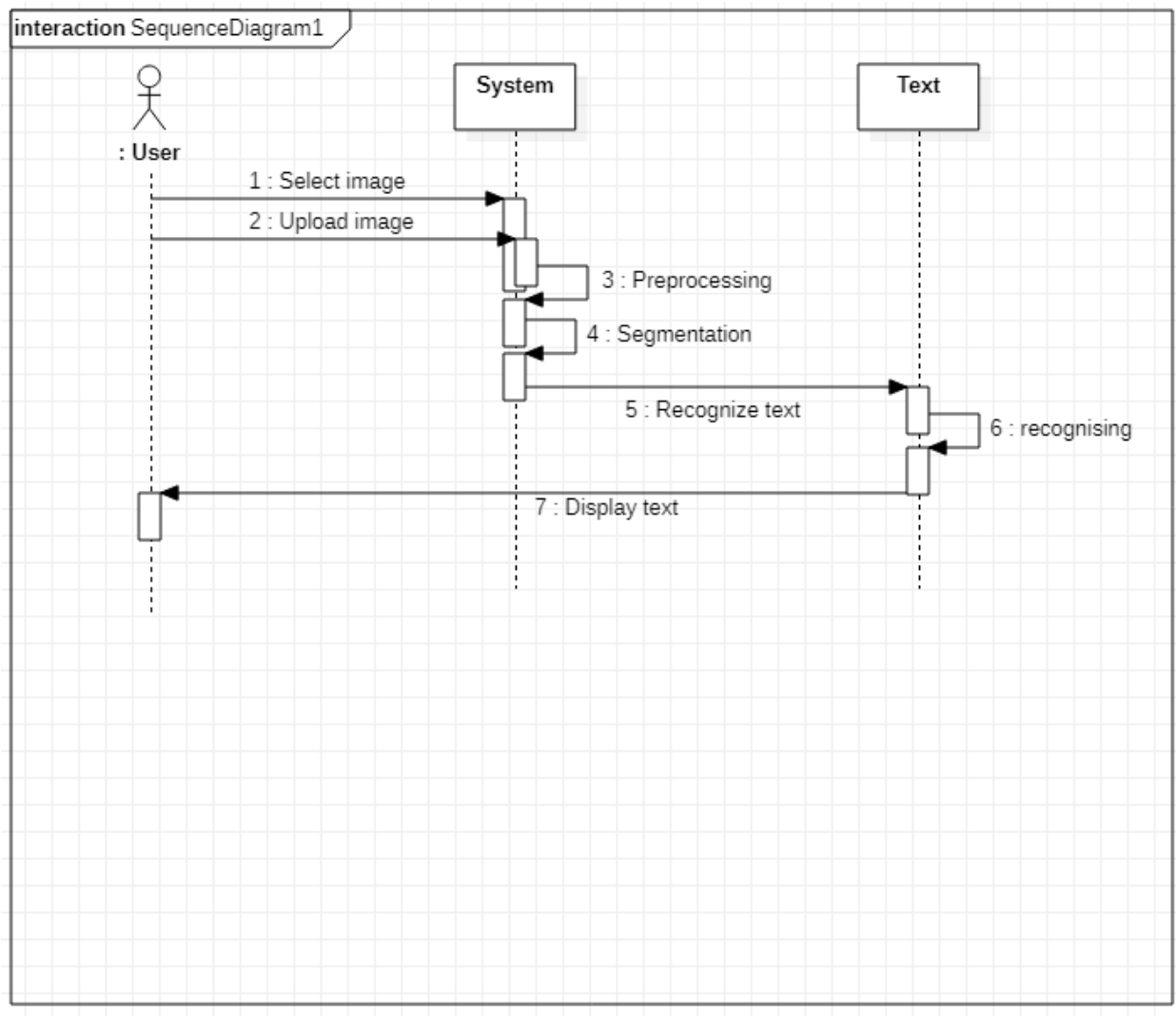
4.3.3 Collaboration Diagram

A collaboration diagram, also known as a communication diagram, is an illustration of the relationships and interactions among software objects in the Unified Modeling Language (UML). These diagrams can be used to portray the dynamic behavior of a particular use case and define the role of each object.

