

USER MANUAL FOR OCR IMAGE RECOGNITION TOOL

Jana Saleh Godieh	0223457
Adnan Ibrahim Sawalha	0221536
Basel Salah Qarout	0225322

Project Supervisor: Dr. Omar Sultan Al-Kadi

May 25, 2025



INTRODUCTION

Welcome to the *OCR Image Recognition Tool*—a versatile application designed to extract text from a digital image. Whether your image is a scanned document, a photograph, or a screenshot, this tool leverages Optical Character Recognition (OCR) technology to transform visual text into editable, machine-readable text.

What This Tool Offers:

- *Universal Text Extraction:* Works with various image formats including JPG, PNG, BMP, and TIFF.
- *Image Preprocessing:* Multiple image enhancement options such as grayscale conversion, noise removal, sharpening, and skew correction to prepare images for better OCR performance.
- *Region of Interest Selection:* Users can target specific areas in an image to improve processing efficiency and focus OCR on relevant text sections.
- *Real-time Results:* Extracted text appears instantly in the designated output area for easy copying and further use.

Scope and Limitations

While the tool performs reliably on clear, printed text images, OCR accuracy may vary with image quality. Handwritten texts and highly complex layouts may not be fully supported in this version. However, the modular design allows for future improvements and feature expansions.

GETTING STARTED

This section will guide you through launching the application and provide a high-level overview of the user interface to help you get comfortable with the main components.

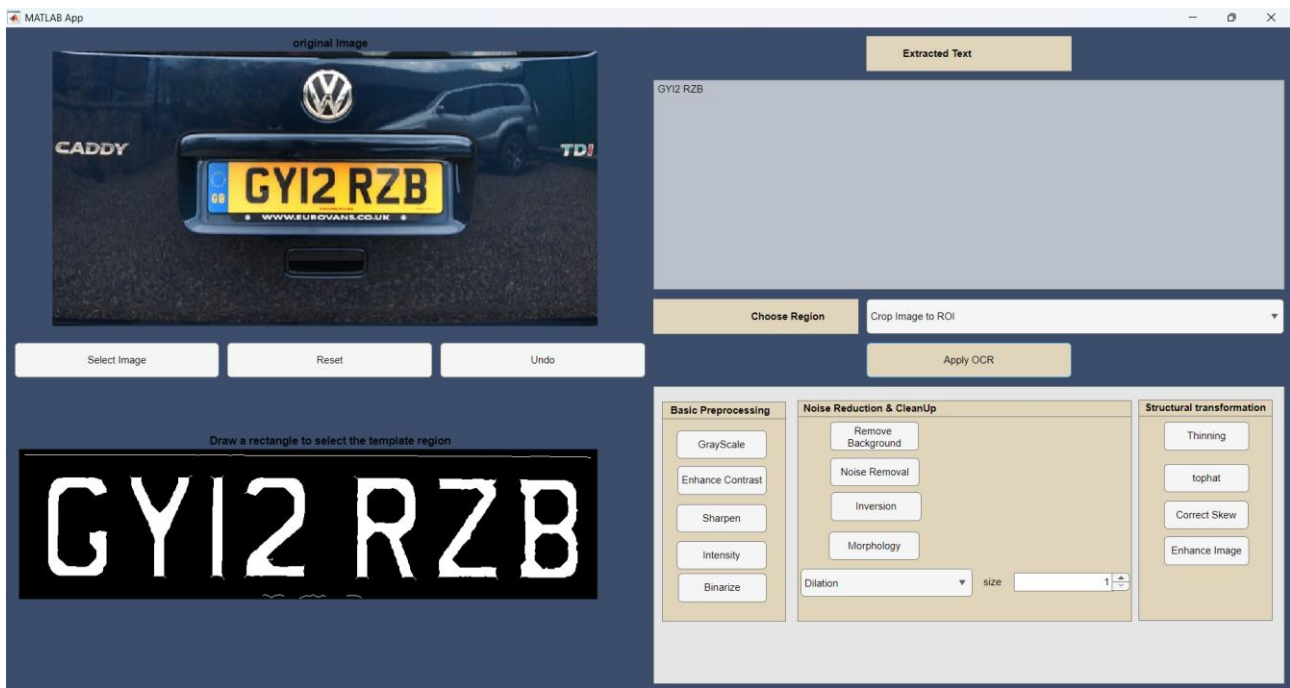


Figure 1: Screenshot of the OCR Image Recognition Tool's user interface

Launching the Application

To begin, simply run the application file in MATLAB. Once opened, the main window will appear, divided into distinct sections designed to streamline your workflow:

