**1. INTRODUCTION**

**1.1 Introduction to the system**

**1.1.1 Project Title**

Digital File Manipulation.

**1.1.2 Category**

Windows Application.

**1.1.3 Overview**

The application provides various manipulation operations on different types of digital files like Images, Audio, Video, Text, Doc, Pdf, xls, csv.

**1.2 Background**

**1.2.1 Introduction of the Company**

Not Applicable.

**1.2.2 Brief note on Existing System**

In existing system, when we cut text from image the file is saved as image itself because we are cropping the image. To highlight the text in image or pdf , we have to use external tools. Image color recognition is difficult in existing system and a lot of time is consumed by existing system. Symbol and Mathematical equations are difficult recognize. User has to use different existing software to achieve the tasks.

* 1. **Objective of the System**
* Objectives are the pre-determined goals or outcomes of the system process.
* The main objective of the system is to develop an application which can be used to manipulate various digital files like image, video, audio, text, doc, pdf, xls and csv. And to provide one stop solution for different types of digital file manipulations.
  1. **Scope of the System**
* Scope is the limitation that a process faces from the beginning to the end.
* Image recognition, symbol and mathematical equation recognition is difficult.
* Clarity of image is must for text recognition.
* Clarity Audio and Video is must for extracting text and to extract audio from video.
  1. **Structure of the System**

**1.5.1 Text Manipulation in an Image**

**1.5.1.1 Extracting Text from Images**

The user provides a clear image as input and the application retrieves the text in that image and saves it in a text file.

**1.5.1.2 Extracting Text from Region of Interest**

The user provides a clear image as input and selects a region in that image and the application will retrieve the text in that image.

**1.5.1.3 Highlighting Text**

The user provides a clear image and a word or a sentence as input, and the application highlights that word/sentence in the image ,else gives a prompt message telling ‘There is no such word or sentence’.

**1.5.2 Spread Sheet Manipulation**

**1.5.2.1 Excel to Text**

The user provides an Excel file as input and the application converts it into a text file and gives it as output.

**1.5.2.2 Csv to Text**

The user provides a Csv file as input and the application converts it into a text file and gives it as output.

**1.5.2.3 Excel to Csv**

The user provides an Excel file and the excel sheet name as inputs and the application converts it into a Csv file and gives it as output.

**1.5.2.4 Search a Column**

The user provides an Excel file required column name as input and the application converts the specified column data into a text file and gives it as output.

**1.5.3 Video-Audio Manipulation**

**1.5.3.1 Extracting Text from Video**

The user provides a video as input and the application retrieves the words spoken in that video and writes it on a text file, which is the output.

**1.5.3.2 Extracting Audio from Video**

The user provides a video as input and the application retrieves the audio in that video and provides it as output.

**1.5.2.3 Extracting Text from Audio**

The user provides an audio as input and the application retrieves the words spoken in that audio and writes it on a text file, which is the output.

**1.5.4 File Manipulation**

**1.5.4.1 Pdf to Text**

The user provides a Pdf file as input and the application converts it into a text file and gives it as output.

**1.5.4.2 Docx to Text**

The user provides a Docx file as input and the application converts it into a text file and gives it as output.

**1.5.4.3 Pdf to Docx**

The user provides a pdf file as input and the application converts it into a Docx file and gives it as output.

**1.5.4.4 Text to Docx**

The user provides a Text file as input and the application converts it into a Docx file and gives it as output.

* 1. **System Architecture**

**USERS**

**GUI**

**User Level**

Commands

Console

Tools

Buttons

Menu

**Digital File Manipulation Application**

File Manipulation

Text Manipulation

**Application**

Video-Audio Manipulation

Spread Sheet Manipulation

**Level**

Mass Storage Device

Image Processing Tools/Packages/Functions/Procedures

**System**

**Operating System**

**Level**

**Processing Hardware**

**1.7 End User**

The End-User can be anyone who requires the functionalities provided by this system. They can give image as an input and perform operations like extracting text, highlighting text. The user can give video and audio as input and perform operations like extracting audio from video, extracting text from video and extract text from audio. They can also input a doc/pdf/text file and convert them to text/doc/pdf.