interaction CommunicationDiagram1

User

1 : Select image
2 : Upload image

3 : Preprocessing
4 : Segmentation

System

7 : display text

5 : recognise text

Text

6 : recognising
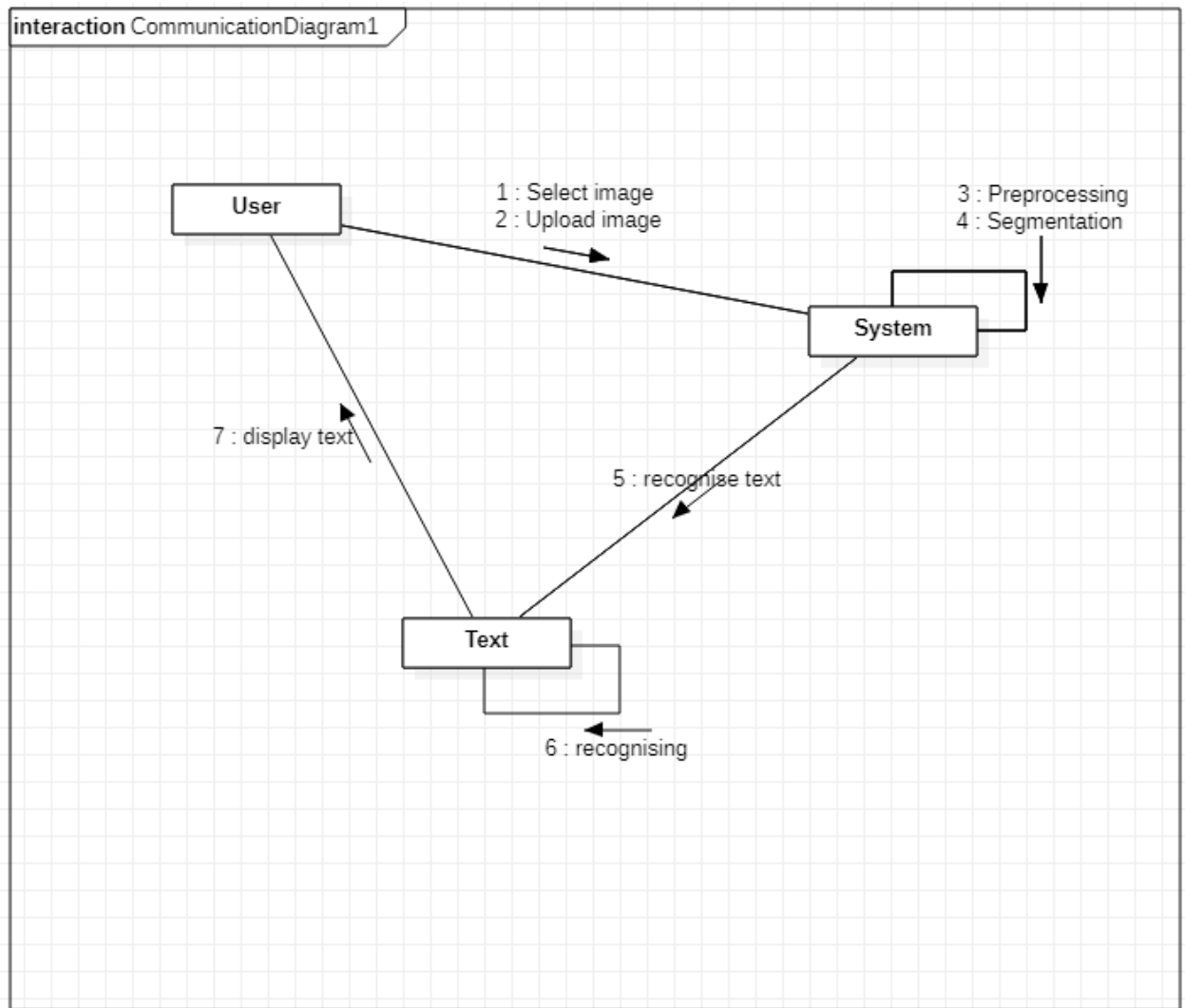
4.3.4 Class Diagram

The class diagram is the main building block of object-oriented modeling. It is used for general conceptual modeling of the structure of the application, and for detailed modeling translating the models into programming code. Class diagrams can also be used for data modeling.
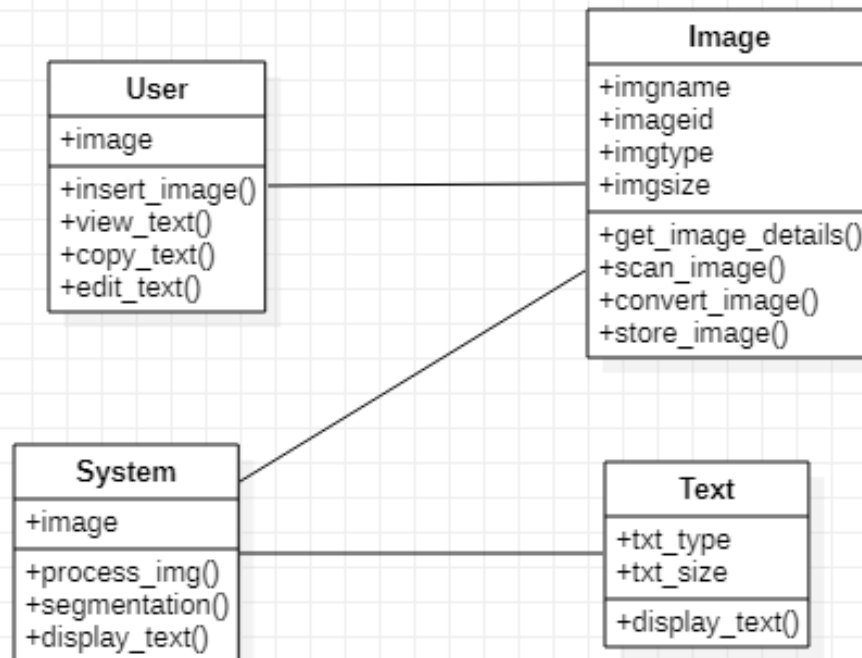
4.3.5   StateTransition Diagram

