


Files



..

sample_data

DogsAsHumanCompanions.pdf

MyPetDog.png

page_1.jpg

page_2.jpg

page_3.jpg

page_4.jpg

page_5.jpg

70.82 GB available

Practical 1

Use pytesseract library in Python for optical character recognition from

1. an image file
2. a multi-page pdf file

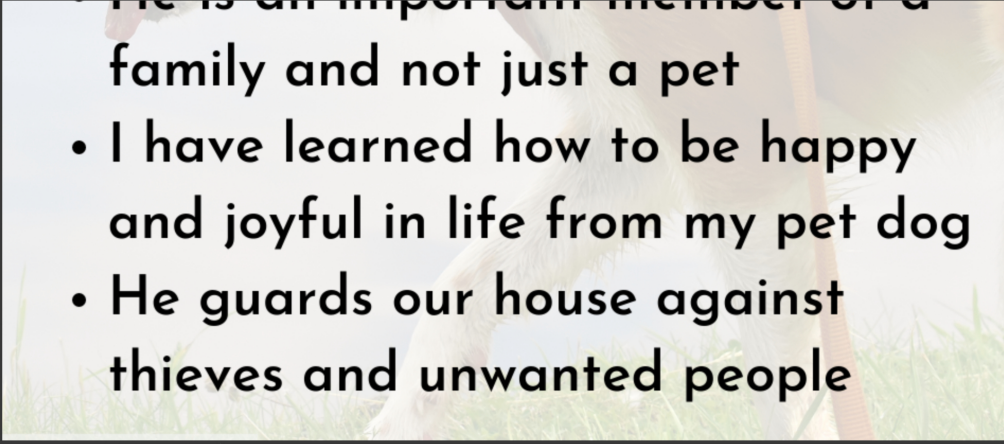
```
[5] 1 # Installation of required packages in google colab
    2 !pip install pytesseract
    3 !sudo apt install tesseract-ocr
```

Requirement already satisfied: pytesseract in /usr/local/lib/python3.7/dist-packages (0.3.8)
Requirement already satisfied: Pillow in /usr/local/lib/python3.7/dist-packages (from pytesseract) (7.1.2)
Reading package lists... Done
Building dependency tree
Reading state information... Done
tesseract-ocr is already the newest version (4.00-git2288-10f4998a-2).
The following package was automatically installed and is no longer required:
 libnvidia-common-460
Use 'sudo apt autoremove' to remove it.
0 upgraded, 0 newly installed, 0 to remove and 40 not upgraded.

```
1 # Importing packages
2 import pytesseract
3 import shutil
4 import os
5 import random
6 try:
7     from PIL import Image
8 except ImportError:
9     import Image
10 from google.colab.patches import cv2_imshow
11 import cv2
12 from matplotlib import pyplot as plt
```

```
[7] 1 # Uploaded png image named 'MyPetDog'
    2 from google.colab import files
    3 # uploaded = files.upload()
```

```
[8] 1 # Setting path and extracting text.
    2 pytesseract.pytesseract.tesseract_cmd = r'/usr/bin/tesseract'
    3 img = cv2.imread('MyPetDog.png')
    4 img1 = Image.open('MyPetDog.png')
    5 cv2_imshow(img)
    6 plt.imshow(img1)
    7 ocrinfo = pytesseract.image_to_string(img1) #can be either img or img1; both works!
    8 print('Text extracted from the image : ', '\n', ocrinfo)
```



Text extracted from the image :

Essay on My Pet Dog

« My pet dog is my best friend in the whole world

« My pet dog shows unconditional love and loyalty towards me and my family

« He is an important member of a family and not just a pet

« I have learned how to be happy and joyful in life from my pet dog

« He guards our house against thieves and unwanted people



PDF to Text

```
[9] 1 !pip install pdf2image
```

Requirement already satisfied: pdf2image in /usr/local/lib/python3.7/dist-packages (1.16.0)
Requirement already satisfied: pillow in /usr/local/lib/python3.7/dist-packages (from pdf2image) (7.1.2)

```
[10] 1 from PIL import Image
2 import pytesseract
3 import sys
4 from pdf2image import convert_from_path
5 import os
6 import pathlib
```

```
[11] 1 # !apt-get install poppler-utils
2 # pathlib.Path('/generated_images/').mkdir(parents=True, exist_ok=True)
```

