Project : OCR

DESCRIPTION

Converting handwritten documents as scanned images or photos (in any format) to legible text document using AI extracting important and critical information into database.

IMPLEMENTATION

Created a Text Recognition model using Tesseract-OCR . Model feeds scanned images and prints the extracted text from the image. Flask web framework is used.

LIBRARIES USED

- Cv2
- Tesseract-ocr
- Pytesseract
- flask

EXPLANATION

The above libraries are imported.

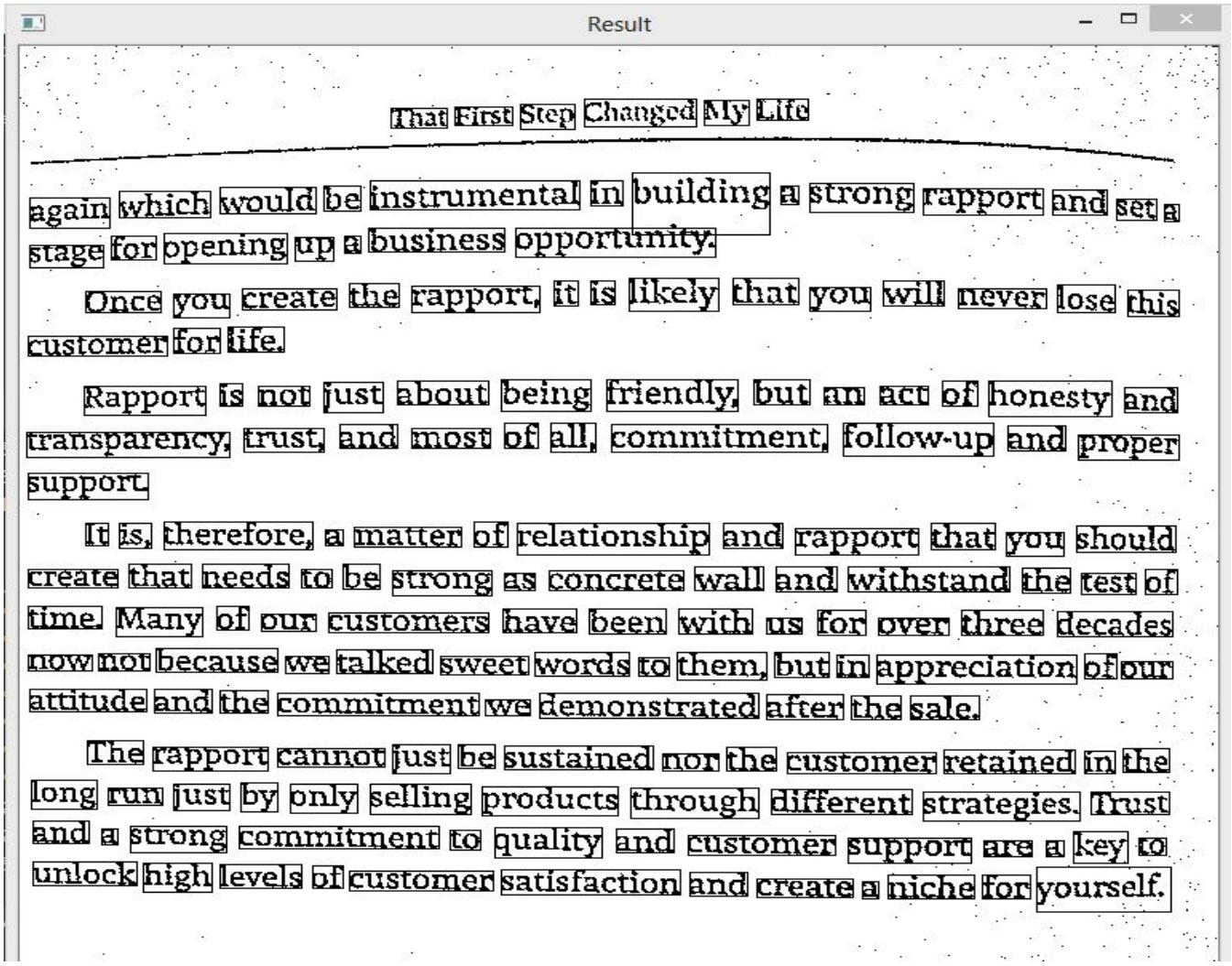
```
img = cv2.imread('trial_2.jpg')
```
- loads image from the specified file

