Course Code : 2CS404

Python for Scientific Computing

AIM : Detect the Bill type from the image and manage all same type of bill in one folder.

Roll Number :

1) 18BCE154 (Bhoomi Patel)
2) 18BCE156 (Devanshi Patel)
3) 18BCE171 (Shubh Patel)

Step Included :

1) Select the image using GUI
2) Scan the image
3) Convert into Grey scale and enhance image readability
4) Recognize text
5) Detection of bill based the text
6) Display type of the bill and saving path
7) Save this bill into Directory

Note : Please run this command in Command Prompt.

Used Libraries

- ✦ SCIKIT-IMAGE
- ✦ NUMPY
- ✦ OPENCV
- ✦ IMUTILS
- ✦ TKINTER
- ✦ TESSERACT
- ✦ OS
- ✦ RE

1. SELECTING THE IMAGE

- Here we used the GUI library called TKINTER to select the image when we run the code the dialog box will be open to choose the image from system.

CODE :

```
window = tkinter.Tk()
window.withdraw()
window.overrideredirect(True)
window.geometry('0x0+0+0')
window.deiconify()
window.lift()
window.focus_force()

window.sourceFile = filedialog.askopenfilename(filetypes
= ( ( "Image Files" , ( "*.jpg" , "*.png" , "*.jpeg" ) )
, ( "All Files" , "*" ) ) , parent = window , initialdir
= "/", title = "Please select a image file")

window.destroy()
img_path = window.sourceFile

text = get_string(img_path)
```

2. Scanning the Image

- Scanning image by cv2.imread(img_path) from opencv library.

CODE :

```
img = cv2.imread(img_path)
```

3. Convert into Grey scale and enhance image readability

- To connect image into grey scale we used cv2.cvtColor (img , cv2.COLOR_BGR2GRAY) from opencv.
- To reduce some noise we applied cv2.dialate( ) and cv2.erode( ) function from opencv.
- To get image only Black and White we applied cv2.adaptiveThreshold( ) function from opencv.

CODE :

```
img = cv2.imread(img_path)
img = cv2.cvtColor(img, cv2.COLOR_BGR2GRAY)
kernel = np.ones(( 1, 1), np.uint8)
img = cv2.dilate(img, kernel, iterations = 1)
img = cv2.erode(img, kernel, iterations = 1)
img = cv2.adaptiveThreshold(img, 255,
cv2.ADAPTIVE_THRESH_GAUSSIAN_C, cv2.THRESH_BINARY, 31,
2)
```

4. Recognize text

- To use pytesseract library first we have to give a path of pytesseract application file by writing pytesseract.pytesseract.tesseract_cmd = r"C:\Program Files (x86)\Tesseract-OCR\tesseract.exe”
- To detect the string form the image we used pytesseract.image_to_string (img) method from pytesseract library.

CODE :

```
pytesseract.pytesseract.tesseract_cmd = r "C:\Program
Files (x86)\Tesseract-OCR\tesseract.exe"
result = pytesseract.image_to_string(img)
```

5. Detection of bill based the text

- To detect type of bills first we define some word that can easily recognize bill type.
- We store all that word in one dictionary.
- For efficient bill recognization we count that in each bill type what is the maximum word count that given bill has and based on the maximum value we decide what type of image it has.
- For doing this we also set a path that where this image is store.

CODE :

```
Bills = { 'Electricity Bill' : ['torrent','electricity
bill','electricity', 'electric', 'kwh','elecrtic usage', 'electric
supply', 'energy', 'generation charges','utility', 'energy charges'] ,
'Mobile Bill' : ['mobile services', 'internet bill', 'mobile' , 'phone'
, 'mobile bill', 'postpaid bill', 'airtel', 'vodafone', 'idea',
'bsnl','jio', 'prepaid','postpaid', 'broadband', 'fixline'] , 'Water
Bill' : [ 'water' , 'water bill', 'water supply', 'water usage', 'bore
well charge', 'water budget', 'water utility bill', 'water
conservation'] , 'Gas Bill' : ['gas','hp gas', 'natural gas', 'gas
bill', 'gas piplines', 'gas charges', 'gas limited', 'gas supply', 'gas
use'] , 'Restaurant Bill' : ['restaurant', 'hotel', 'dhaba', 'panjabi',
'chinese', 'kathiyawadi', 'italian', 'maxican', 'coffee', 'tea', 'fast
food', 'dinner', 'lunch','breakfast', 'food', 'cafe'] , 'Shopping Bill'
: ['mall','shop', 'supermarket', 'shopping', 'megamarket', 'dmart',
'retail store', 'cloth', 'store'] , 'Hospital Bill' : ['hospital',
'medical', 'laboratory', 'pharmacy', 'surgical', 'physical
therapy','therapy', 'recovery', 'blood', 'emergency', 'surgery',
'pathology', 'health', 'radiology', 'cardiology'] , 'Newspaper Bill' :
['newspaper','news', 'papers', 'the times of india', 'hindustantimes',
'the indian express', 'news paper', 'divyabhasker', 'sandesh','gujarat
samachar', 'magazine'] }


Guess = { "Electricity Bill" : 0 , "Mobile Bill" : 0 , "Water Bill" : 0
, "Gas Bill" : 0 , "Restaurant Bill" : 0 , "Shopping Bill" : 0 ,
"Hospital Bill" : 0 , "Newspaper Bill" : 0 }
```

6. Display type of the bill and saving path

- Print type of bill on the Command Prompt.
- Print the path where image will be saved.

CODE :

```
path = "./Bills/"
subpath = ""
for i in Bills :
    count = 0
    for j in Bills[i] :
        if j in text :
            Guess[i] += 1
value = list(Guess.values())
bill = list(Guess.keys())
flag = False
print("\n\n\n\t\t\t***" , end = '  ')
if max(value) != 0 :
    for i in Guess:
        if Guess[i] == max(value) :
            if flag :
                path += "OTHER/"
                subpath += "OR "
                print("OR",end = ' ')
            subpath += (i + " ")
            print(i , end = ' ')
            flag = True
        path += (subpath.strip( ) + '/')
else :
    print("Please Upload Clear Image..... Not
Detected" , end = ' ')
    path += ("OTHER/")
print(" ***")
```

7. Save this bill into Directory

- Save the image to the created path.

CODE :

```
now = datetime.now()
dt_string = now.strftime("%Y_%m_%d_%H_%M_%S")

if not os.path.exists(path) :
    os.makedirs(path)

path += ( 'IMG_' + dt_string + '.jpg' )
print("\nYour Bill is saved in this Location :",path)
cv2.imwrite(path,img)
```