

MANUAL for OCR Program (v0.25)

1. Introduction

This is a program based on OCR (Optical character recognition) engine called Tesseract. *"Tesseract is an open source [text](https://en.wikipedia.org/wiki/Optical_character_recognition) Engine, available under the [Apache](<http://www.apache.org/licenses/LICENSE-2.0>) It can be used directly, or (for programmers) using an [API](<https://github.com/tesseract-ocr/tesseract/blob/master/src/api/baseapi.h>) to extract printed text from images. It supports a wide variety of languages."* (cited from <https://github.com/tesseract-ocr/tesseract/wiki>)

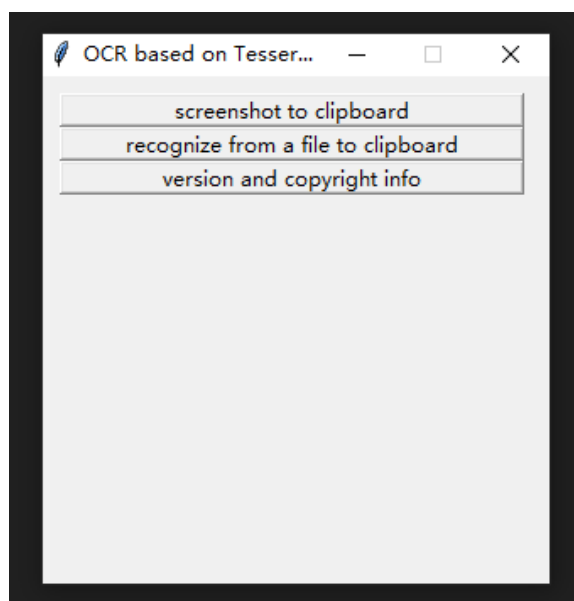
We tried our best to make it easier as much as possible, and we add the following features to the program:

- (1). Add the UI for the engine, provide easier selection.
- (2). Add the function of "Screenshot" and directly to the clipboard.
- (3). Remove some of the language pack to make the file from 0.5GB to 100MB (the file removed can be downloaded from the Github." <https://github.com/tesseract-ocr/tesseract/tree/master/tessdata>")
- (4). Integrate the program to make optimize the file structure. (Special care for *Virgo*)
- (5). Be able to work on the computer without python environment. Minimize the modification on the user's computer.

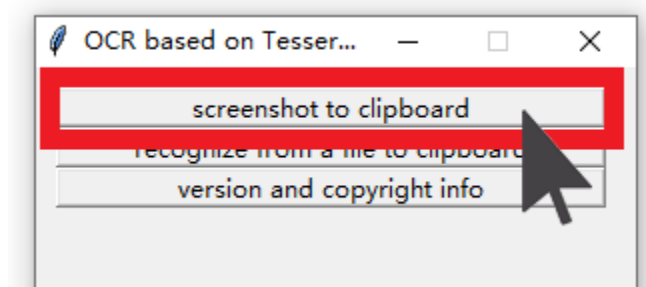
For any future support and release please see "<https://github.com/YZhu2022/OCR-Project/>"

2. The usage of program

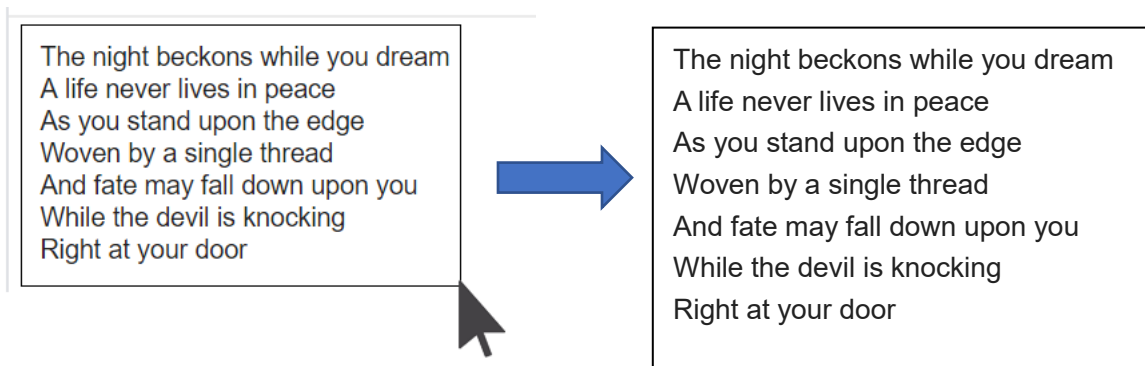
- (1) Double click the "Picker.exe"



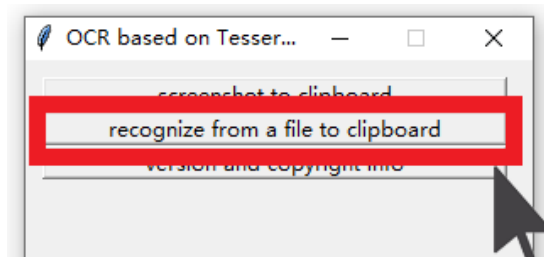
(2) Select the function you want. If you want to take a screenshot, press the first button.



(3) Select the zone you want. Then press the paste button on your computer. (Usually Ctrl + V)



(4) Also, you can select a file directly. Click on the second Button. And press "Paste".



(5) If want to see the copyright and other related information, please click on the final button.