The name of the diagram itself clarifies the purpose of the diagram and other details. It describes different states of a component in a system. The states are specific to a component/object of a system. A State Chart diagram describes a state machine. State machine can be defined as a machine which defines different states of an object and these states are controlled by external or internal events.
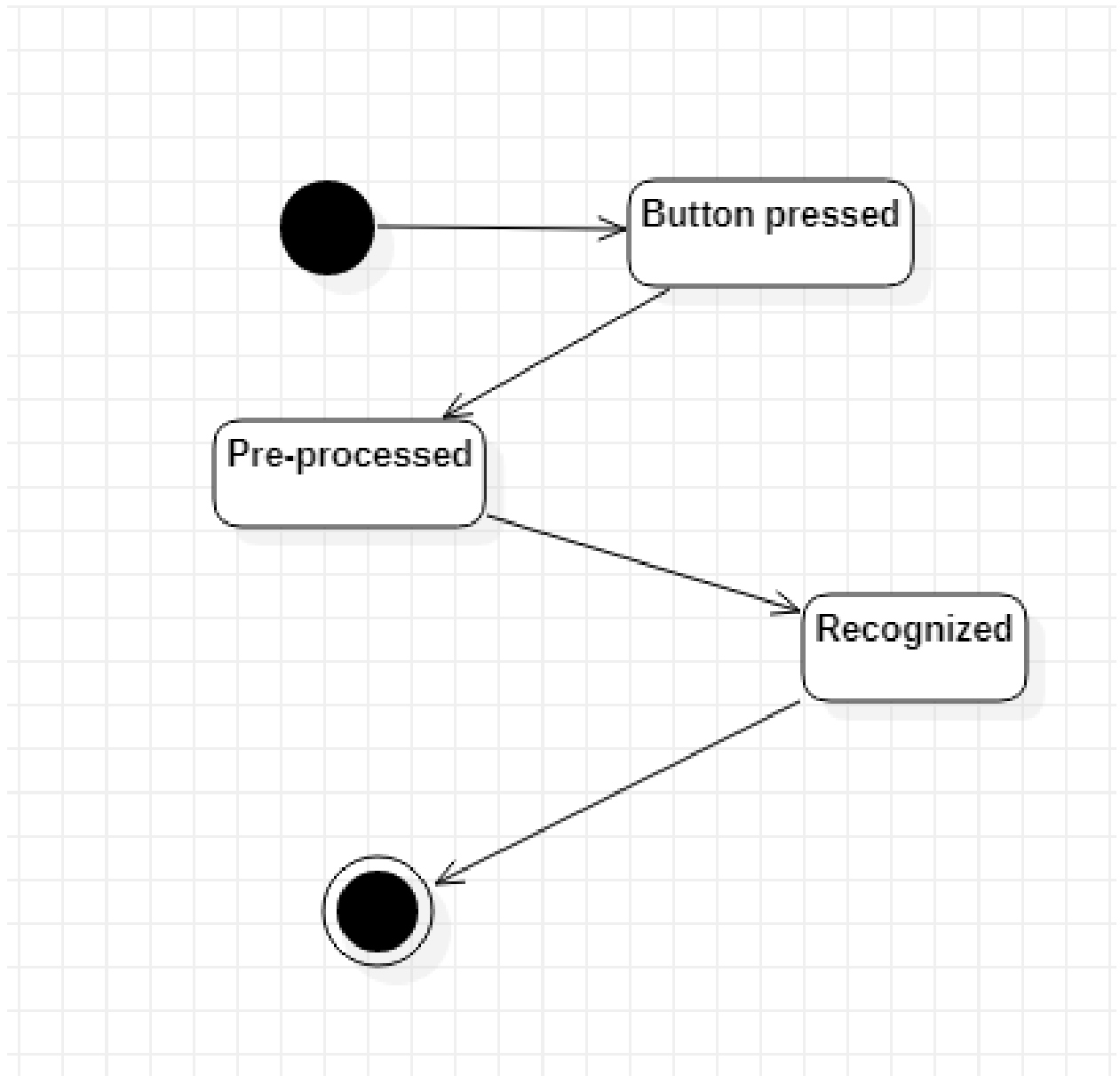
4.3.6 Component Diagram

A component diagram, also known as a UML component diagram, describes the organization and wiring of the physical components in a system. Component diagrams are often drawn to help model implementation details and double-check that every aspect of the system's required function is covered by planned development.
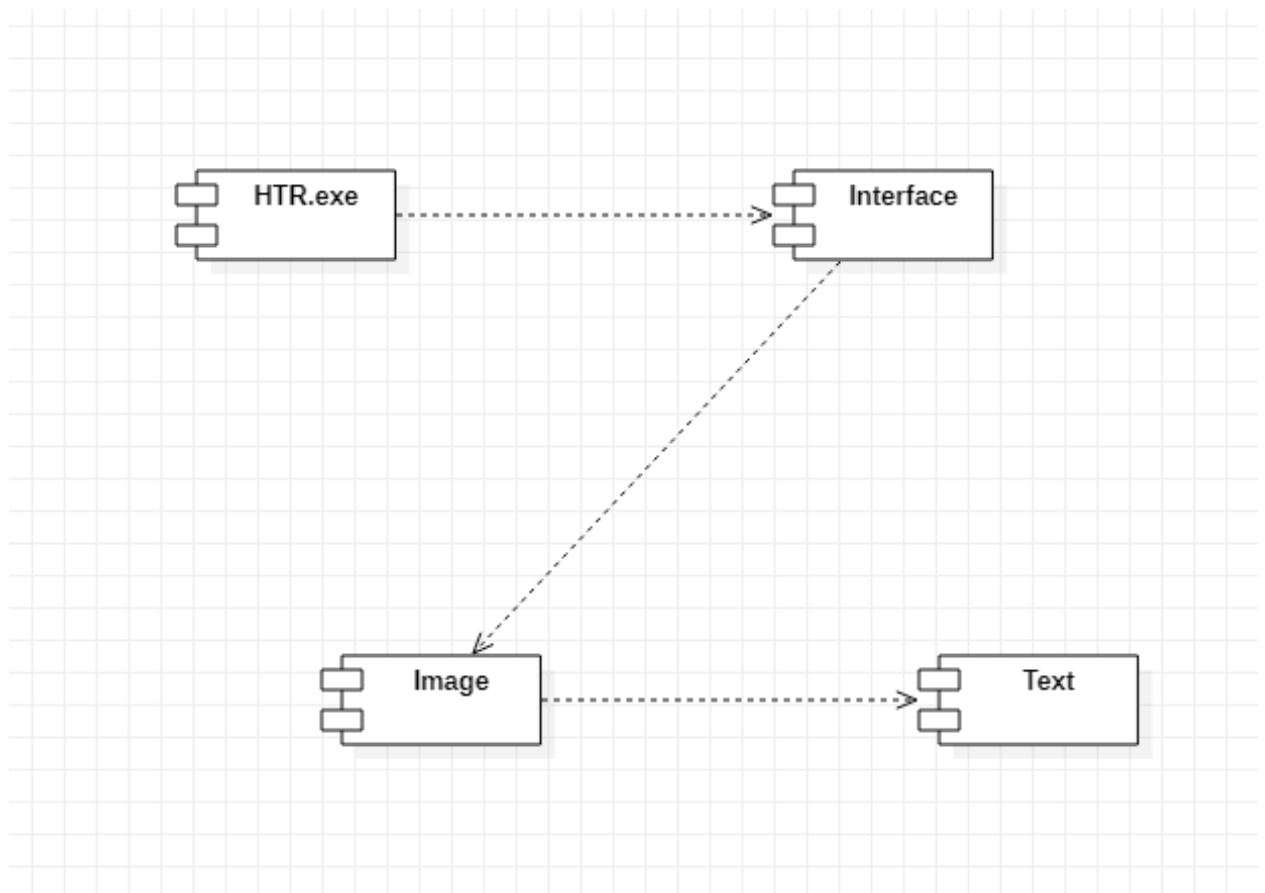
4.3.7　Deployment Diagram

