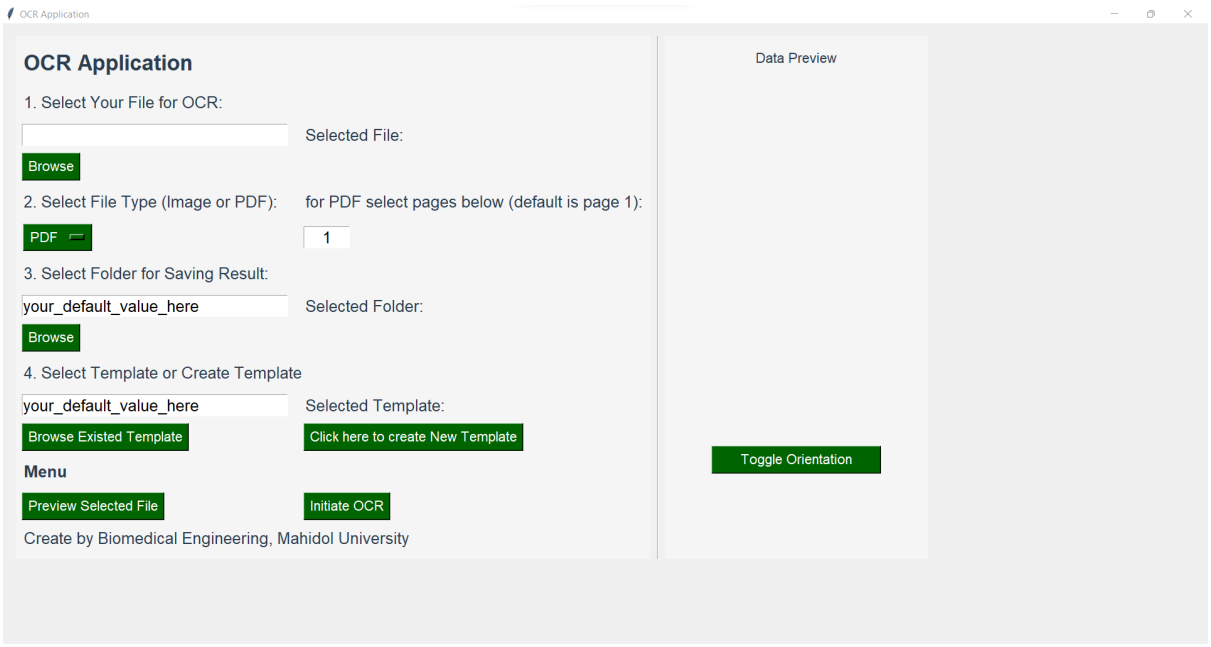
OCR Application User Manual

Introduction

OCR (Optical Character Recognition) application is designed to facilitate optical character recognition on images and PDF files. OCR is a technology that converts different types of documents, such as scanned paper documents, PDFs, or images captured by a digital camera, into editable and searchable data.



The OCR application provides a comprehensive interface to streamline the document processing workflow. Users can seamlessly navigate through the following key steps:

1. Browse Files and Folders:

Initiate the process by selecting the target file or folder containing the documents to be processed. This step sets the foundation for subsequent OCR operations.

2. File Type Selection:

Choose the desired file type for the OCR output. This flexible option accommodates various document formats, enhancing compatibility with diverse user requirements.

3. Select or Create Templates:

Tailor the OCR process to specific document structures by selecting existing templates or creating new ones. Templates provide a standardized framework for accurate data extraction.

4. Folder Selection:

Specify the output folder where the processed documents and OCR results will be saved. Organize the workflow efficiently by directing the application to the desired destination.

5. Preview File:

Visualize the document before initiating OCR to ensure accuracy and make any necessary adjustments. This step allows users to verify the selected file's content and layout.

