

Análise de Imagens e Visão Computacional

Prof. Henrique Batista da Silva

Introdução

Introdução

Um problema de visão computacional é usar uma imagem para inferir algo sobre o mundo.

O objetivo é “entender” o conteúdo visual de uma imagem.

Referência: Simon J. D. Prince. **Computer Vision: Models, Learning, and Inference 1st Edition**

Introdução



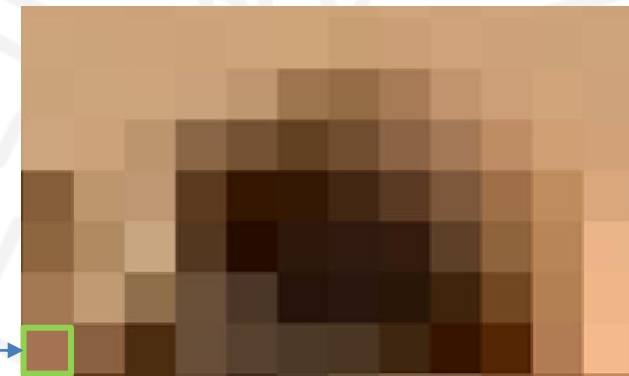
Fonte: canaltech.com.br

Introdução

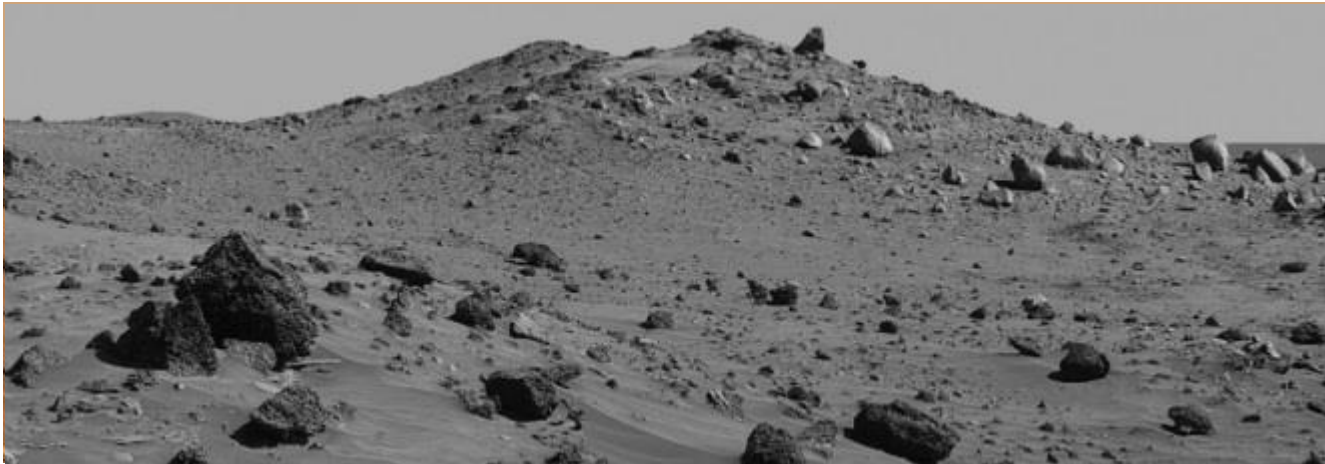


Fonte: canaltech.com.br

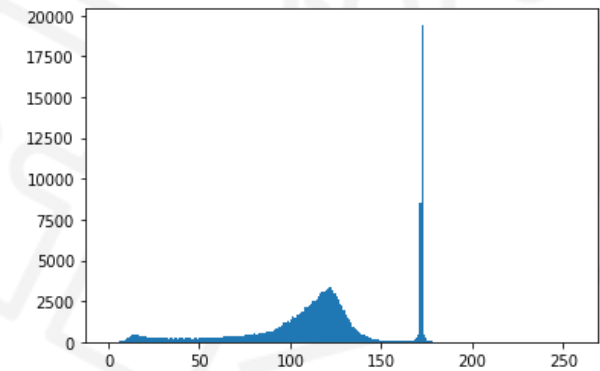
Pixel: menor unidade de
informação de uma imagem