A deployment diagram is a UML diagram type that shows the execution architecture of a system, including nodes such as hardware or software execution environments, and the middleware connecting them. Deployment diagrams are typically used to visualize the physical hardware and software of a system.
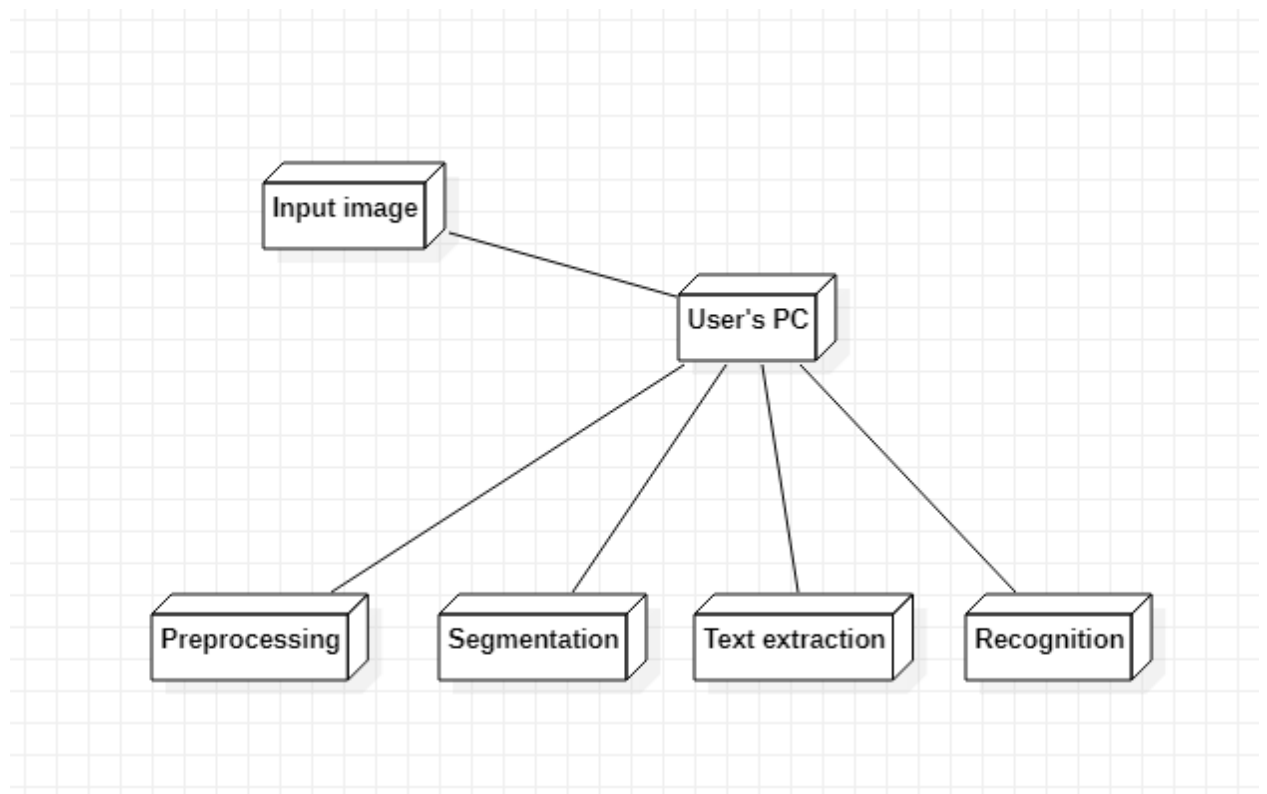
4.3.8 Summary

Detailed design of project has been described in this chapter including the Data Flow Diagrams and the UML Diagram explaining all the design details of the project. Conclusion of the project has been explained in the next chapter.