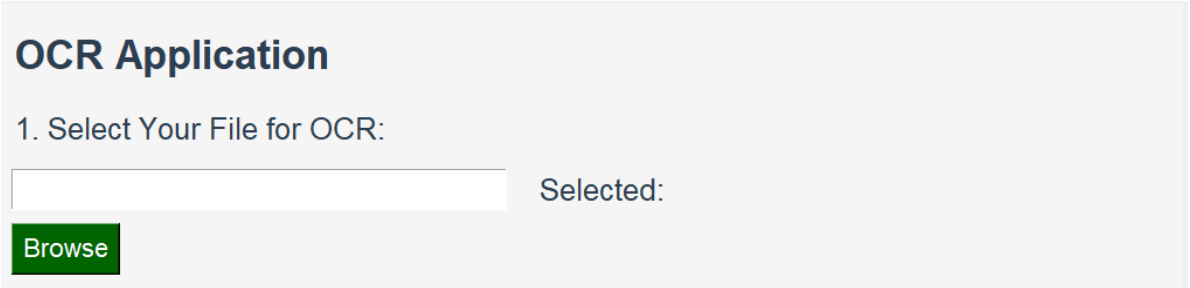
6. Initiate OCR:

Execute the Optical Character Recognition process to extract text and data from the documents. The application employs advanced algorithms to enhance accuracy and efficiency, delivering a seamless experience for users seeking to convert images or scanned documents into editable text.

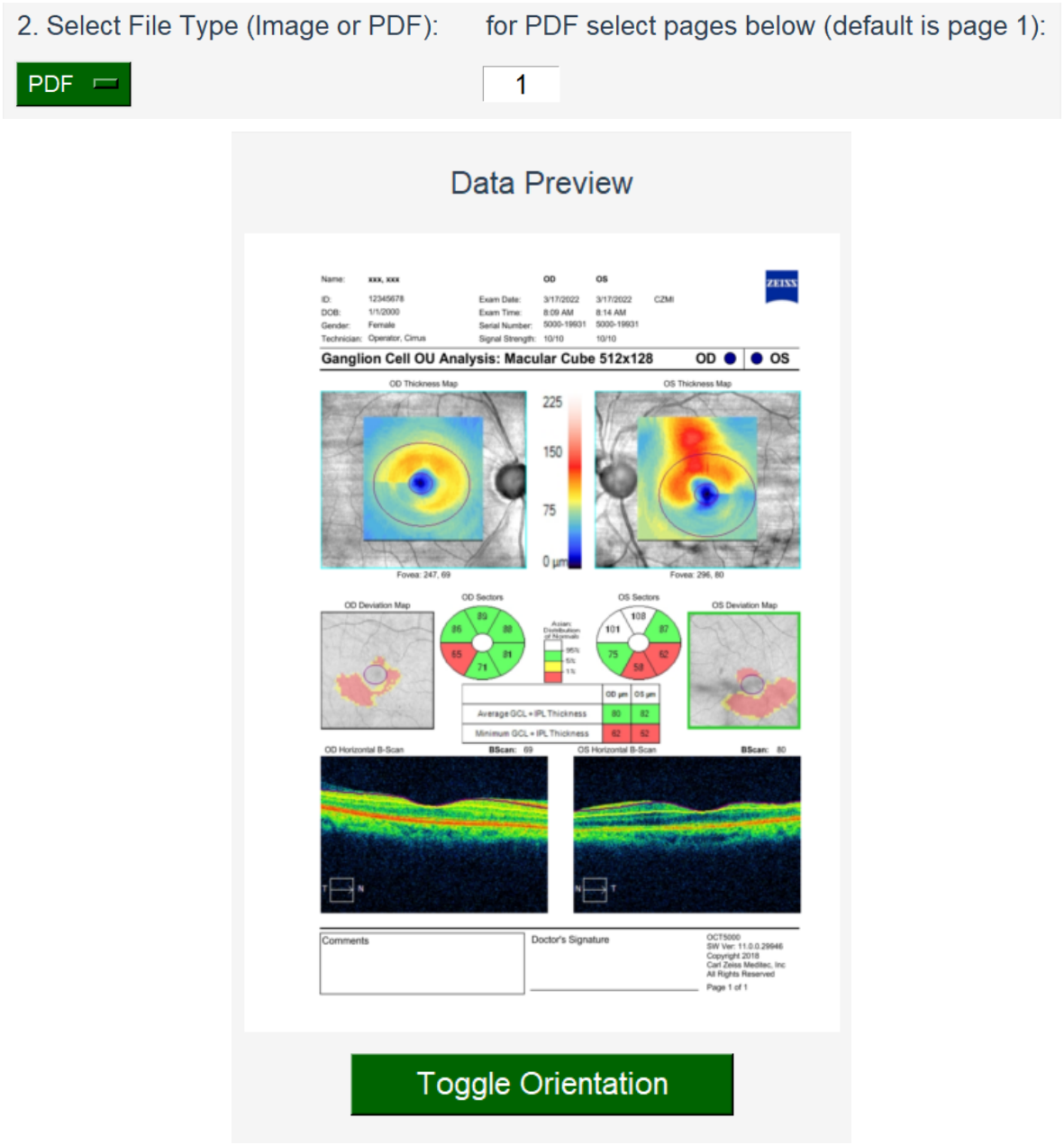
This interface follows a logical top-to-bottom sequence, guiding users through each step of the OCR workflow. The user-friendly design ensures a smooth experience by presenting options in a clear and intuitive manner. As users progress through the workflow, they are encouraged to click the prominent green button at each stage, moving them forward in the sequence. This straightforward approach minimizes complexity and promotes a step-by-step understanding of the process. By clicking the green button, users advance confidently, fostering a sense of control and facilitating efficient navigation through the OCR application's sequential interface.

