

Visão por Computador 2023/2024
Projeto de Sopa de Letras
Bruno Duarte - 118326
Tomás Taxa -121863

O nosso projeto de Visão por Computador é um programa que resolve sopas de letras.

O programa tem os seguintes passos de execução:

- Primeiro lemos a imagem, onde assumimos que as letras estarão a preto e todas maiúsculas, isto pois quando tínhamos letras em minúsculo em alguns casos havia confusão entre o 'L' minúsculo e o 'I' maiúsculo, dependendo da *font* que usávamos.

- Após lermos a imagem, usamos operações de dilatação até termos um aglomerado com as letras da sopa, e assumimos que os restantes aglomerados são as palavras que queremos procurar. Para achar esses aglomerados, bem como os tamanhos deles usamos a função `cv2.connectedComponentsWithStats()`, que nos dá o número de aglomerados, bem como a posição mais acima e à direita e uma largura e comprimento para criarmos um retângulo que envolva todo o aglomerado.

- Após isso, temos que achar as letras. Inicialmente usávamos operações de fecho (3 de dilatação e 3 de erosão) para as letras ficarem mais preenchidas e assim se identificarem melhor, no entanto verificámos que no caso das palavras que queríamos às vezes 2 letras ficavam coladas, o que não era desejável, pelo que decidimos detetar as letras como estas apareciam na imagem usando a função `cv2.connectedComponentsWithStats()`, que nos retorna o número de aglomerados, (neste caso letras), dentro da imagem, e outros dados úteis para construir um retângulo à volta da letra. Usando esses mesmos dados cortamos a letra para verificarmos a qual carater corresponde.

- Para isso, inicialmente usámos `TemplateMatching`, no entanto isso revelou-se não muito preciso em alguns casos, por exemplo 'V' e 'Y', em que de vez em quando tínhamos uma troca entre ambos. Por isso decidimos usar a livreria `Tesseract`, que usa modelos de aprendizagem para detetar letras, à custa de ser relativamente mais lento. Obtendo o retângulo que engloba a letra chamamos a função `pytesseract.image_to_string()` que nos devolve os caracteres presentes numa imagem. Nós colocámos como opções para essa função que só queríamos caracteres únicos, pois passamos uma letra de cada vez, bem como só queríamos as letras maiúsculas e o 'L' minúsculo e o número zero. O 'L' e o 0 foram adicionados pois em alguns casos os O's eram mais "magros", pelo que o programa não os detetava, então nós colocamos o zero como sendo mais uma possibilidade para 'O'. O 'L' minúsculo foi adicionado pelo mesmo motivo mas para a letra 'I'.

- Obtendo todas as letras da sopa e das palavras desejadas usamos um algoritmo para resolver a sopa de letras e no fim damos print a tuplos que correspondem a uma palavra, a posição inicial e final da mesma, e ao sentido (de trás para frente ou vice-versa) no qual a palavra aparece. Este algoritmo funciona bem, achando sempre solução.

Um erro que tivemos foi que às vezes ao achar as letras a letra 'A' era considerada mais baixa que as outras, pelo que há casos em que ela aparece no fim da linha quando não devia, pelo que algumas palavras ficam sem solução devido a essa troca, no entanto a maioria das mesmas tem solução.