**1.8 Software/Hardware need for the development**

**Software:**

Python, Pytessaract, ffmpeg, Windows OS.

**Hardware:**

The system must have more than 2GB of RAM.

**1.9 Software/Hardware need for the implementation**

**Software:**

Pytessaract, ffmpeg, Windows OS.

**Hardware:**

The system must have more than 2GB of RAM, and the screen resolution must be more than 1600X900 pixels.

**2. SOFTWARE REQUIREMENT SPECFICAION**

**2.1. Introduction**

Software Requirement Specification is a description of a software to be developed. It layouts functional and non-functional requirements and may include the set of use cases that describes the user interaction that they must provide. It helps the client to understand their own needs. It provides justification of final product. It helps to bridge the communication gap between developer and user.

The Purpose of Software Requirement Specification document is to provide a detailed overview of the software product and its parameter and goals. It specifies all the requirement of the application.

The boundaries of software products are defined set of requirements. The software development team designing implements tests and delivers these requirements to you. A requirement is an atomic unit of software product from the view point of the user. As a rule, requirements are always correct, unambiguous, verifiable and traceable requirements are numbered and prioritized.

**2.2. Overall Description**

**2.2.1. Product perspective**

**2.2.1.1. System interfaces:**

This application runs in the latest version of python, ffmpeg and pytesseract on windows OS.

**2.2.1.2 User interfaces**

This applications GUI provides menus, toolbars, buttons, frames, containers, grids allowing for easy control by a keyboard and a mouse.

**2.2.1.3 Hardware interfaces**

Not Applicable.

**2.2.1.4 Software interfaces**

Not Applicable.

**2.2.1.5 Communications interfaces**

This application requires internet connectivity as communication interface.

**2.2.1.6 Interfaces with Server**

Not Applicable.

**2.2.2. Product Functions**

The product function relates physical output of a production process to physical input as features of production.

The general function of this application is to manipulate the digital files as per the user requirements. This application provides various functionalities like text manipulation on the image, spreadsheet manipulation, video-audio manipulations and file conversions like converting text, pdf and word document with each other.

**2.2.3. User characteristics**

The End User must have basic knowledge about the usage of the system, which is how to operate the application and the user must install the pytesseract and ffmpeg engine in their system.

**2.2.4. General constraints**

The software works on any windows computer system. And it requires clear image for manipulation and audio quality must be good and video quality must be good.

The system must have more than 2GB of RAM, and the screen resolution must be more than 1600X900 pixels. The system must have external softwares called pytesseract and ffmpeg and requires internet connectivity.

**2.2.5. Assumptions and Dependencies**

These factors are not design constraints on the software but any changes to these factors can affect the requirement in the SRS.

The image quality must be good for image and text manipulation. Video and Audio quality must be good. The system must have more than 2GB of RAM and must run on latest version of OS and the system must have pytesseract engine and ffmpeg software.

**2.3. Special Requirements (Software / Hardware - if any)**

This application requires a software called pytesseract which is an external engine and ffmpeg software.

**2.4. Functional requirements**

In the functional requirement section, the functional capabilities of the system are described. In this organization, the functional capabilities for all the modes of operation of the software are given. For each functional requirement, the required inputs, desired outputs and processing requirements will have to be specified.

