

Relatório 1 - Vídeo e Leitura: Fundamentos de Visão Computacional

Lucas Scheffer Hundsdorfer

Descrição da atividade

Dentro do vídeo passado, é feito a diferenciação entre o que é processamento de imagem e o que é visão computacional. E a diferença que é apontada, é que o processamento de imagem é quando a entrada é uma imagem e a saída também uma imagem, como exemplo foi citado a aplicação de um filtro:



Original Image



Filtered Image

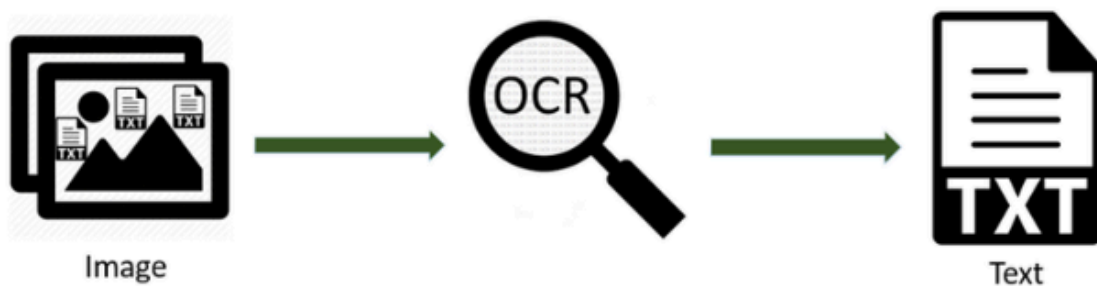
Pelo outro lado, visão computacional tem como entrada uma imagem também mas a saída é informação. O exemplo citado foi o reconhecimento de faces, onde a imagem de entrada é um rosto e a saída seria a identidade dessa pessoa. Também é citado que existem alguns casos onde a saída também pode ser uma imagem como uma foto panorama com a junção de várias imagens de um mesmo cenário, a entrada são várias imagens e a saída um panorama, porém isso acontece pelo fato que o computador precisa entender onde cada foto vai juntar na outra, e é essa a informação que ele recebe e já faz a junção das imagens.

Agora se aprofundando no que é de fato visão computacional, se trata de um campo dentro da área de inteligência artificial que faz o que foi dito já, recebe imagens e devolve informação, como se o computador conseguisse interpretar a imagem. Funciona basicamente com 4 passos: coleta de dados, pré-processamento dos dados, escolha do modelo e por fim o treinamento do modelo. Vou usar de exemplo que quero criar um modelo de identificação de doenças oculares, então o primeiro passo seria coletar imagens de olhos e suas classes, como se o olho está normal, se tem catarata entre outros. Já a parte de pré-processamento que consiste em melhorar fazer uma limpeza nos dados aumentando a qualidade das imagens, como ajustar o brilho e reduzir o ruído já que é de suma importância que os dados estejam corretos e utilizáveis para que tenha um resultado agradável. O terceiro passo é a escolha do modelo adequado, aqui é a escolha do modelo pode ser feita de várias maneiras, como escolher o tipo de rede neural que vai ser utilizada como uma CNN ou uma RNN e isso vai depender do tipo de dado que você tem, no meu exemplo como são imagens o mais recomendado é a CNN. Após decidir qual rede neural é necessário escolher se vai ser criado um modelo apenas para essa tarefa ou se vai utilizar de um modelo pré treinado para essa tarefa específica. E o quarto e último passo é o treinamento do modelo, envolve colocar o modelo para rodar com base no dataset separado anteriormente e medir a performance dele no dataset de teste para analisar o quão assertivo o modelo foi, envolve toda uma parte de otimização de parâmetros para que consiga maximizar o resultado do modelo.

Dentro do capítulo de introdução do livro do Szeliski foram passados algumas áreas de aplicação da visão computacional e vou falar um pouco sobre algumas:

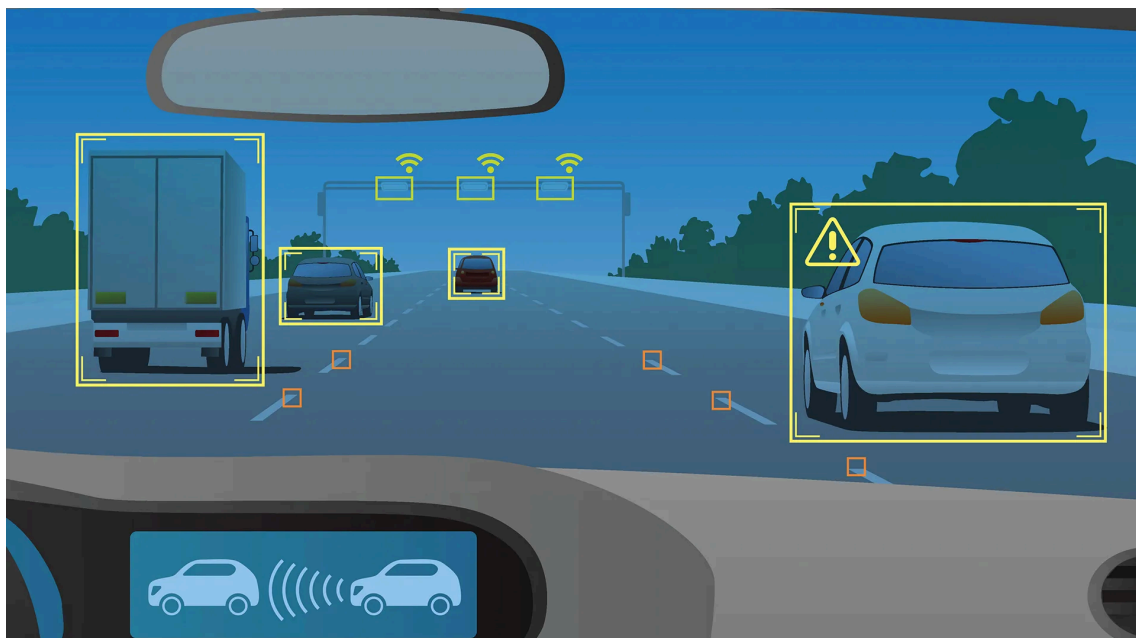
- **Optical character recognition (OCR)**