Reading package lists... Done
Building dependency tree
Reading state information... Done
poppler-utils is already the newest version (0.62.0-2ubuntu2.12).
The following package was automatically installed and is no longer required:
libnvidia-common-460
Use 'apt autoremove' to remove it.
0 upgraded, 0 newly installed, 0 to remove and 40 not upgraded.

```
[12] 1 PDF_file = "DogsAsHumanCompanions.pdf"
2 pages = convert_from_path(PDF_file, 500)
3 image_counter = 1
4 for page in pages:
5     filename = "page_"+str(image_counter)+".jpg"
6     page.save(filename, 'JPEG')
7     image_counter = image_counter + 1
8
9 filelimit = image_counter-1
10
11 extracted_text = ''
12
13 for i in range(1, filelimit + 1):
14     filename = "page_"+str(i)+".jpg"
15     text = str((pytesseract.image_to_string(Image.open(filename))))
16     text = text.replace('\n', ' ')
17     extracted_text += text
```

```
[13] 1 print('Extracted text : \n',text)
```

Extracted text :
their pets, the dog owners averaged 35.3 hours per week and the cat owners averaged 33.2 hours. For dog owners, 44% of this time was estimated as play, as compared with 36% for cat owners (J. Angus, personal communication).

Fig. 12.5. Attraction of young children to animals. Young toddlers respond to both mechanical and live dogs, but a real dog elicits the stronger interest (Kidd & Kidd, 1987). Photograph: Joan Borinstein.

Touch

A study of three- to four-year-old children's interactions with dogs revealed that 67% of these interactions involved body contact with the dog, such as putting a hand on the dog, patting it or hitting it. In contrast, vocal and verbal behavior comprised only 9% of the interactions (Milot & Filiatre, 1986). In a subsequent study touching was again the most frequent behavior shown in the presence of a dog, accounting for 40% of all child-dog interactions (Filiatre et al., 1988).

In an analysis of 1105 photographs of dogs or cats in a family setting submitted to a national photographic contest, Katcher & Beck (1985) found that 97% of the pictures illustrated people and animals touching each other, generally with the heads of the animal and human close together. Over 92% showed a dyadic relationship, with one person and one animal occupying the center of the photograph. Touching was also a primary mode of interaction with a dog in a study of nursing home residents (Neer, Dorn & Grayson, 1987). Of the nine different types of interaction recorded involving the dog, grooming and touching were the two most commonly employed by residents.

The value of dogs for different types of people

Albert & Bulcroft's (1987, 1988) Rhode Island study found that households with children at home tended to have more pets than either widows or families with an 'empty nest', or with an infant. However, feelings of attachment to the pet were lowest in families where children were at home. Although pet ownership was highest among households containing large families, attachment to pets was highest among people living alone and among couples who did not have children living at home. The authors noted that the single, divorced and widowed individuals and childless couples who were most attached to their pets also expressed more anthropomorphic attitudes to their pets, particularly in relation to dogs. In a longitudinal study of older people (a population that experiences increasing losses), Lago, Connell & Knight (1985) found that persons who stayed at home and spent more time with the animal also became more attached and formed a stronger relationship with it.

Stronger Relationship with TV

An 'invisible cord' often seems to connect a dog to its owner (Serpell, 1986a). Almost invariably, dogs are more attentive to their owners than their owners are to them. In a study of ten families' interactions with their dogs, the associations between the dog and the adult family members were found to differ between families with and without children (Smith, 1983). In childless families the people and the dog interacted more

✓ 0s completed at 11:11 PM