4.3.9 Output
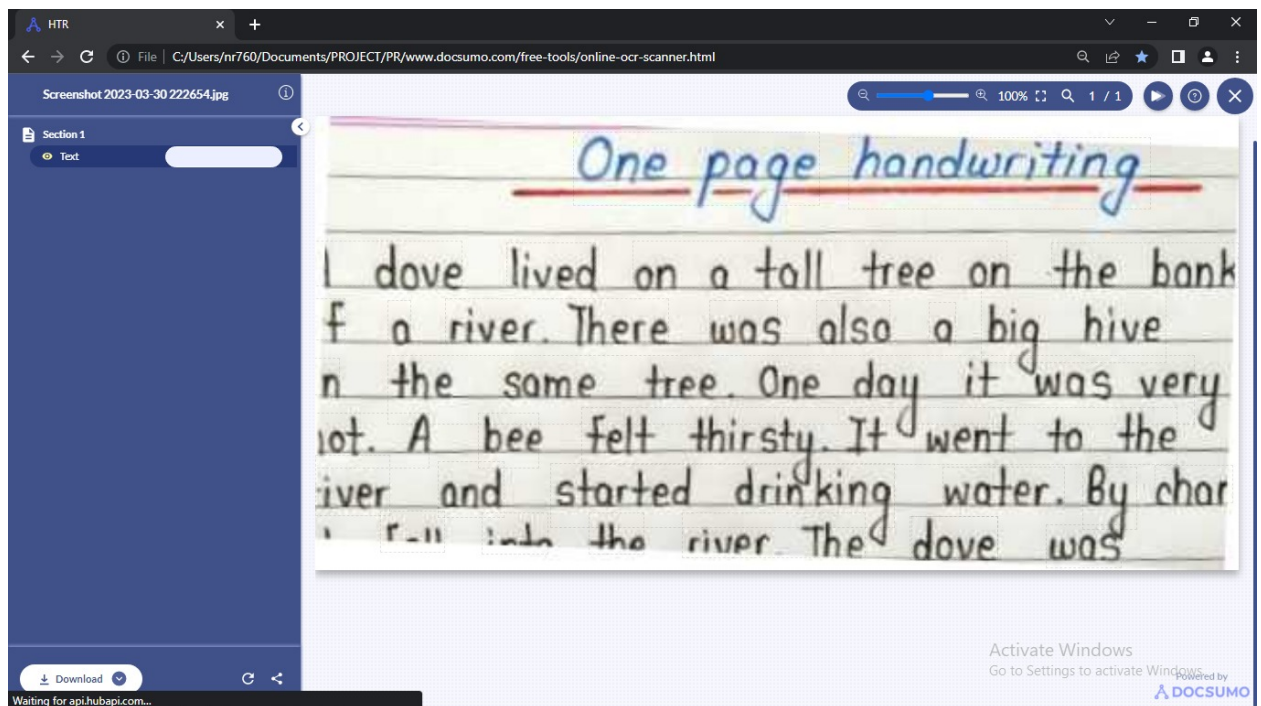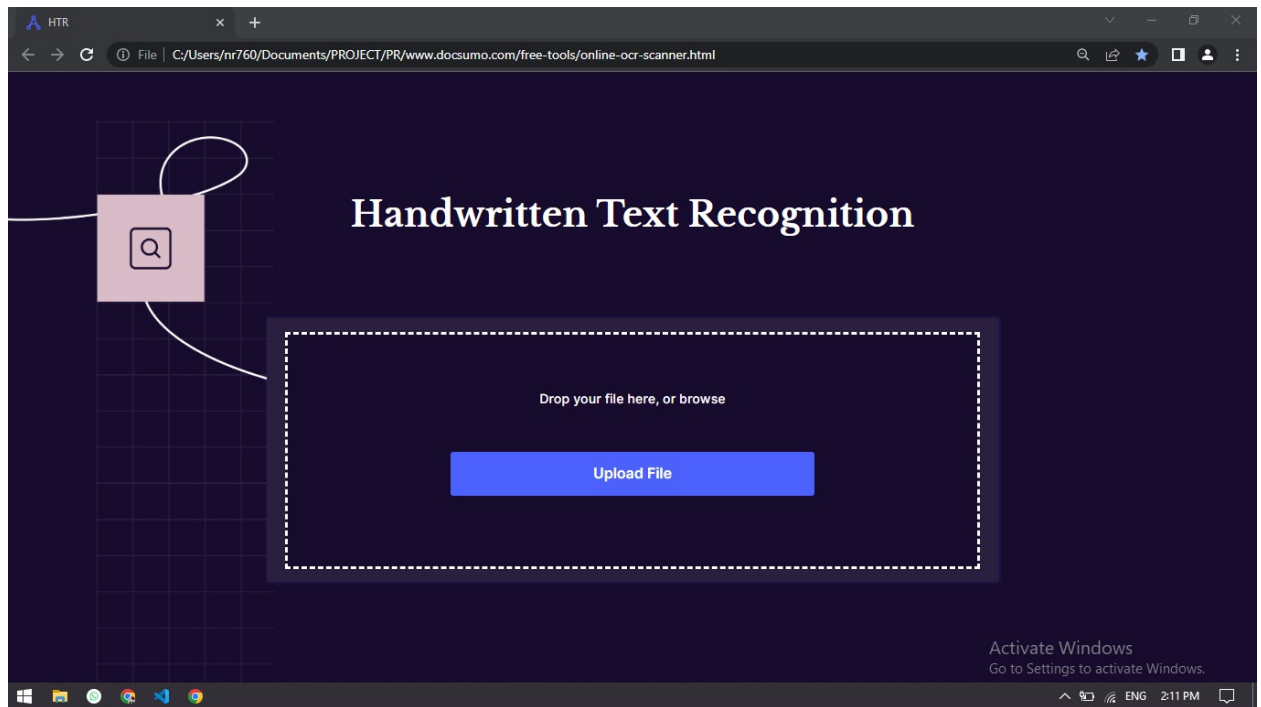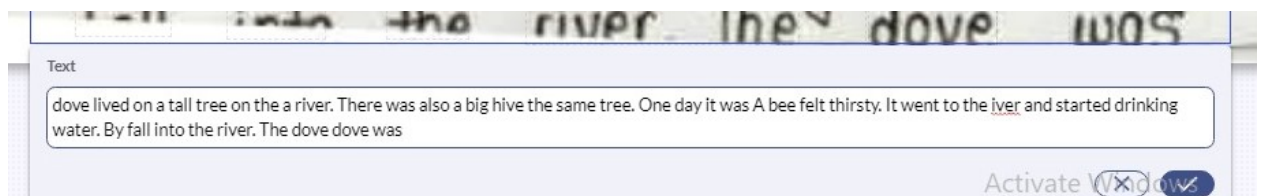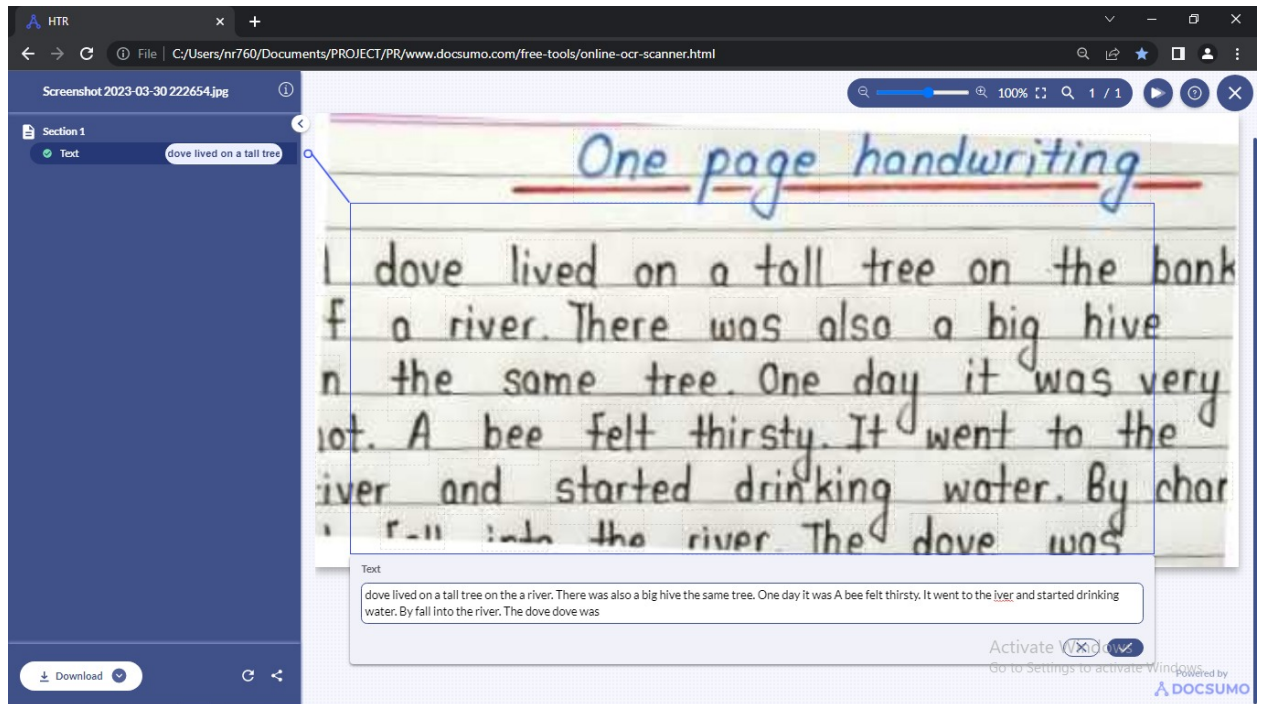
4.4 code

```html
<body>
  <div class="exit-intent-popup">...
  </div>
  <div class="modal-wrapper">...
  </div>
  <div id="free-tool-modal-wrapper" style="opacity:0" class="free-tool-modal-wrapper">...
  </div>
  <div id="review-popup" class="popups-2"></div>
  <div id="content-tool" class="content-tool">

    <div class="main-wrapper" style="background-image:url(&quot;https://global-uploads.webflow.com/636bdbebfc681f083e923f81/63a753bfc7b14439be15bede_Free%20OCR.png&quot;)"
      <div style="background-image:url(&quot;https://global-uploads.webflow.com/636bdbebfc681f083e923f81/63a753bfc7b14439be15bede_Free%20OCR.png&quot;)" class="section b
        <div class="site-wrapper">
          <div class="section-container">
            <div class="ft-hero-text-container">...
            </div>
            <div class="div-block-287">
              <div class="form-block-2 w-form">
                <form id="email-form-4" name="email-form-4" data-name="Email Form 4" method="get" class="toolform">
                  <div id="firsttoolcard" class="firsttoolcard">
                    <div id="error-message" class="errormessage">This is some text inside of a div block.</div>
                    <div class="text_span"><strong class="bold-text-7">Drop your file here, or browse</strong></div>
                    <div class="text-block-58"></div><a id="pdf_button" href="#" class="pdf_button w-button">Upload File</a></div>
                    <div id="downloadPdf" class="downloadcard">...
                  </div>
                  <div id="secondtoolcard" class="secondtoolcard">
                    <div class="div-block-230">
                      <div class="loading">
                        <div class="div-block-35">
                          <div id="timer-app" class="div-block-252"></div>
                        </div>
                      </div>
                    </div>
                    <div id="info-message" class="infomessage">Please wait for few seconds. Your document is getting processed.</div>
                  </div>
                  <div id="recaptcha-err-msg" class="recaptcherrmsg">
                    <div class="div-block-235"><img src="https://global-uploads.webflow.com/6358e85b37c8c34b8bd103ae/639ab33bdca83b6457754cd5_paper.svg" lo
                    <div class="div-block-236 error">
                      <div id="error-message1" class="errormessage _2nd">Error! Your file couldnot be processed further due to some internal server error
                      <div tooltipster="top" title="Reset" class="tooltip-2">
                        <div id="reset-btn" class="resetbtn"><img src="https://global-uploads.webflow.com/6358e85b37c8c34b8bd103ae/639ab33bdca83b5c7175
                      </div>
                    </div>
                  </div>
                  <div id="thirdtoolcard" class="thirdtoolcard">

                    <div class="div-block-237">
                      <div id="download-err" class="text-block-66">File Couldn&#x27;t be downloaded ! Try again</div>
```

```javascript
        <script defer type="text/javascript">
          let realFileBtn = document.getElementById("real_file");
          let customBtn = document.getElementById("pdf_button");
          let customTxt = document.getElementById("custom_text");
          let errMsg = document.getElementById("error-message");
          let errMsg1 = document.getElementById("error-message1");
          let firsttoolcard = document.getElementById("firsttoolcard");
          let secondtoolcard = document.getElementById("secondtoolcard");
          let thirdtoolcard = document.getElementById("thirdtoolcard");
          let recaptchaErrMsg = document.getElementById("recaptcha-err-msg");
          let resetBtn = document.getElementById("reset-btn");
          let resetBtn1 = document.getElementById("reset-btn1");
          let resetBtn2 = document.getElementById("reset-btn2");
          let downloadErr = document.getElementById("download-err");
          let file, review_url, data, files;
          let reviewBlock = document.getElementById("review-block");
          let check = true;
          let reviewErr = document.getElementById("review-err");
          let freeToolForm = document.getElementById("free-tool-form");
          let downloadTextSingle = document.getElementById("downloadtext-single");
          let downloadBtnCSV = document.getElementById("download-csv");
          let downloadBtnJSON = document.getElementById("download-json");
          let downloadBtnExcel = document.getElementById("download-excel");
          let downloadDropdown = document.getElementById("download-dropdown");
          //elements required for function thirdtoolcard
          let reviewImage = document.getElementById("reviewImg");