**2.4.1 Text Manipulation in an Image**

**2.4.1.1 Extracting Text from Images**

**a. Input:** Input is a clear image.

**b. Process:** Process include extracting text from the given image.

**c. Output:** Output is the text extracted from the image.

**2.4.1.2 Extracting Text from Region of Interest**

**a. Input:** Input is a clear image.

**b. Process:** Process include extracting text from the region of interest from the given image.

**c. Output:** Output is the text extracted from the specified region in the image.

**2.4.1.3 Highlighting Text**

**a. Input:** Input is a clear image as input.

**b. Process:** Process include Highlighting text in the given image.

**c. Output:** Output is the text highlighted in the image or a message saying that “the given text is not found”.

**2.4.2 Spread Sheet Manipulation**

**2.4.2.1 Excel to Text**

**a. Input:** Inputs is an Excel file.

**b. Process:** Process include extracting text from the given Excel file to save it as a text file.

**c. Output:** Output is a text file.

**2.4.2.2 Csv to Text**

**a. Input:** Input is a Csv file.

**b. Process:** Process include extracting text from the given Csv file to save it as a text file.

**c. Output:** Output is a text file.

**2.4.2.3 Excel to Csv**

**a. Input:** Input is an Excel file and the sheet name.

**b. Process:** Process include extracting text from the given Excel file to save it as a Csv file.

**c. Output:** Output is a Csv file.

**2.4.2.4 Search a Column:**

**a. Input:** Input is an Excel File and the required column name.

**b. Process:** Process include finding the given column name in the excel file and writing that column data on a text file.

**c. Output:** Output is a Text file.

**2.4.3 Video-Audio Manipulation**

**2.4.3.1 Extracting Text from Video**

**a. Input:** Input is a video with clear audio.

**b. Process:** Process include extracting audio and then extracting text from that audio.

**c. Output:** Output is text file extracted from the video.

**2.4.3.2 Extracting Audio from Video**

**a. Input:** Input is a video with clear audio.

**b. Process:** Process include extracting audio from the video.

**c. Output:** Output is the audio file extracted from the video.

**2.4.3.3 Extracting Text from Audio**

**a. Input:** Input is a clear audio file.

**b. Process:** Process include extracting text from the audio.

**c. Output:** Output is the text file extracted from the audio.

**2.4.4 File Manipulation**

**2.4.4.1 Pdf to Text**

**a. Input:** Input is a pdf file.

**b. Process:** Process include extracting text from the given pdf file to save it as a text file.

**c. Output:** Output is a text file.

**2.4.4.2 Docx to Text**

**a. Input:** Input is a Docx file.

**b. Process:** Process include extracting text from the given Docx file to save it as a text file.

**c. Output:** Output is a text file.

**2.4.4.3 Pdf to Docx**

**a. Input:** Input is a pdf file.

**b. Process:** Process include extracting text from the given pdf file to save it as a Docx file.

**c. Output:** Output is a Docx file.

**2.4.4.4 Text to Docx**

**a. Input:** Input is a text file.

**b. Process:** Process include converting the text file to Docx file.

**c. Output:** Output is a Docxx file.

**2.5. Design Constraints**

The client environment may restrict the designer to include some design constraints that must be followed.

**2.5.1. Hardware Constraint**

The system must have more than 2GB of RAM, and the screen resolution must be more than 1600X900 pixels.

**2.5.2. Software Constraint**

The system must have Windows OS, ffmpeg and pytesseract .

**2.5.3. Fault Tolerance**

Fault tolerance requirements can place a major constraint on how the system is to be designed. Fault tolerance requirements often make the system more complex and expensive, so they should be minimized.

If any fault occurs due to blur image or blur audio or video, then the application shows an error message stating the fault and asks the user to input clear image, audio or video.

**2.5.4. Security**

Currently security requirements have become essential and major for all types of systems. Security requirements place restrictions on the use of certain commands, control access to database, provide different kinds of access, requirements for different people, require the use of passwords and cryptography techniques, and maintain a log of activities in the system.

**2.5.5. Standard Compliance**

It specifies the requirements for the standard the system must follow. The standards may include the report format, Type of Navigations, Naming Conventions for Button, access keys, shortcut keys.

**2.6. System Attributes**

* **Availability**

Availability refers to the percentage of time that the infrastructure, system, or solution remains operational under normal circumstances in order to serve its intended purpose.

* **Portability**

Portability, in relation to software, is a measure of how easily an application can be transferred from one computer environment to another. A computer software application is considered portable to a new environment if the effort required to adapt it to the new environment is within reasonable limits.

* **Reliability**

Reliability refers to the probability that the system will meet certain performance standards in yielding correct output for a desired time duration.

* **Maintainability**

Maintainability refers to the ease with which you can repair, improve and understand software code. Software maintenance is a phase in the software development cycle that starts after the customer has received the product.

* **Scalability**

Software scalability is an attribute of a tool or a system to increase its capacity and functionalities based on its users’ demand. Scalable software can remain stable while adapting to changes, upgrades, overhauls, and resource reduction.