IMAGE PRE-PROCESSING

```
img = cv2.resize(img, None, fx=0.25, fy=0.25)
```
- image resizing helps in reducing the number of pixels from an image

```
thresh_img = cv2.adaptiveThreshold(img, 255,
cv2.ADAPTIVE_THRESH_GAUSSIAN_C, cv2.THRESH_BINARY, 11, 11)
```
-Grayscale image is passed as input to adaptiveThreshold to separate desirable foreground image objects from the background based on difference in pixel intensities of each region.

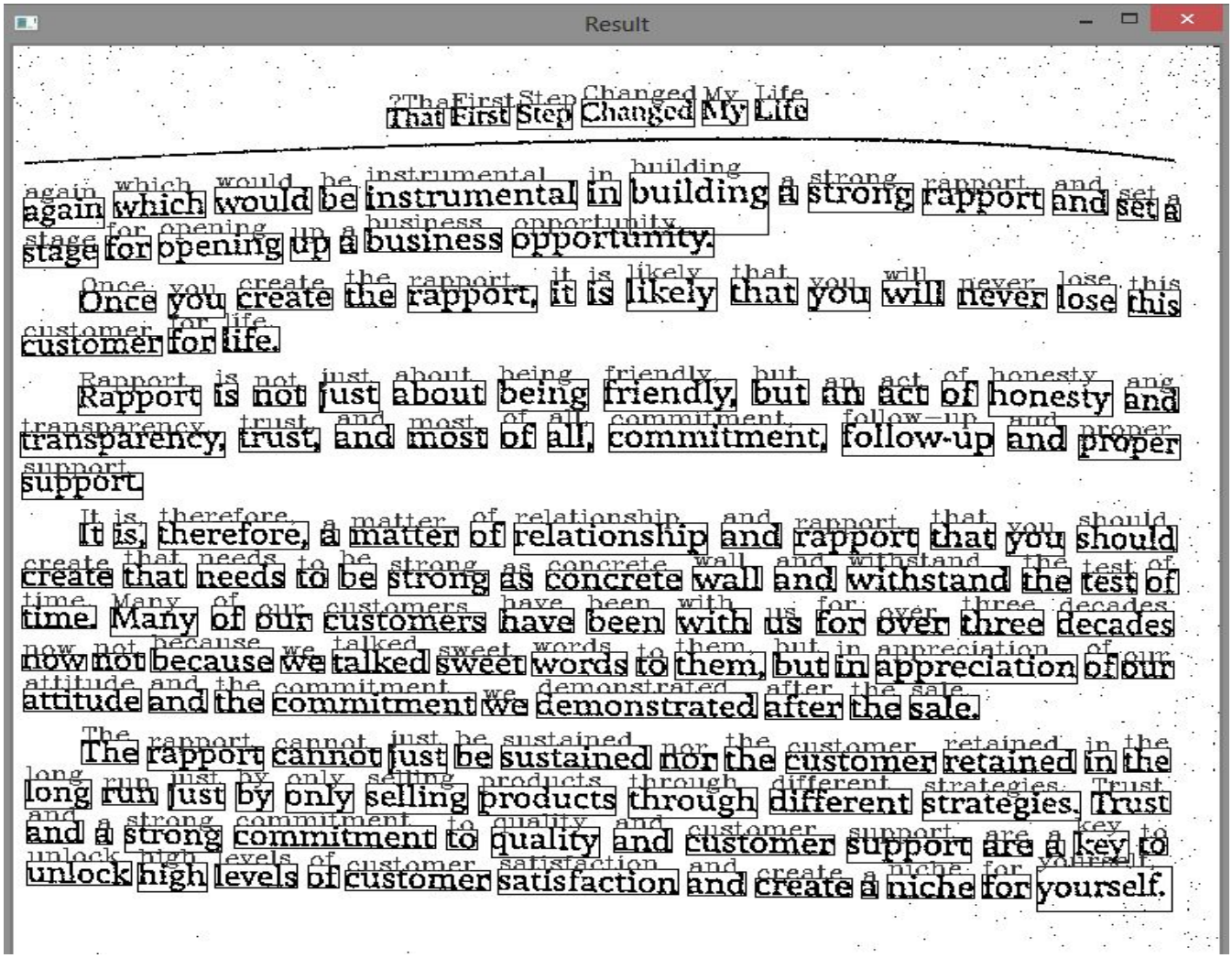
Image after threshold is shown below ,



`boxes = pytesseract.image_to_data(thresh_img)` - Returns string containing box boundaries

```
cv2.putText(thresh_img, b[11], (x, y), cv2.FONT_HERSHEY_COMPLEX, 0.5, (50, 50, 255), 1)
```

Image after putText() ,



`extracted_text = pytesseract.image_to_string(thresh_img)` - Image_to_string() extracts the whole content from image and prints them

FILES

- main.py
- app.py
- templates/upload.html

app.py :-

```
from flask import Flask, render_template, request
```

Flask libraries are imported , render_template is used to render HTML document which will be displayed in the browser.