          let fileImageExcel = document.getElementById("file-image-xlsx");
          let fileImageJSON = document.getElementById("file-image-json");
          let fileImageCSV = document.getElementById("file-image-csv");
          let fileNameField = document.getElementById("file-name");
          let copyLink = document.getElementById("copy-link");
          let copyLinkText = document.getElementById("copy-link-text");
          let ratingSection = document.getElementById("rating-section");
          let downloadPdf = document.getElementById("downloadPdf");
          let trySampleBlock = document.getElementById("try_sample-block");
          let requestData = null;
          let requestButtonData = null;
          let downloadPDFText = document.getElementById("downloadpdf-text");
          //let loader = document.getElementById("loader");

          var path = window.location.pathname;
          var blogParams = path.split("/");

          document.addEventListener("DOMContentLoaded", function() {
            let reviewPopup = document.getElementById("review-popup");
            let ratingSection = document.getElementById("rating-section");
            let downloadPdfButton = document.getElementById("downloadPdfButton");
            let startOver = document.getElementById("startover");
```

```
<div class="html-embed-6 w-embed w-iframe w-script">
    <script defer type="text/javascript">
        var pathName = window.location.pathname;
        var path = pathName.split("/");
        var count = path.length;
        var documentType = path[count - 1];

        async function secondtoolcard1(freefile) {
            firsttoolcard.style.display = "none";
            secondtoolcard.style.display = "flex";
            trySampleBlock.style.display = "none";
            apicall(freefile);
        }

        async function thirdtoolcard1(apiRes1) {
            var path = window.location.pathname;
            var blogContentArray = path.split("/");
            secondtoolcard.style.display = "none";
            thirdtoolcard.style.display = "block";
            if (documentType === "others__frYRR") {
                downloadTextSingle.style.display = "flex";
                downloadDropdown.style.display = "none";
                data = apiRes1.data.document;
                const filename = data[0].title;
                review_url = data[0].review_url;
            } else {
                data = apiRes1.data.document;
                const filename = data[0].title;
                review_url = data[0].review_url;
                popUpUrl(review_url);
            }
            if (/Android|webOS|iPhone|iPad|iPod|BlackBerry|IEMobile|Opera Mini/i.test(navigator.userAgent)) {
                resetBtn1 = resetBtn2;
                reviewImage.classList.add("style.img.filter=invert(1)");
            }
            var token = review_url.substring(review_url.indexOf("=") + 1);
            freeToolForm.addEventListener('submit', function(event) {
                disableBtn(requestButtonData);
                if (requestData == "single_text") {
                    downloadApiTextCall(data);
                } else {
                    downloadApiCall(data, requestData)
                }
                toggle(requestButtonData);
            });
            copyLink.addEventListener("click", function() {
                copyToClipboard(review_url);
```

```
<script {
    try {
        hubspot.require(['enviro', 'PortalIdParser'], function(Enviro, PortalIdParser) {
            var enabledForQA = false;
            var effectiveSampleRate = window.newRelicSamplingOverride != null ? parseFloat(window.newRelicSamplingOverride, 10) : 0.13;
            var isSampled = effectiveSampleRate == null || effectiveSampleRate === 1 || Math.random() <= effectiveSampleRate;
            var isDeployed = Enviro.deployed('newrelic');
            var disableNRCookies = false;
            if (isDeployed) {
                try {
                    isDeployed = localStorage.getItem('NEWRELIC_DEPLOYED') !== "false";
                } catch (e) {}
            }
            if (disableNRCookies) {
                window.NREUM || (NREUM = {});
                NREUM.init = {
                    privacy: {
                        cookies_enabled: false
                    },
                    ajax: {
                        deny_list: ["bam-cell.nr-data.net"]
                    }
                };
            } else {
                window.NREUM || (NREUM = {});
                NREUM.init = {
                    privacy: {
                        cookies_enabled: true
                    },
                    ajax: {
                        deny_list: ["bam-cell.nr-data.net"]
                    }
                };
            }
            if (isSampled && isDeployed && (Enviro.getInternal('newrelic') === 'prod' || enabledForQA)) {

                var licenseKey = 'f9d051f404';
                var applicationID = '205242107';