**2.7. Other Requirements (if any)**

Not Applicable.

**3. System Design (Functional Design)**

**3.1 Introduction**

* System design is the process of defining the architecture,

Module interface and data for a system to satisfy specified requirements.

* The purpose of the design phase is to plan the solution of the problem

Specified by the requirements documents.

* This is the first step that moving from problem domain to the solution domain.
* The design of the system is essentially a blueprint or a plan for a solution for the system.

**3.2 Assumption and Constraints**

An assumption is a condition you think to be true and constraints is fixed limitations of project development.

* All the functional requirement collected from client are sufficient for the project life-cycle.
* All the Non functional and Specific requirement specified in SRS are well enough for the development of system.
* Time constraint.

**3.3 Functional decomposition**

Functional decomposition is the process of taking a complex process and breaking it down into its smaller, simpler parts. Using functional decomposition large or complex functionalities are more easily understood. It is mainly used during project analysis phase, so each phase can be viewed as software. So this has modular with some sub modules.

**3.3.1 System software architecture.**

**Digital File manipulation**

File Manipulation

Video-Audio Manipulation

Spread Sheet Manipulation

Text manipulation in an Image

CSV To TEXT

PDF to TEXT

Extracting audio from video

Extracting text from Images

XLS To TEXT

DOCX to TEXT

Extracting text from video

Text from Region of interest

Pdf to Docx

XLS To CSV

TEXT to DOCX

Extracting text from audio

Search Column

Highlighting text

**3.3.2 System Technical Architecture**

USER

APPLICATION

DISK

**3.3.3 System Hardware Architecture**

Hard copy Device

Monitor

Computer

Standard input Device

Mass Storage

**3.3.4 External Interfaces**

Not applicable

**3.4 Description of the programs**

**3.4.1 Context flow Diagram**

In CFD entire system is considered as a single process. Context flow Diagram shows input and outputs of the system. It shows all the external entities that interact with the system and how the data flow between this external entities and system.

User

File

Manipulated Digital File

**3.4.2 Data Flow Diagram (DFD’s Level 1,Level 2,Level 3)**

**Data Flow Diagram(DFD)**

Data flow diagram shows the flow of the data through system. Data flow diagram also called the data flow graphs. It views a system as a function that transforms the inputs into desired outputs. It aims to capture the transformation that takes place within a system to the input data so that eventually the output data is produced.

|  |  |  |
| --- | --- | --- |
| **SYMBOLS** | **NAMES** | **DESCRIPTION** |
|  | Process | It performs transformation of the data from one state to another. |
|  | Source/Sink | It represents the external entity that may either source or sink. |
|  | Flow of data | It represents the flow of data from source to destination. |
|  | Data Source/Data | It is placed where data is stored. |

**Level 0 DFD**

File

Image File

Manipulated text

Xls or Csv File

Audio, Video File

Converted File

Manipulated File

Manipulated Audio and Video

**USER**

**Level 1 DFD**

Text Manipulation

Spread Sheet Manipulation

Audio and Video Manipulation

File Manipulation

**3.5** **Description of Components**

**3.5.1 Functional Component-1 Text Manipulation**

**3.5.1.1 Extracting Text From Images**

USER

SELECTED FILE

disk

3.5.1.1.1 Input

User Provide the Clear Images

3.5.1.1.2 Process Definition

Application retrieves the text in the image

3.5.1.1.3 Output

Application saves the text in Text file format.

3.5.1.1.4 Interface with other functional component

Independent module

3.5.1.1.5 Resource Allocation

System Internal Storage

3.5.1.1.6 User interface

Buttons ,Frames ,Entry box, Message box, Label.

**3.5.1.2 Extracting Text from Region of Interest**

USER

SELECTED FILE

disk

3.5.1.2.1 Input

User Provide the clear Image

3.5.1.2.2 Process Definition

Here user specifies the region in that image.

3.5.1.2.3 Output

Application Will retrieve the text in that Image

3.5.1.2.4 Interface with other functional component

Independent module

3.5.1.2.5 Resource Allocation

System Internal Storage

3.5.1.2.6 User interface

Button, Frames, Entry box, Message box, Label.

