

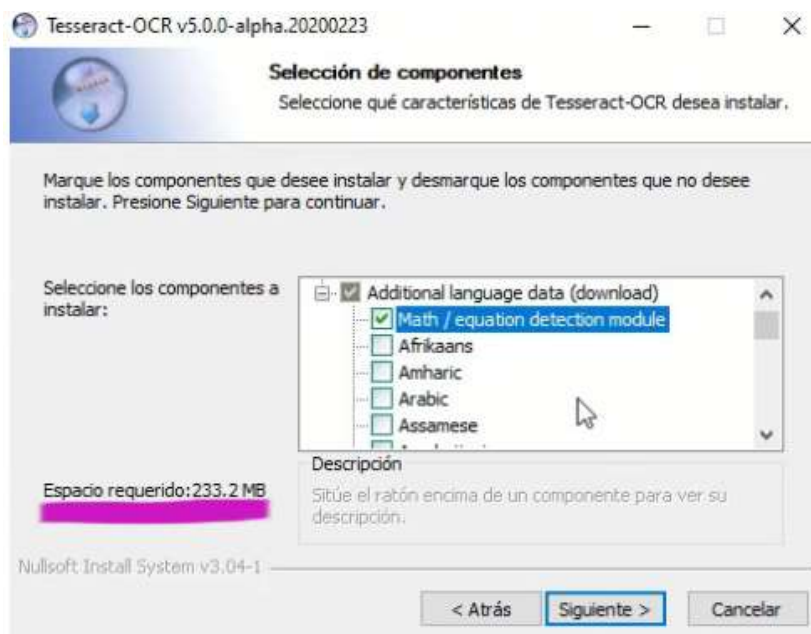
Reconocimiento Óptico de Caracteres (OCR)

Como primer paso vamos a instalar Tesseract – OCR del siguiente link <https://github.com/UB-Mannheim/tesseract/wiki>

Una vez descargado, ejecutamos el instalador. Escogemos un idioma para realizar el proceso.



Aceptamos los términos y en la sección de *Selección de componentes* vamos a ver que por defecto se instalan los datos del idioma inglés. Si deseamos otros idiomas vamos a expandir *Adicional language data*, en donde podremos elegir aquellos que deseemos instalar.



Una vez finalizada la instalación vamos a dar clic en Terminar.

CÓDIGO

```
lector de texto.py - C:\Users\kv460\Documents\lector detexto\lector de texto.py (3.8.5)
File Edit Format Run Options Window Help

import cv2
import pytesseract } Librería
#donde esta el programa instalado
pytesseract.pytesseract.tesseract_cmd = r'C:/Program Files/Tesseract-OCR/tesseract.exe'
#leer imagen
pruebal = cv2.imread('a.jpg')
#reconocimiento optico de caracteres esto se almacena
#en la variable text. y aqui imprimimos lo que se encuentre
text = pytesseract.image_to_string(pruebal, lang='spa')
print('Texto: ',text)
#vizualizamos la imagen
cv2.imshow('Image',pruebal)
#cerramos y borramos todo
cv2.waitKey(0)
cv2.destroyAllWindows()
```

Para evitar errores, cuando no se detecta el programa instalado.

Se realiza el reconocimiento óptico en idioma español

RESULTADO



Imagen donde se detectará el texto

```
Python 3.8.5 (default, Sep 3 2020, 21:29:08) [MSC v.1916 64 bit (AMD64)] on win
32
Type "help", "copyright", "credits" or "license()" for more information.
>>>
===== RESTART: C:\Users\kv460\Documents\lector detexto\lector de texto.py =====
Texto:

ARDUINO -

EFD

OPEN=SOURCE ELECTRONICS PLATFORM
DUDO CCODOO

E DESIGNED ANO.
HA cesto mm rraix
```

Texto detectado

Probamos con otra imagen con mas texto

```
import cv2
import pytesseract
#donde esta el programa instalado
pytesseract.pytesseract.tesseract_cmd = r'C:/Program Files/Tesseract-OCR/tessera
#leer imagen
pruebal = cv2.imread('tex.jpg')
#reconocimiento optico de caracteres esto se almacena
#en la variable text. y aqui imprimimos lo que se encuentre
text = pytesseract.image_to_string(pruebal, lang='spa')
print('Texto: ', text)
#vizualizamos la imagen
cv2.imshow('Image', pruebal)
#cerramos y borramos todo
cv2.waitKey(0)
cv2.destroyAllWindows()
```

Solo se modificó la imagen que se debe leer.

RESULTADO

TEXTOS DESCRIPTIVOS

DEFINICIÓN

- Un texto es un discurso oral o escrito que tiene coherencia interna propia
- Se otorga información al lector u oyente de forma que pueda 'fabricar' sus propias representaciones mentales.
- Consta de descripciones que sirven al interlocutor para hacerse una idea de aquello de lo que está hablando.