                window.NREUM || (NREUM = {}), __nr_require = function(t, e, n) {
                    function r(n) {
                        if (!e[n]) {
                            var o = e[n] = {
                                exports: {}
                            };
                            t[n][0].call(o.exports, function(e) {
                                var o = t[n][1][e];
                                return r(o || e)
                            }, o, o.exports)
```

```javascript
                filter: invert(1);
            }
        </style>

        <script defer type="text/javascript">
            async function downloadApiCall(data, type) {
                const baseUrl = "https://app.docsumo.com";
                downloadErr.style.display = "none";
                let serverErr;
                let thirdtoolcard = document.getElementById("thirdtoolcard");
                const review_url = data[0].review_url;
                var token = review_url.substring(review_url.indexOf("=") + 1);
                var docId = data[0].doc_id;
                let url = new URL(`${baseUrl}/api/v1/mew/documents/download/?type=${type}`);
                const body = {
                    doc_ids: [docId],
                };
                try {
                    let downloadApiRes;
                    await fetch(url, {
                        method: "POST",
                        headers: {
                            'token': token,
                            'Content-Type': 'application/json',
                        },
                        body: JSON.stringify(body),
                    })
                    .then((response) => response.json())
                    .then((data) => {
                        downloadApiRes = data;
                        const download_url = downloadApiRes.data.download_url;
                        download(download_url);
                    })
                    .catch((error) => {
                        serverErr = error;
                    });
                    if (serverErr || downloadApiRes.error != "") {
                        downloadErr.style.display = "block";
                        setTimeout(async function() {
                            downloadErr.style.display = "none";
                        }, 5000);
                    }
                } finally {
                    thirdtoolcard.style.display = "flex";
                }
            }

            let trySample = document.getElementById('try-sample');
            let captchaBlock = document.getElementById('captcha-block');
            let downloadTextBtn = document.getElementById('download-text');
            let infoMessage = document.getElementById('info-message');

            firsttoolcard.addEventListener("dragenter", dragenter, false);
            firsttoolcard.addEventListener("dragover", dragover, false);
```

```javascript
            function startTimer(timePeriod) {

                const TIME_LIMIT = timePeriod;
                let timePassed = 0;
                let timeLeft = TIME_LIMIT;
                let timerInterval = null;
                infoMessage.style.display = 'block';
                timer.style.display = 'flex';
                timerInterval = setInterval(() => {
                    timePassed = timePassed += 1;
                    timeLeft = TIME_LIMIT - timePassed;
                    timer.innerHTML = `
                        <span id="base-timer-label" class="base-timer__label">${formatTime(
                            timeLeft
                        )}</span>`;

                    if (timeLeft === 0) {
                        clearInterval(timerInterval);
                        onTimesUp();

                        timeLeft = timePeriod;
                        timePassed = 0;
                    }
                    document.getElementById("base-timer-label").innerHTML = formatTime(
                        timeLeft
                    );

                }, 1000);
            }

            function formatTime(time) {
                const minutes = Math.floor(time / 60);
                let seconds = time % 60;

                if (seconds < 10) {
                    seconds = `0${seconds}`;
                }
                return `${minutes}:${seconds}`;
            }

            //apicall for free tool upload document
        </script>
    </div>
    <div class="browser-extension-div"></div>
    <div class="ocr-video-cover w-condition-invisible">
        <div style="padding-top:56.17021276595745%" class="w-embed-youtubevideo"><iframe src="https://www.youtube.com/embed/uCu5sCsQ6qs?rel=0&amp;controls=1&amp;autoplay=0&amp;mute=0&amp;start=4" frameBorder="0" style=
            allowfullscreen="" title="OCR Chrome Extension by Docsumo| Data extraction from web-pages with one screenshot"></iframe></div>
    </div>
</div>
</div>
    <div data-w-id="175b68cb-ebb9-f83b-bd5d-dec5d569a3b5" class="nav-scroll-trigger"></div>
```