1. Image Displays

- *Original Image Display*: Shows the uploaded image before any processing.
- *Modified Image Display*: Shows the image after applying any preprocessing or transformations.

2. Preprocessing Panels

The right side contains three main panels to enhance the image before OCR:

- *Basic Preprocessing*: Common image adjustments like grayscale, contrast, sharpening, etc.
- *Noise Reduction & Cleanup*: Tools to remove noise, background, and perform morphological operations.

- *Structural Transformation*: Advanced operations including thinning, skew correction, and image enhancement.

3. Controls and Workflow

- *Select Image Button*: Browse and load images.
- *Choose Region Dropdown*: Define how regions of interest (ROI) are handled—process whole image, crop ROI, or preprocess only ROI.
- *Undo and Reset Buttons*: Undo last changes or reset to the original image.
- *Apply OCR Button*: Initiate text extraction on the processed image.
- *Extracted Text Area*: Displays the recognized text for review and copying.

FEATURES AND FUNCTIONALITY

This section explains the components of the user interface and their functions, enabling you to make the most out of the OCR Image Recognition Tool.



1. Original Image Display

Located on the left side, this panel displays the image you load into the application. It shows the raw, unprocessed image so you can visually inspect the input before applying any enhancements or OCR.

2. Select Image Button

Click this button to browse your computer and select an image file to load. Supported formats include JPG, PNG, BMP, and TIFF.

3. Reset Button

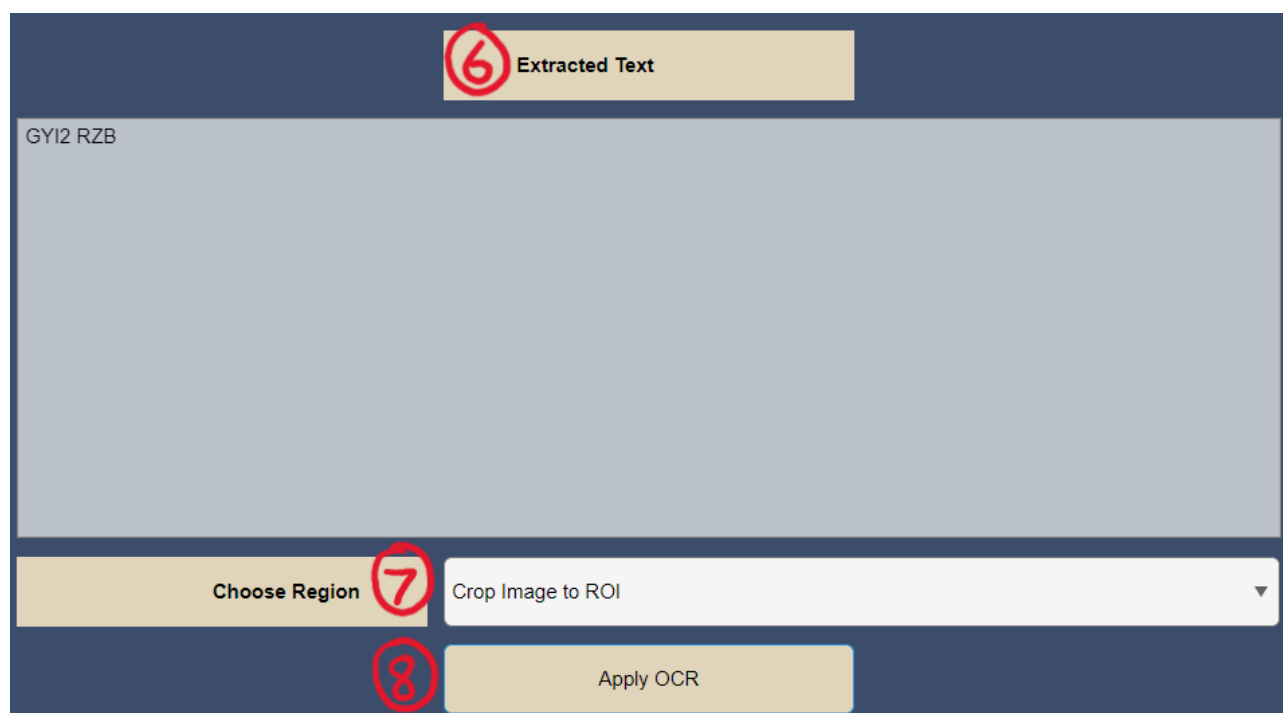
Returns the modified image back to the original, clearing all applied changes.