Reconhecimento óptico de caracteres é o que o próprio nome já diz, um reconhecedor de texto que basicamente extrai textos de imagens, documentos escaneados, pdfs e entre outros. Utiliza algoritmos de visão computacional para detectar e segmentar caracteres também em ambientes não controlados, como placas de ruas ou rótulos de produtos.



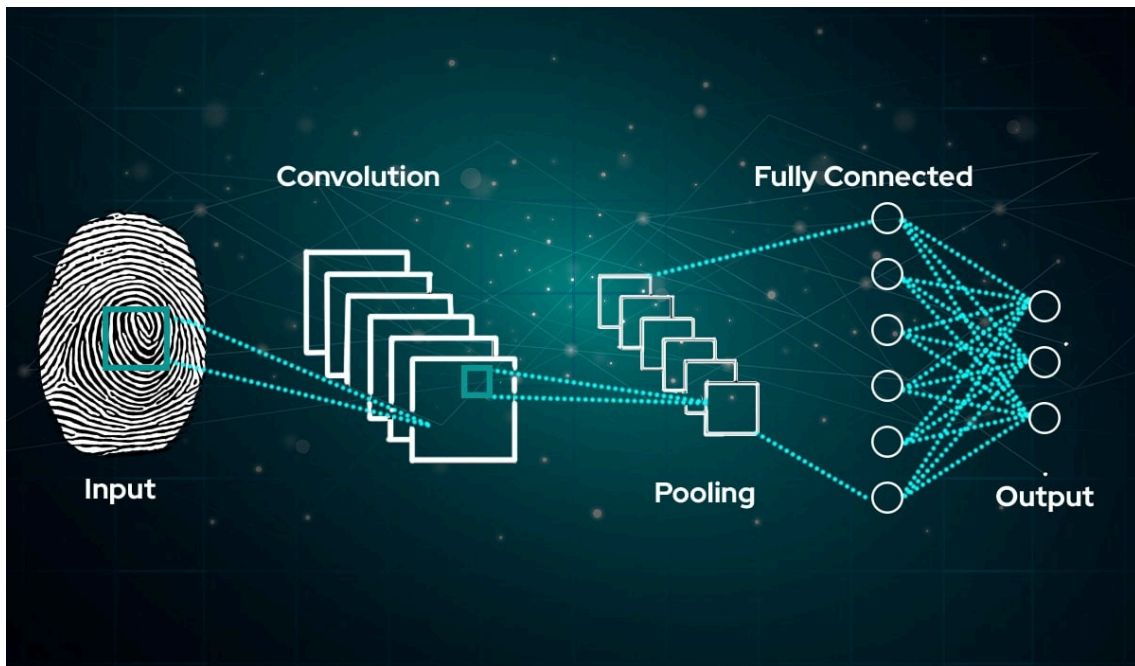
- **Self-driving vehicles**

Os veículos autônomos utilizam de algoritmos da visão computacional para que o computador aja como os olhos do carro usando câmeras e sensores para que consiga interpretar e reconhecer objetos para que o carro consiga fazer o percurso corretamente.



- **Fingerprint recognition and biometrics**

Algo que já é usado no dia a dia, presente em quase todos os celulares modernos hoje, e é baseado em visão computacional, o reconhecimento de padrões baseados em dados que fazem com que exista essas biometrias hoje.



Dificuldades

Nenhuma dificuldade, o máximo que posso dizer já que não sou fluente é que o conteúdo é em inglês mas é algo bem tranquilo de entender.

Conclusões

A conclusão que eu chego é que o que foi dito no vídeo é uma boa forma de descrever o que é a visão computacional, entrada de imagem que retorna informações. Com essa explicação dá para ter um bom entendimento do que é e de como funciona e do porque a visão computacional está tão presente no dia a dia sem mesmo sabermos, já que é algo valioso demais fazer com o computador compreenda o que está na imagem e te retorne informações.

Referências

 Image Processing VS Computer Vision: What's The Difference?

<https://szeliski.org/Book/download.php>

[What Is Computer Vision? | IBM](#)

[What Is Optical Character Recognition \(OCR\)? | IBM](#)

[Understanding Computer Vision in Self-Driving Cars](#)