Getting Started

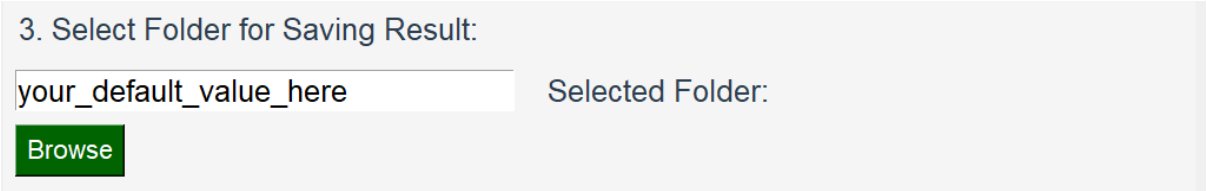
Prerequisite:
Setup the software program by following the README.txt in cloned folder



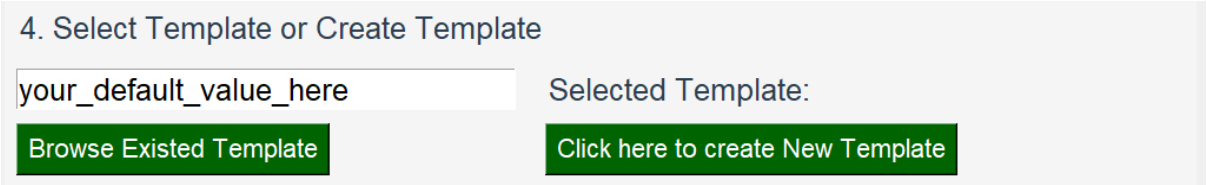
1. Select File for OCR
    a. Click on the "Browse" button next to the "Select Your File for OCR" label.
    b. Choose the file that you want to perform OCR on.
    c. The selected file path will be displayed in the "Selected File" section.