**3.5.1.3 Highlighting Text**

USER

SELECTED FILE

disk

3.5.1.3.1 Input

User Provide the clear Image and word or sentence.

3.5.1.3.2 Process Definition

Here user highlight the specified text in an Image.

3.5.1.3.3 Output

Application will return the image with highlighted text.

3.5.1.3.4 Interface with other functional component

Independent module

3.5.1.3.5 Resource Allocation

System Internal Storage.

3.5.1.3.6 User Interface

Button, Frames, Entry box, Message box, Label.

3.5.**2 Functional Component 2 Spread Sheet Manipulation**

**3.5.2.1 Csv To Text**

USER

SELECTED FILE

disk

3.5.2.1.1 Input

User Provides a Csv File using independent module.

3.5.2.1.2 Process Definition

Application process the Csv file and converts to Text file.

3.5.2.1.3 Output

Application will returns the text file and create the text file in original location.

3.5.2.1.4 Interface with other functional component

Independent module

3.5.2.1.5 Resource Allocation

System Internal Storage

3.5.2.1.6 User interface

Buttons , Frames ,Entry box, Message box, Label.

**3.5.2.2 Xls To Text**

USER

SELECTED FILE

disk

3.5.2.2.1 Input

User Provide the Xls file using independent module.

3.5.2.2.2 Process Definition

Application process the Xls file and convert to Text format.

3.5.2.2.3 Output

Application will returns the text file and create the text file

In the original location.

3.5.2.2.4 Interface with other functional component

Independent module

3.5.2.2.5 Resource Allocation

System Internal Storage

3.5.2.2.6 User interface

Buttons , Frames ,Entry box, Message box, Label.

**3.5.2.3 Xls To Csv**

USER

SELECTED FILE

disk

3.5.2.3.1 Input

User Provide the Xls file using independent module.

3.5.2.3.2 Process Definition

Application converts the Xls file to Csv file.

3.5.2.3.3 Output

Application return the Csv file and create the Csv file in original location.

3.5.2.3.4 Interface with other functional component

Independent module

3.5.2.3.5 Resource Allocation

System Internal Storage

3.5.2.3.6 User interface

Buttons , Frames ,Entry box, Message box, Label.

**3.5.2.4 Search column**

USER

SELECTED FILE

disk

3.5.2.3.1 Input

User Provide the Xls file using independent module.

3.5.2.3.2 Process Definition

Application process the xls file and search the columns..

3.5.2.3.3 Output

Application return the searched columns.

3.5.2.3.4 Interface with other functional component

Independent module

3.5.2.3.5 Resource Allocation

System Internal Storage

3.5.2.3.6 User interface

Buttons , Frames ,Entry box, Message box, Label.

**3.5.3 Functional Component3 Video and Audio Manipulation**

**3.5.3.1 Extracting the Text from Video**

USER

SELECTED FILE

disk

3.5.3.1.1 Input

User Provide the Video as input.

3.5.3.1.2 Process Definition

Application retrieves the words spoken in that video.

3.5.3.1.3 Output

Application will writes the text on a text file.

3.5.3.1.4 Interface with other functional component

Independent module

3.5.3.1.5 Resource Allocation

System Internal Storage

3.5.3.1.6 User interface

Buttons , Frames ,Entry box, Message box, Label.

**3.5.3.2 Extracting the Audio from Video**

USER

SELECTED FILE

disk

3.5.3.2.1 Input

User Provide the Video as input in this application.

3.5.3.2.2 Process Definition

Application will retrieves audio in the video.

3.5.3.2.3 Output

Application will produce the audio as output.

3.5.3.2.4 Interface with other functional component

Independent module

3.5.3.2.5 Resource Allocation

System Internal Storage

3.5.3.2.6 User interface

Buttons , Frames ,Entry box, Message box, Label.

**3.5.3.3 Extracting the Text from Audio**

USER

SELECTED FILE

disk

3.5.3.3.1 Input

User Provide the audio as input to this application.

3.5.3.3.2 Process Definition

Application retrieves the words spoken in the audio.

3.5.3.3.3 Output

Application will writes the text in text file.