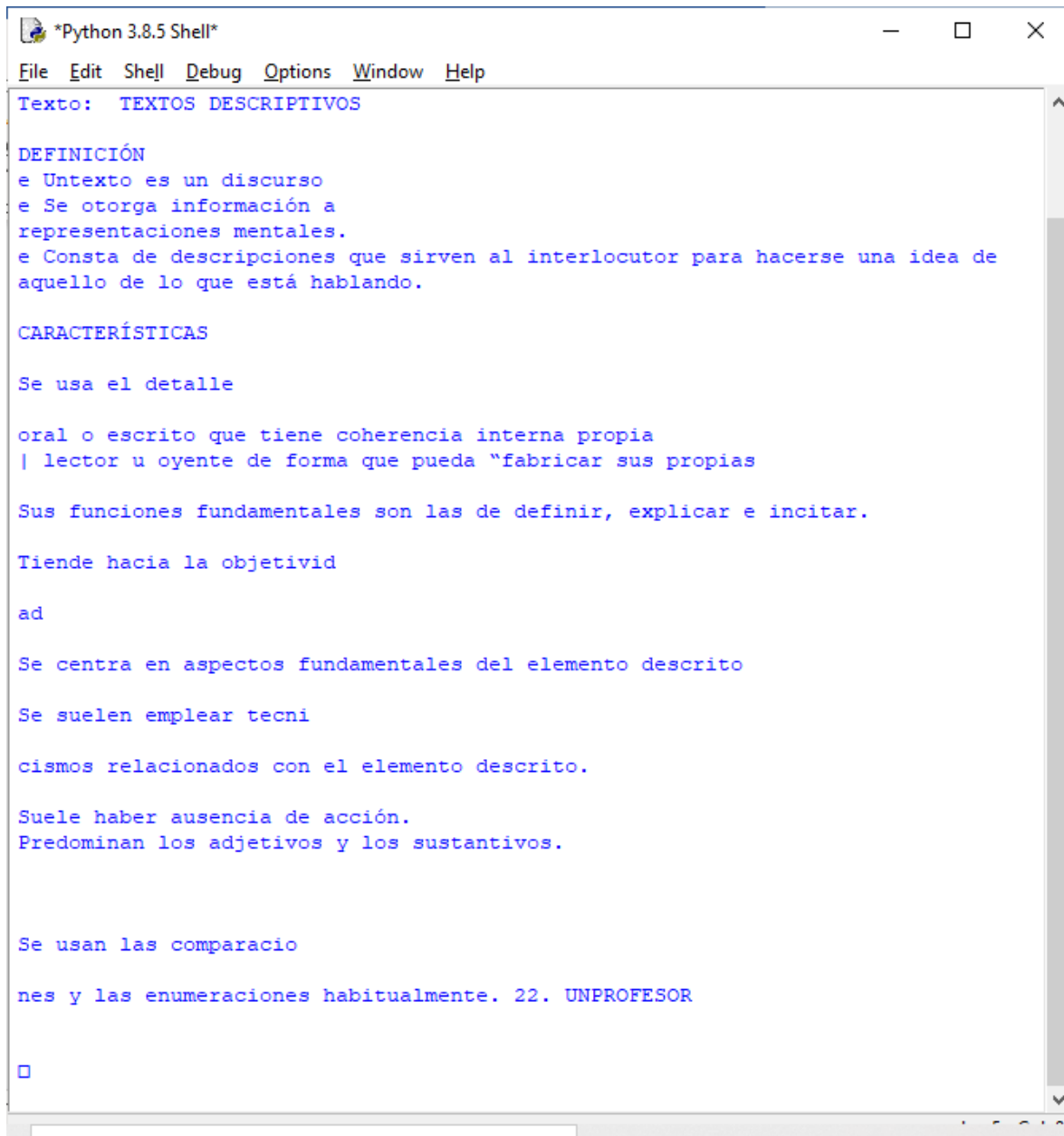
CARACTERÍSTICAS

- Se usa el detalle
- Sus funciones fundamentales son las de definir, explicar e incitar.
- Tiende hacia la objetividad
- Se centra en aspectos fundamentales del elemento descrito
- Se suelen emplear tecnicismos relacionados con el elemento descrito.
- Suele haber ausencia de acción.
- Predominan los adjetivos y los sustantivos.
- Se usan las comparaciones y las enumeraciones habitualmente.



Imagen donde se detectará el texto

Texto detectado



```
*Python 3.8.5 Shell*
File Edit Shell Debug Options Window Help
Texto: TEXTOS DESCRIPTIVOS

DEFINICIÓN
e Un texto es un discurso
e Se otorga información a representaciones mentales.
e Consta de descripciones que sirven al interlocutor para hacerse una idea de aquello de lo que está hablando.

CARACTERÍSTICAS

Se usa el detalle

oral o escrito que tiene coherencia interna propia
| lector u oyente de forma que pueda "fabricar sus propias

Sus funciones fundamentales son las de definir, explicar e incitar.

Tiende hacia la objetividad

ad

Se centra en aspectos fundamentales del elemento descrito

Se suelen emplear tecnicismos relacionados con el elemento descrito.

Suele haber ausencia de acción.
Predominan los adjetivos y los sustantivos.

Se usan las comparaciones y las enumeraciones habitualmente. 22. UNPROFESOR

□
```