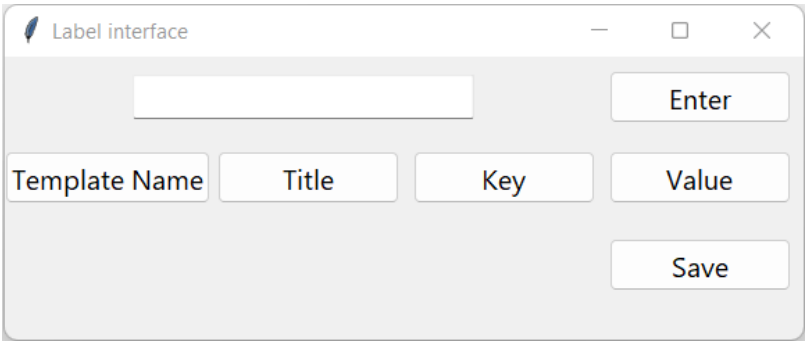
2. Select File Type
    a. Choose the file type (Image or PDF) using the dropdown menu.
    b. If "PDF" is selected, the page number can be specified for processing (default is page 1).
    c. The selected file can be previewed on the right hand side “Data Preview” section
    d. Click on the "Toggle Orientation" button to switch between landscape and portrait orientations. The displayed file in the "Data Preview" section will be resized accordingly.

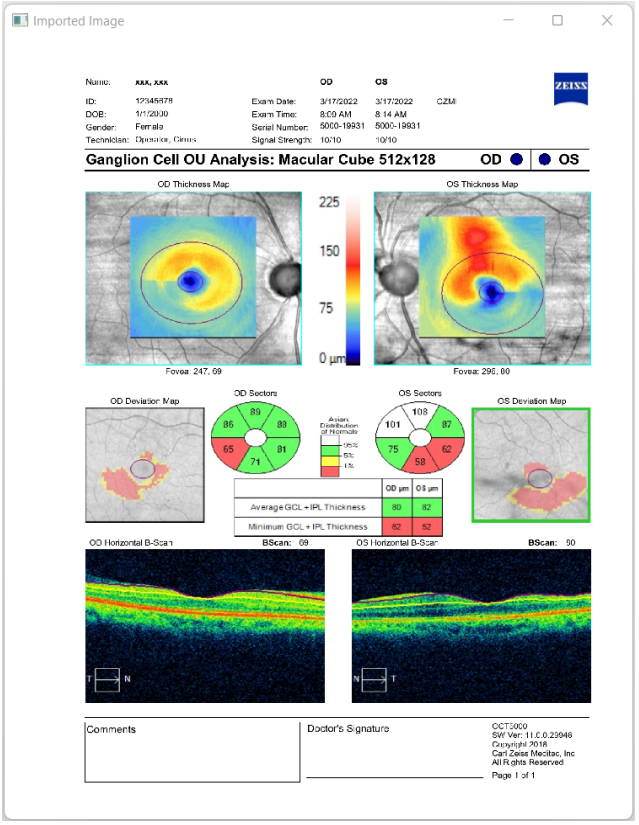


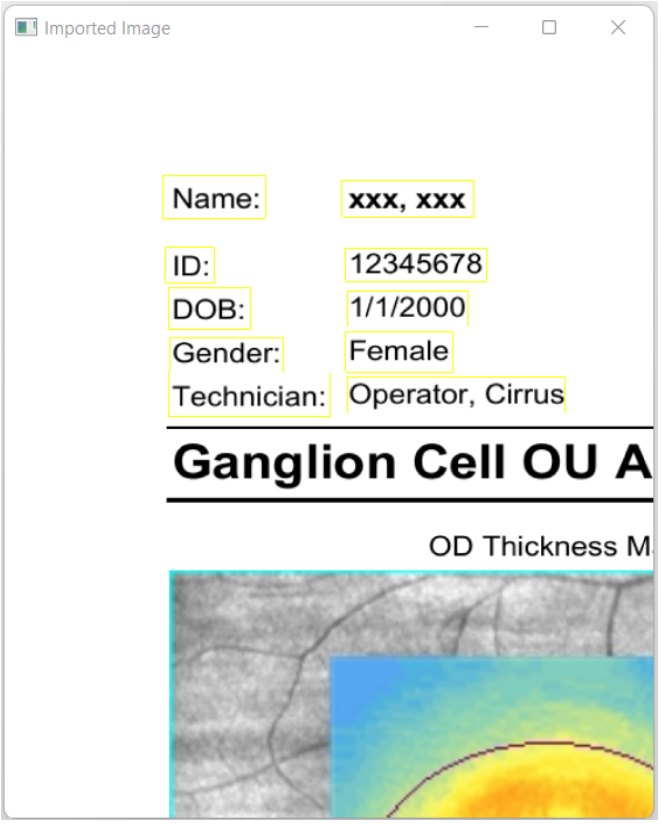
3. Select Folder for Saving Result
    a. Click on the "Browse" button next to the "Select Folder for Saving Result" label.
    b. Choose the folder where to save the OCR results.
    c. The selected folder path will be displayed in the "Selected Folder" section.



