

02 Preparando o ambiente

cursos.alura.com.br/course/visao-computacional-reconhecimento-texto-ocr-opencv/task/113539

Iremos utilizar o Tesseract OCR no Google Colaboratory com auxílio da biblioteca Pytesseract e os dados virão diretamente do GitHub. Para utilizar o Tesseract OCR no Colab é necessário instalá-lo e instalar bibliotecas complementares, como fizemos em aula.

```
!pip install opencv-python==4.6.0
!sudo apt install tesseract-ocr
!pip install pytesseract==0.3.9
```

Após esse passo, o ambiente virtual deverá ser reiniciado para que as mudanças, no caso as instalações, sejam realmente efetivadas. Depois de reiniciar o ambiente é necessário fazer a importação das bibliotecas.

```
import pytesseract
import numpy as np
import cv2
from google.colab.patches import cv2_imshow
```

Os dados serão importados diretamente de um repositório do GitHub, então utilizaremos o comando

```
!git clone :
```

```
! git clone https://github.com/sthemonica/text-recognize
```

Na aba Arquivos do Colab, todos os arquivos usados durante o curso, inclusive nas atividades, estão disponíveis.

[Discutir no Fórum Próxima Atividade](#)