`ALLOWED_EXTENSIONS = {'png', 'jpg', 'jpeg'}` -Extensions are set such that it would allow files to be uploaded of specific type

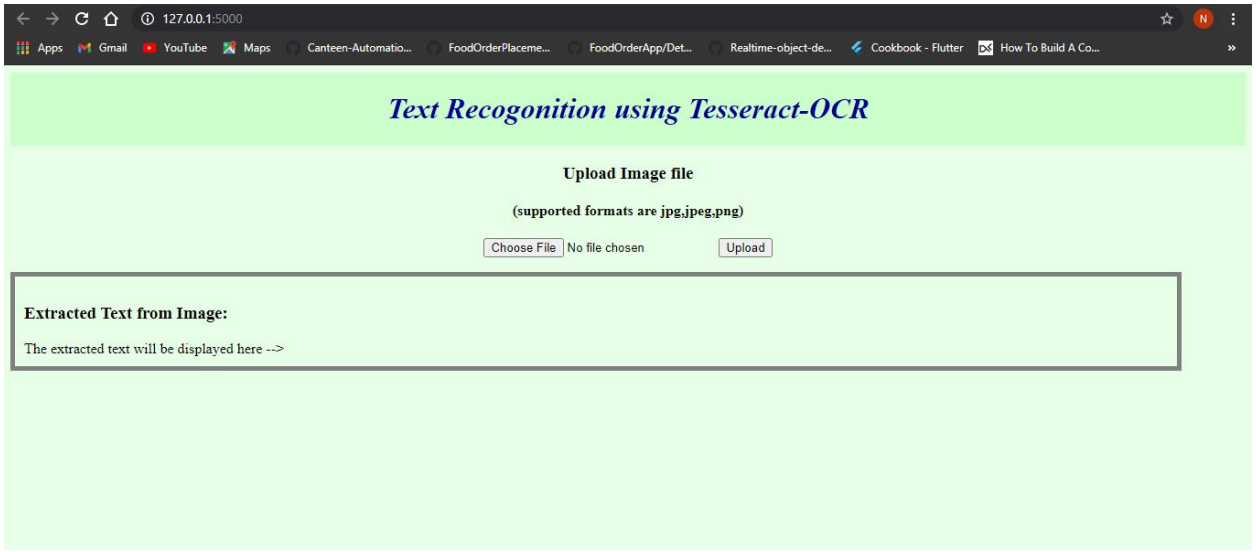
```
# route and function to handle the upload page
@app.route('/', methods=['GET', 'POST'])
def upload_page():
    if request.method == 'POST':
        # check if there is a file in the request
        if 'file' not in request.files:
            return render_template('upload.html', msg='No file selected')
        file = request.files['file']
        # if no file is selected
        if file.filename == '':
            return render_template('upload.html', msg='No file selected')

        if file and allowed_file(file.filename):
            # call the OCR function on it
            file = secure_filename(file.filename)
            extracted_text = main.ocr_main(file)

            # extract the text and display it
            return render_template('upload.html', msg='Successfully processed', extracted_text=extracted_text)
```