4. Select Template or Create Template
    a. Choose an existing JSON template by clicking on the "Browse Existed Template" button.
    b. Alternatively, create a new template by clicking on "Click here to create New Template."







Create a new template

TKinter or Label interface window consist of input typing section for some mode, the status display, and provided buttons to select the mode:

- Template Name: Enter the template name from typing.
- Title: Enter the title from typing.
- Key: Enter a key for typing or selected region for OCR as stored as a key.
- Value: Capture the selected region and store it in template for subsequent OCR.

The CV2 or imported image window is used for

- Use for key value and value input
- Cropping ROI of value and storing in template file for OCR in the subsequent step
- The cropping can also be perform on key value which giving the text from OCR directly instead of ROI

Step-by-Step Creating New Template

Step 1: Enter Template and Title Name

- The template, title, and key value can be entered by the typing input.
- Type the template name and press Enter button.
- Then, type the title name and press Enter button.
- The title name can be added for each section of document such as patient information, table, and circle chart.
- The key value can be add either by typing and OCR cropping

Step 2: Enter Region Information

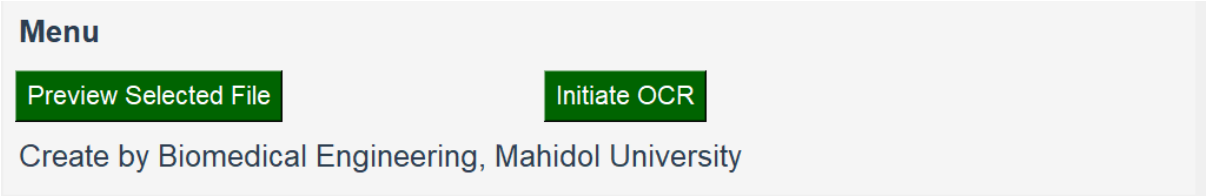
- The script will display a graphical or cv2 interface for image cropping.
- Use the mouse to select regions of interest (ROI) by clicking and dragging.
- Press Enter on keyboard to capture the ROI and perform OCR on the selected region.
- For the key value mode, the image in ROI performs OCR and is stored in text.
- For the value mode, the ROI is stored in the template file.

Step 3: Save the Template

- When all information has been entered for all regions, press the "Save" button to save the template data to a JSON file.
- The script will display a success message and the path to the saved JSON file.

Step 4: Exiting the Program After Finishing Template

- To exit the program, press ESC while the image cropping interface is active.
- Alternatively, close the main Tkinter window.



Menu for OCR process

1. Preview Selected File
    a. Click on the "Preview Selected File" button to display the selected file in the "Data Preview" section.
    b. The file will be displayed with the current orientation.
2. Initiate OCR
    a. After selecting the file and providing other necessary information, click on the "Initiate OCR" button.
    b. The OCR process will start, and the results will be saved in the specified folder.

Troubleshooting

- If you encounter any issues during the OCR process, an error message will be displayed in the "Error" section.
- Ensure that all required fields are filled before initiating OCR.

Contact

For assistance or inquiries, please contact the Biomedical Engineering Department at Mahidol University.

Acknowledgments

This OCR Application is created and maintained by the Biomedical Engineering Department at Mahidol University. Special thanks to all contributors for their efforts in developing and improving this tool.