4. Undo Button

Reverts the last applied change, allowing you to step back through your edits.

5. Modified Image Display

This panel shows the current state of the image after any preprocessing or transformations you apply. It helps you track how your edits affect the image and prepares it for OCR.



6. Extracted Text Area

This read-only text box displays the OCR results. You can copy the text from here for use in other documents or applications.

7. Choose Region Dropdown

This dropdown lets you select how to handle regions of interest (ROI) in your image:

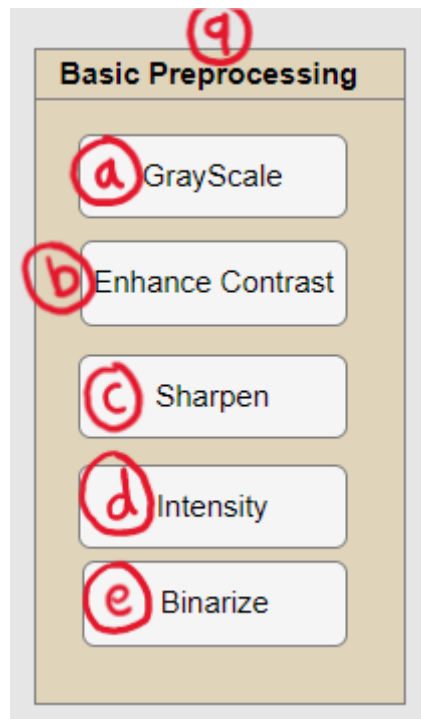
- *None*: Process the entire image.
- *Crop Image to ROI*: Select a rectangular region to crop and process only that part.
- *Preprocess ROI Only*: Apply image preprocessing only within the selected region before running OCR.

When you choose either ROI option, you'll be prompted to draw a rectangle over the Modified Image panel to specify the area.



8. Apply OCR Button

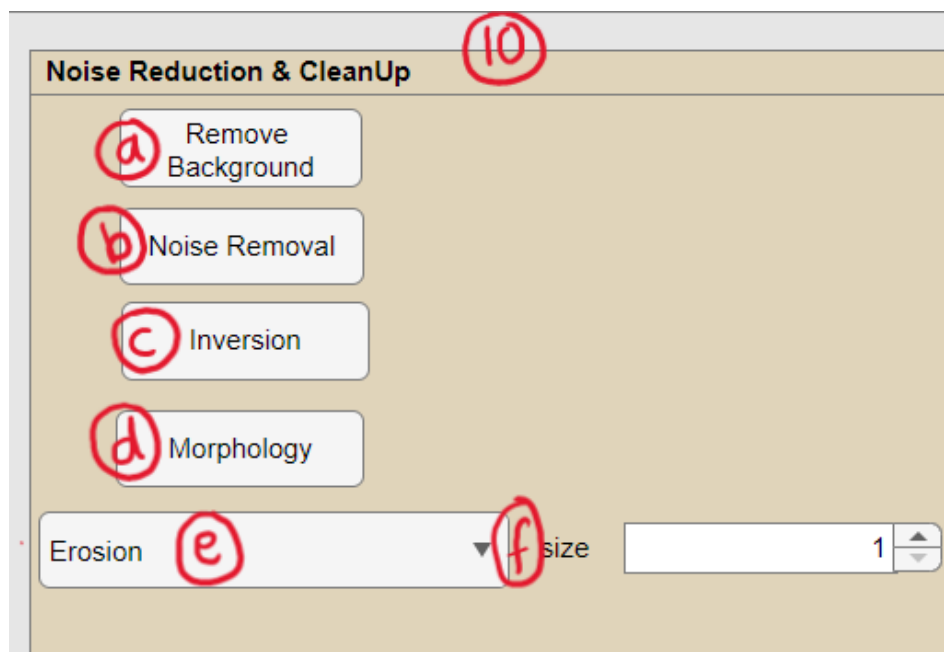
Once preprocessing is complete, click this button to extract text from the processed image. The recognized text will appear in the Extracted Text Area.