The following function takes the uploaded image , main.ocr_main() is called the text in the image is extracted and returned to ‘extracted_text’ which is the passed to render_template()

OUTPUT



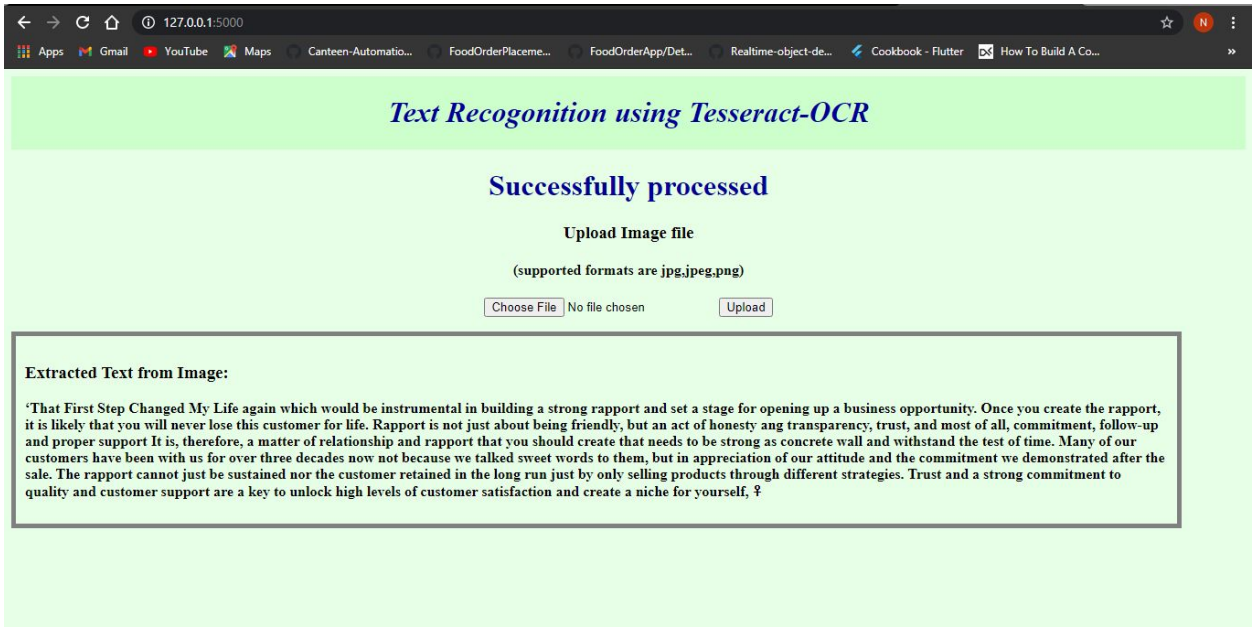
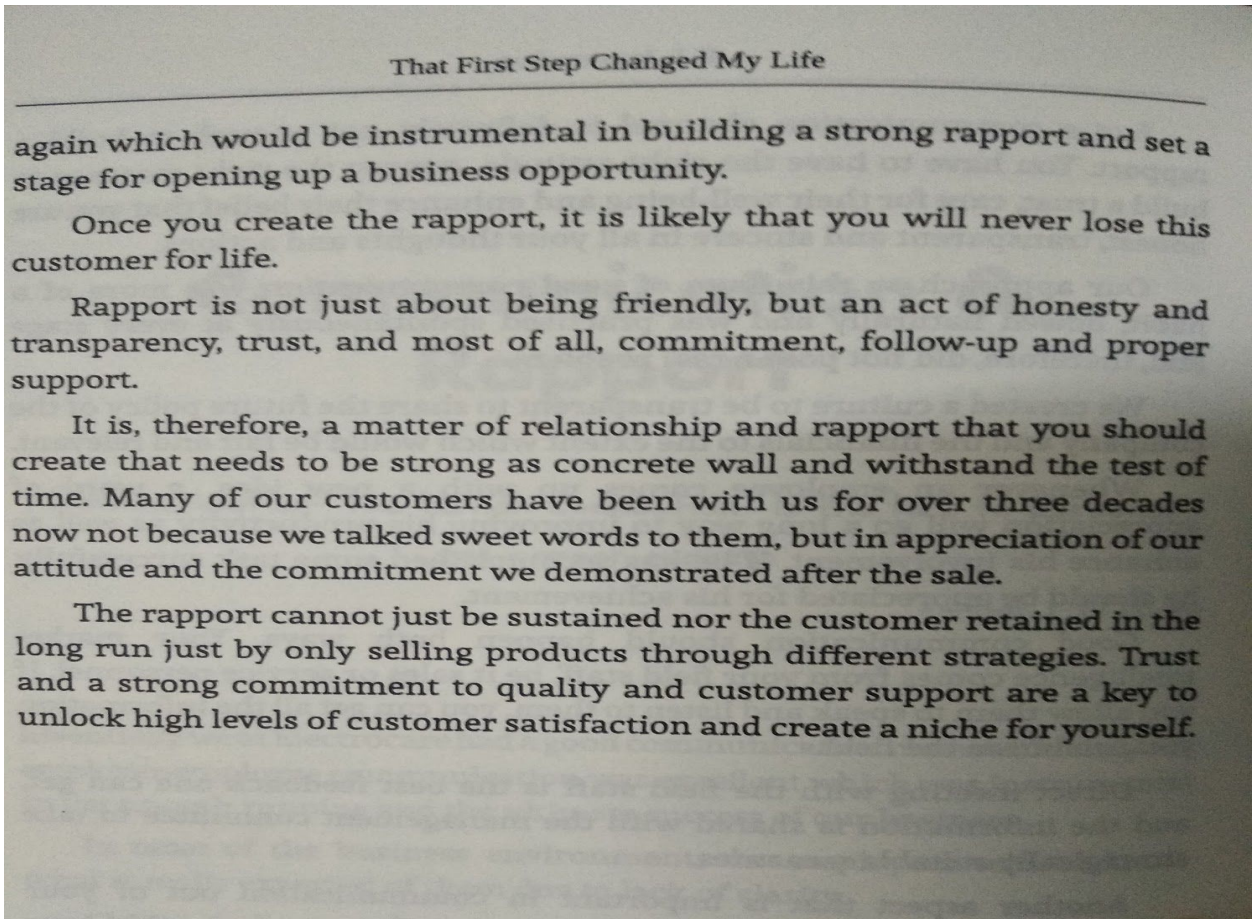

