

PROJECT DEVELOPMENT PHASE

PROJECT DEVELOPMENT – DELIVERY OF SPRINT-1

DATE	08- NOVEMBER-2022
TEAM ID	PNT2022TMID44898
PROJECT NAME	REAL-TIME COMMUNICATION SYSTEM POWERED BY AI FOR SPECIALLY ABLED

IMPORTING NECESSARY LIBRARIES:-

```
In [1]: import cv2
import pytesseract
import os
from PIL import image
import sys
```

READING IMAGE WITH DATA FILES:-

```
In [ ]: def get_string(img_path):
        #read image with opencv
        img=cv2.imread(img_path)
```

REMOVING NOISE FROM DATASET:-

```
In [ ]: #convert to gray
img=cv2.cvtColor(img,cv2.COLOR_BGR2GRAY)
#apply dilation & erosion to remove some noise
kernel=np.ones((1,1),np.uint8)
img=cv2.dilate(img,kernel,iterations=1)
img=cv2.erode(image,kernel,iterations=1)
```

RECOGNISING THE DATASET & CHANGING TEXT TO READ:-

```
In [ ]: #Write the image after apply opencv to do some...
cv2.imwrite("thres.png",img)
#recognize text with tesseract for python
result=pytesseract.image_to_string(image.open("thres.png"))
os.remove("thres.png")

return result
```



CHANGE TEXT TO READ:

For information on obtaining SAFOD core, cuttings and other samples please go to the [EarthScope Web site](#). All PI's currently involved in SAFOD should automatically receive email updates on timetables for requesting SAFOD samples from the EarthScope National Office. If you are not receiving this information or if you are not currently involved in SAFOD and wish to be, please see the Earthscope website or contact the NSF EarthScope Program Coordinator, Greg Anderson (greander@nsf.gov)

CHANGE TEXT TO READ:

Sample distribution, Phases 1 and 2 only (for Phase 3 core distribution and analyses go to www.earthscope.org)

DISPLAY IMAGES FROM DATASET:-

```
In [ ]: if __name__ == '__main__':
        from sys import argv

        if len(argv)<2:
            print("usage python image-to-text.py relative-filepath")
        else:
            print('---start recognize text from image---')
            for i in range(1,len(argv)):
                print(argv[i])
                print(get_string(argv[i]))
                print()

            print()
            print('-----Done-----')
```

SAMPLE IMAGES:-

Sign_img=cv2.imread(train_data_path+'0/0_234.jpeg')

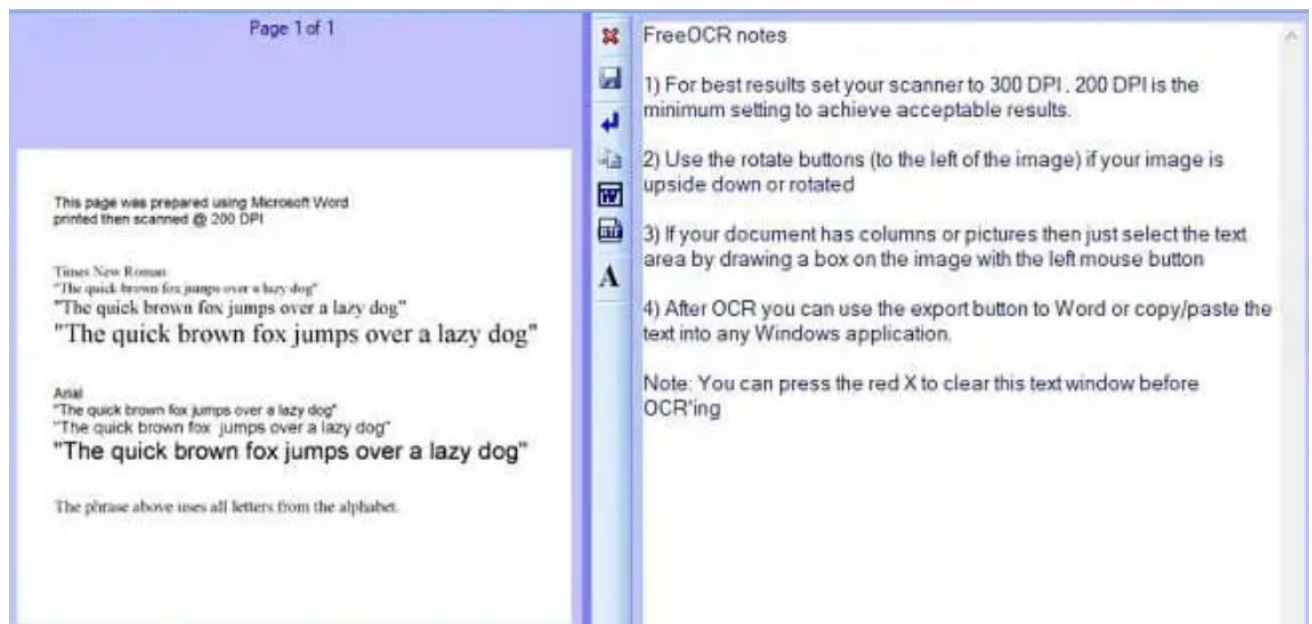
Display(Text_img,'a')

The left image is a green document titled "Endorsement 2456 Leased mortgage additional interest". It contains a paragraph of text explaining the endorsement, followed by a table with columns for "TYPE", "MODEL", "YEAR", "MAKE OF VEHICLE", and "IDENTIFICATION NUMBER". The table has one row with the following data: "1", "2012", "MITS", "5A3CJ566XNR088876". Below the table, there is a section for "LIMITS OF LIABILITY" with two rows: "Property Damage" and "Bodily Injury", both with a limit of "\$100,000 Each". The document also includes a section for "Additional Interest" and a section for "STATEWIDE NATIONAL INSURANCE CO.".

The right image is a white document titled "Endorsement 2456 Leased mortgage additional interest". It contains a paragraph of text explaining the endorsement, followed by a table with columns for "ENTRY NO.", "MODEL YEAR", "MAKE OF VEHICLE", and "IDENTIFICATION NUMBER". The table has one row with the following data: "1", "2012", "MITS", "5A3CJ566XNR088876". Below the table, there is a section for "LIMITS OF LIABILITY" with two rows: "Property Damage" and "Bodily Injury", both with a limit of "\$100,000 Each". The document also includes a section for "Additional Interest" and a section for "STATEWIDE NATIONAL INSURANCE CO.".

Sign_img=cv2.imread(train_data_path+'0/0_235.jpeg')

Display(Text_img,'b')



```
Sign_img=cv2.imread(train_data_path+'0/0_236.jpeg')
```

```
Display(Text_img,'c')
```