9. Basic Preprocessing Panel

These buttons provide essential image enhancements:

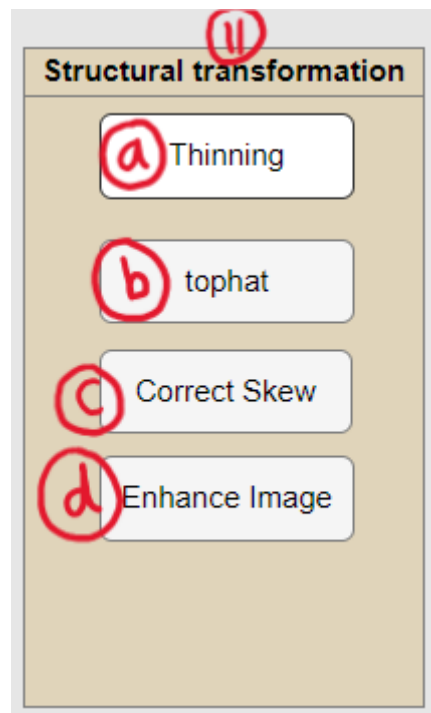
- a) *Grayscale*: Converts the image to grayscale, reducing complexity for OCR.
- b) *Enhance Contrast*: Improves the contrast to make text more distinguishable.
- c) *Sharpen*: Sharpens edges to highlight text features.
- d) *Intensity*: Adjusts the image intensity for clearer visuals.
- e) *Binarize*: Converts the image into black and white for easier text recognition.



10. Noise Reduction & Cleanup Panel

These tools help clean up your image before OCR:

- a) *Remove Background*: Eliminates unwanted background details using morphological operations.
- b) *Noise Removal*: Applies a median filter to reduce background noise.
- c) *Inversion*: Inverts the image colors, which can be useful for certain image types.
- d) *Morphology*: Performs morphological operations (Erosion, Dilation, Opening, Closing).
- e) *Size*: configures the structuring element size of the morphological operations.



11. Structural Transformation Panel

These options provide advanced image processing techniques:

- a) *Thinning*: Reduces the thickness of characters to simplify shapes.
- b) *Tophat*: Applies a tophat filter to enhance bright regions on a dark background.
- c) *Correct Skew*: Detects and corrects tilted or rotated text to align it properly for OCR.
- d) *Enhance Image*: Enlarges the image to improve resolution for OCR.

CONCLUSION

The OCR Image Recognition Tool offers a practical and efficient solution for extracting text from a wide range of image formats. Its combination of intuitive controls and powerful image preprocessing capabilities allows users to enhance images, focus on regions of interest, and obtain accurate text extraction results with minimal effort.

By providing options such as noise reduction, contrast enhancement, skew correction, and morphological transformations, the application addresses common challenges in image-based text recognition and helps improve overall OCR accuracy.

While this version is optimized for clear, printed text in standard image formats, future enhancements could expand its capabilities to handle handwritten text, more complex document layouts, and multi-language recognition, further increasing its versatility and usefulness.