127.0.0.1:5000
Apps
Gmail
YouTube
Maps
Canteen-Automatio...
FoodOrderPlaceme...
FoodOrderApp/Det...
Realtime-object-de...
Cookbook - Flutter
How To Build A Co...
Text Recogonition using Tesseract-OCR
Successfully processed
Upload Image file
(supported formats are jpg,jpeg,png)
Choose File
No file chosen
Upload
Extracted Text from Image:
'That First Step Changed My Life again which would be instrumental in building a strong rapport and set a stage for opening up a business opportunity. Once you create the rapport, it is likely that you will never lose this customer for life. Rapport is not just about being friendly, but an act of honesty ang transparency, trust, and most of all, commitment, follow-up and proper support It is, therefore, a matter of relationship and rapport that you should create that needs to be strong as concrete wall and withstand the test of time. Many of our customers have been with us for over three decades now not because we talked sweet words to them, but in appreciation of our attitude and the commitment we demonstrated after the sale. The rapport cannot just be sustained nor the customer retained in the long run just by only selling products through different strategies. Trust and a strong commitment to quality and customer support are a key to unlock high levels of customer satisfaction and create a niche for yourself, ?

Image used:-


That First Step Changed My Life
again which would be instrumental in building a strong rapport and set a stage for opening up a business opportunity.
Once you create the rapport, it is likely that you will never lose this customer for life.
Rapport is not just about being friendly, but an act of honesty and transparency, trust, and most of all, commitment, follow-up and proper support.
It is, therefore, a matter of relationship and rapport that you should create that needs to be strong as concrete wall and withstand the test of time. Many of our customers have been with us for over three decades now not because we talked sweet words to them, but in appreciation of our attitude and the commitment we demonstrated after the sale.
The rapport cannot just be sustained nor the customer retained in the long run just by only selling products through different strategies. Trust and a strong commitment to quality and customer support are a key to unlock high levels of customer satisfaction and create a niche for yourself.