Chapter 5

Testing

5.1 Black Box Testing

Black Box Testing is not a type of testing; it instead is a testing strategy, which does not need any knowledge of internal design or code etc. As the name "black box" suggests, no knowledge of internal logic or code structure is required. The types of testing under this strategy are totally based/focused on the testing for requirements and functionality of the work product/software application. The base of the Black box testing strategy lies in the selection of appropriate data as per functionality and testing it against the functional specifications in order to check for normal and abnormal behavior of the system

5.1.1 Functional Testing

In this type of testing, the software is tested for the functional requirements. The tests are written in order to check if the application behaves as expected.

5.1.2 Stress Testing

The application is tested against heavy load such as complex numerical values, large number of inputs,large number of queries etc. which checks for the stress/load the applications can withstand.

5.1.3 Load Testing

The application is tested against heavy loads or inputs such as testing of web sites in order to find out at what point the web- site/application fails or at what

5.1.4 Recovery Testing

Recovery testing is basically done in order to check how fast and better the application can recover against any type of crash or hard- ware failure etc. Type or extent of recovery

---

is specified in the requirement specifications.

5.2   White Box Testing

White box testing strategy deals with the internal logic and structure of the code. White box testing is also called as glass, structural, open box or clear box testing. The tests written based on the white box testing strategy incorporate coverage of the code written, branches, paths, statements and internal logic of the code etc. In order to implement white box testing, the tester has to deal with the code and hence is needed to possess knowledge of coding and logic i.e. internal working of the code. White box test also needs the tester to look into the code and find out which unit/statement/chunk of the code is malfunctioning. at what

5.2.1   Unit Testing

The developer carries out unit testing in order to check if the particular module or unit of code is working fine. The Unit Testing comes at the very basic level as it is carried out as and when the unit of the code is developed or a particular functionality is built.

5.2.2   Branch Coverage

No software application can be written in a continuous mode of coding, at some point we need to branch out the code in order to perform a particular functionality. Branch coverage testing helps in validating of all the branches in the code and making sure that no branching leads to abnormal behavior of the application.

5.2.3   Security Testing

Security Testing is carried out in order to find out how well the system can protect itself from unauthorized access, hacking cracking, any code damage etc. which deals with the code of application. This type of testing needs sophisticated testing.

5.3   Manual Testing

It is oldest and most rigorous types of testing it is performed by human sitting in front of a computer carefully going through application screens, trying various usage and input combination, comparing the results to be expected behaviour and recording and observations about the project. The manual testing corresponding to project is to test whether that system give accurate classification of dataset. Manual testing is testing of the software

---

where tests are executed manually by a QA analyst. It is performed to discover bugs in software under development. In manual testing, the tester checks all the essential features of the given application or software. In this process, the software testers execute the test cases and generate the test reports without the help of any automation software testing tools. It is a classical method of all testing types and helps and bugs in software systems. It is generally conducted by an experienced tester to accomplish the software testing process.

• The initial investment in the Manual testing is comparatively lower. • Manual testing is not as accurate because of the possibility of the human errors. • Manual testing proves useful when the test case only needs to run once or twice. • No need for programming in Manual Testing. • While testing a small change, an automation test would require coding which could be time consuming. While you could test manually.

5.4 Automated Testing

In Automated Software Testing, testers write code test scripts to automate test execution. Testers use appropriate automation tools to develop the test scripts and validate the soft ware. The goal is to complete test execution in a less amount of time. Automated testing entirely relies on the pre scripted test which runs automatically to compare actual result with the expected results. This helps the tester to determine whether or not an application performs as expected. Automated testing allows you to execute repetitive task and regression test without the intervention of manual tester. Even though all processes are performed automatically, automation requires some manual export to create initial testing scripts. • The initial investment in the automated testing is higher. • Automated testing is a reliable method, as it is performed by tools and scripts. There is no testing fatigue. • Programming knowledge is a must in automation testing. • Automation testing is useful when frequently executing the same set of test cases. • Testing coverage can be increased because automation testing tool never forgets to check even the smallest unit.

Chapter 6

Conclusion

We have learning for recognition of English characters. It has been found that recognition of handwritten character becomes difficult due to presence of odd characters or similarity in shapes for multiple characters. Scanned image is pre-processed to get a cleaned image and the characters are isolated into individual characters.

Preprocessing work is done in which normalization, filtration is performed using processing steps which produce noise free and clean output. Managing our evolution algorithm with proper training, evaluation other step wise process will lead to successful output of system with better efficiency. Use of some statistical features and geometric features through neural network will provided better recognition result of English characters. This work will be helpful to the researchers for the work towards other script.

Bibliography

1. https://towardsdatascience.com/build-a- handwritten-text-recognition-system-using -tensorflow-2326a3487cd5

2. https://www.slideshare.net/chiranjeeviadi/hand-written-character-recognition-using -neural-networks

3. https://nanonets.com/blog/handwritten-character-recognition/

4. Chirag Patel, Ripal Patel, Palak Patel, "Handwritten Character Recognition Using Neural Networks", International Journal of Scientific Engineering Research Volume 2, Issue 5, May-201 1.