3.5.3.3.4 Interface with other functional component

Independent module

3.5.3.3.5 Resource Allocation

System Internal Storage

3.5.3.3.6 User interface

Buttons , Frames ,Entry box, Message box, Label.

**3.5.4 Functional Componnet4 File Manipulation**

**3.5.4.1 PDF to TEXT**

USER

SELECTED FILE

disk

3.5.4.1.1 Input

User Provide the PDF file as input.

3.5.4.1.2 Process Definition

Application will converts into text file.

3.5.4.1.3 Output

Application will produce the text file as output and create a text file in particular folder.

3.5.4.1.4 Interface with other functional component

Independent module

3.5.4.1.5 Resource Allocation

System Internal Storage

3.5.4.1.6 User interface

Buttons , Frames ,Entry box, Message box, Label.

**3.5.4.2 DCOX to TEXT**

Input

USER

SELECTED FILE

disk

3.5.4.2.1 Input

User Provide the word file as input to this application.

3.5.4.2.2 Process Definition

Application process the word file and converts to text file format.

3.5.4.2.3 Output

Application will creates the text file in particular folder.

3.5.4.2.4 Interface with other functional component

Independent module

3.5.4.2.5 Resource Allocation

System Internal Storage

3.5.4.2.6 User interface

Buttons , Frames ,Entry box, Message box, Label.

**3.5.4.3 Pdf to Docx**

USER

SELECTED FILE

disk

3.5.4.3.1 Input

User Provide the pdf file as input.

3.5.4.3.2 Process Definition

Application converts into docx format.

3.5.4.3.3 Output

Application will create the docx file in particular folder.

3.5.4.3.4 Interface with other functional component

Independent module

3.5.4.3.5 Resource Allocation

System Internal Storage

3.5.4.3.6 User interface

Buttons , Frames ,Entry box, Message box, Label.

**3.5.4.3 TEXT to DOCX**

USER

SELECTED FILE

disk

3.5.4.3.1 Input

User Provide the text file as input.

3.5.4.3.2 Process Definition

Application converts into Dcox format.

3.5.4.3.3 Output

Application will create the Dcox file in particular folder.

3.5.4.3.4 Interface with other functional component

Independent module

3.5.4.3.5 Resource Allocation

System Internal Storage

3.5.4.3.6 User interface

Buttons , Frames ,Entry box, Message box, Label.

**4 .DETAILED DESIGN**

**4.1 Introduction :**

During detailed design, the internal logic of each modules specified in system design is decided. During this phase further details of the modules are decided. Design of each of the modules usually specified in a high level description language which is independent of the language in which software eventually be implemented.

**4.2 Structure of software package :**

**Digital File manipulation**

CSV To TEXT

Extracting text from video

XLS To TEXT

Text manipulation in an Image

Spread Sheet Manipulation

Video-Audio Manipulation

File Manipulation

Extracting text from Images

Extracting text from Region of interest

Highlighting text

XLS To CSV

Extracting audio from video

Extracting text from audio

PDF to TEXT

DOCX to TEXT

Pdf to Docx

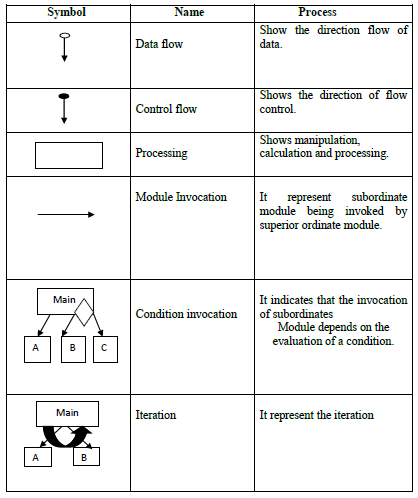
Search Column

TEXT to DOCX

**4.3 Module decomposition of software :**

**Structure chart:**

Structure chart is a top-down modular design, consist of squares representing different models in a system and lines .Structure chart shows how program has been partitioned into manageable modules hierarchy and organization of those modules and communicational interface.



**Flow chart :**

Flow chart is a graphical representation of solution to the given problems. A Flowchart is pictorial representation of an algorithm, workflow or process. The diagrammatic representation illustrates a solution model to given problem. It uses the following symbol.

|  |  |  |
| --- | --- | --- |
| **Symbol** | **Name** | **Purpose** |
|  | Terminator | It indicates the start and end process |
|  | Input/Output | Input/Output data. |
|  | Decision | It represents a comparison or question that determines an alternate path to be followed. |
|  | Flow direction | Shows the direction of data flow. |
|  | Processing | It represents manipulation, calculation, or information processing. |
|  | Direction access storage | File storage. |
|  | Preparation(Looping) | An instruction or group of instruction. |
|  | In-page |  |
|  | Off-page |  |
|  | Delay |  |

**4.3.1 Text manipulation:**

**4.3.1.1 Extracting Text from image:**

**4.3.1.1.1 Input:**

Single Image with jpg, jpeg, png, webp format from disk.

**4.3.1.1.2 Procedural details:**

**Flow Chart:**

Read Image File

start

Save the text file

Display the path of text file

stop

Extracting Text from image

disk

**4.3.1.1.3 File I/O interfaces**

Not applicable

**4.3.1.1.4 Outputs**

**Extracted text file.**

**4.3.1.1.5 Implementation aspects**

Not applicable.

**4.3.1.2 Extracting Region of interest:**

**4.3.1.2.1 Input:**

Single Image with jpg, jpeg, png, webp format from disk.

**4.3.1.2.2 Procedural details:**

**Flow Chart**

Read Image File

start

Save the text file

Display the path of text file

stop

Extracting region of interest

disk

**4.3.1.2.3 File I/O interfaces**

Not applicable

**4.3.1.2.4 Outputs**

**Extracted text file.**

**4.3.1.2.5 Implementation aspects**

Not applicable.

**4.3.1.3 Highlighting text from Image:**

**4.3.1.3.1 Input:**

Single Image with jpg, jpeg, png, webp format from disk.

**4.3.1.3.2 Procedural details:**

**Flow Chart:**

Read Image File

start

Save the text file

Display the path of text file

stop

Highlighting text from image

disk

**4.3.1.3.3 File I/O interfaces**

Not applicable

**4.3.1.3.4 Outputs**

Extracted text file.

**4.3.1.3.5 Implementation aspects**

Not applicable.

**4.3.2. Spreadsheet manipulation:**

**4.3.2.1 CSV to TEXT:**

**4.3.2.1.1 Input:**

Single .csv format file from disk.

**4.3.2.1.2 Procedural details:**

**Flow chart:**

Read CSV File

start

Save the text file

Display the path of text file

stop

CSV to TEXT

disk

**4.3.2.1.3 File I/O interfaces**

Not applicable

**4.3.2.1.4 Outputs**

Extracted text file from csv file.

**4.3.2.1.5 Implementation aspects**

Not applicable.

**4.3.2.2 XLS to TEXT:**

**4.3.2.2.1 Input:**

Single .xlsx format file from disk.

**4.3.2.2.2 Procedural details:**

**Flow chart:**

Read XLSX File

start

Save the text file

Display the path of text file

stop

XLSX to TEXT

disk

**4.3.2.2.3 File I/O interfaces**

Not applicable

**4.3.2.2.4 Outputs**

Extracted text file from xlsx file.

**4.3.2.2.5 Implementation aspects**

Not applicable

**4.3.2.3 XLS to CSV:**

**4.3.2.3.1 Input:**

Single .xlsx format file from disk.

**4.3.2.3.2 Procedural details:**

**Flow chart:**

Read XLSX File

start

Save the csv file

Display the path of csv file

stop

XLSX to CSV

disk

**4.3.2.3.3 File I/O interfaces**

Not applicable

**4.3.2.3.4 Outputs**

Extracted csv file from xlsx file.

**4.3.2.3.5 Implementation aspects**

Not applicable

**4.3.2.4 Search Column:**

**4.3.2.4.1 Input:**

Single .xlsx format file from disk.

**4.3.2.4.2 Procedural details:**

**Flow chart:**

Read XLSX File

start

Save the xlsx file with search columns.

Display the path of xlsx file

stop

SEARCH COLUMN

disk

**4.3.2.4.3 File I/O interfaces**

Not applicable

**4.3.2.4.4 Outputs**

Saved xlsx file with searched columns.

**4.3.2.4.5 Implementation aspects**

Not applicable

**4.3.3. Audio and Video manipulation:**

**4.3.3.1 Extracting audio from Video :**

**4.3.2.1.1 Input**:

Single video file from disk.

**4.3.2.1.2 Procedural details:**

**Algorithm:**

Step 1: start

Step 2: input video file

Step 3: extracting the audio from video

Step 4: save the audio file to disk

Step 5: display the path of stored audio file

Step 6: stop

**4.3.2.1.3 File I/O interfaces**

Not applicable

**4.3.2.1.4 Outputs**

Extracted audio file from video..

**4.3.2.1.5 Implementation aspects**

Not applicable

**4.3.3.2 Extracting text from Video :**

**4.3.2.2.1 Input**:

Single video file from disk.

**4.3.2.2.2 Procedural details:**

**Algorithm:**

Step 1: start

Step 2: input video file

Step 3: extracting the text from video

Step 4: save the text file to disk

Step 5: display the path of stored text file

Step 6: stop

**4.3.2.2.3 File I/O interfaces**

Not applicable

**4.3.2.2.4 Outputs**

Extracted audio file from video..

**4.3.2.2.5 Implementation aspects**

Not applicable

**4.3.3.3 Extracting text from audio :**

**4.3.3.1.1 Input**:

Single audio file from disk.

**4.3.3.1.2 Procedural details:**

**Algorithm:**

Step 1: start

Step 2: input audio file

Step 3: extracting the text from audio

¿ Step 4: save the text file to disk

Step 5: display the path of stored text file

Step 6: stop

**4.3.3.1.3 File I/O interfaces**

Not applicable

**4.3.3.1.4 Outputs**

Extracted audio file from video..

**4.3.3.1.5 Implementation aspects**

Not applicable

**4.3.4 File Manipulation**

**4.3.4.1 PDF TO TEXT :**

**4.3.4.1.1 Input**:

User provide the single pdf file from disk.

**4.3.4.1.2 Procedural details:**

**Structure Chart:**

USER

Read the pdf file

PDF TO TEXT

SAVE THE TEXT FILE

DISK

DISPLAY THE PATH OF FILE

**4.3.4.1.3 File I/O interfaces**

Not applicable

**4.3.4.1.4 Outputs**

Converted text file from pdf file

**4.3.4.1.5 Implementation aspects**

Not applicable

**4.3.4.2 DOCX TO TEXT :**

**4.3.4.2.1 Input**:

User provide the single DOCX file..

**4.3.4.2.2 Procedural details:**

**Structure Chart:**

USER

Read the pdf file

DOCX TO TEXT

SAVE THE TEXT FILE

DISK

DISPLAY THE PATH OF FILE

**4.3.4.2.3 File I/O interfaces**

Not applicable

**4.3.4.2.4 Outputs**

Converted text file from docx file

**4.3.4.2.5 Implementation aspects**

Not applicable

**4.3.4.3 Pdf to Docx:**

**4.3.4.3 .1 Input**:

User provide the single Pdf file..

**4.3.4.3.2 Procedural details:**

**Structure Chart:**

USER

Read the pdf file

Pdf to Docx

SAVE THE TEXT FILE

DISK

DISPLAY THE PATH OF FILE

**4.3.4.3.3 File I/O interfaces**

Not applicable

**4.3.4.3.4 Outputs**

Converted Docx file from Pdf file.

**4.3.4.3.5 Implementation aspects**

Not applicable

**4.3.4.4 TEXT TO DOCX :**

**4.3.4.4 .1 Input**:

User provide the single Text file..

**4.3.4.4.2 Procedural details:**

USER

Read the pdf file

TEXT TO DOCX

SAVE THE TEXT FILE

DISK

DISPLAY THE PATH OF FILE

**4.3.4.4.3 File I/O interfaces**

Not applicable

**4.3.4.4.4 Outputs**

Converted docx file from text file.

**4.3.4.4.5 Implementation aspects**

Not applicable