Introdução



Fonte: canaltech.com.br

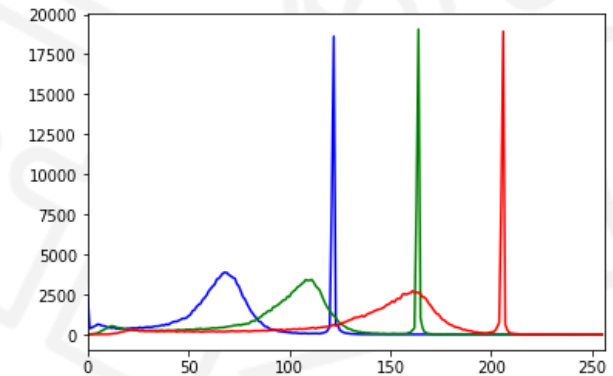


Histograma de cor da imagem (escala de cinza)

Introdução



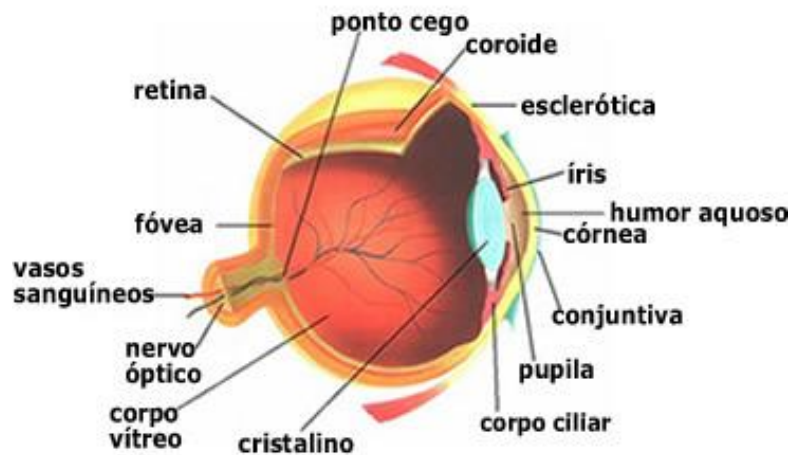
Fonte: canaltech.com.br



Histograma de cor da imagem (RGB)

Introdução

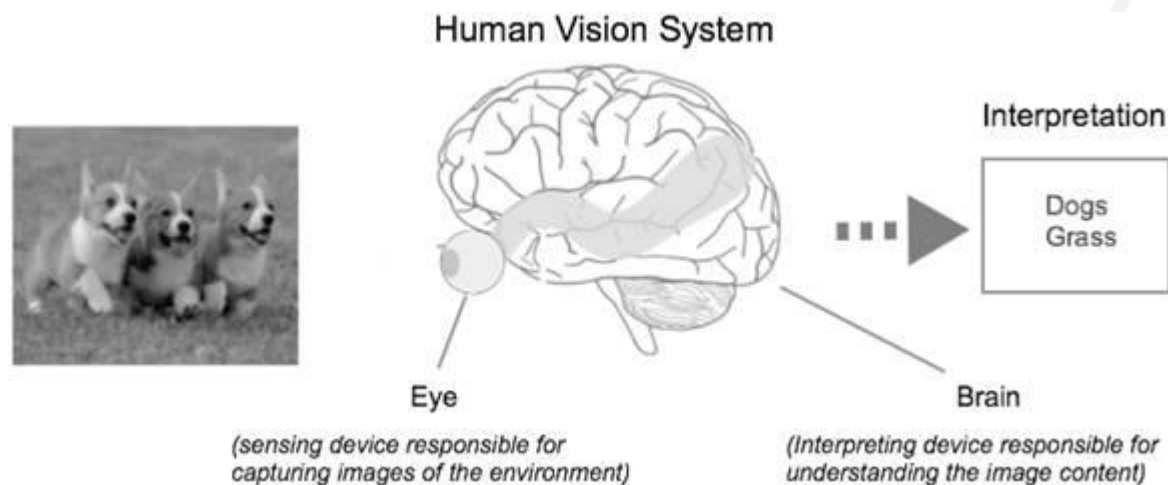
O métodos de visão computacional tentam imitar a capacidade do olho humano



Fonte da imagem: <https://brasilecola.uol.com.br/oscincosentidos/visao.htm>

Introdução

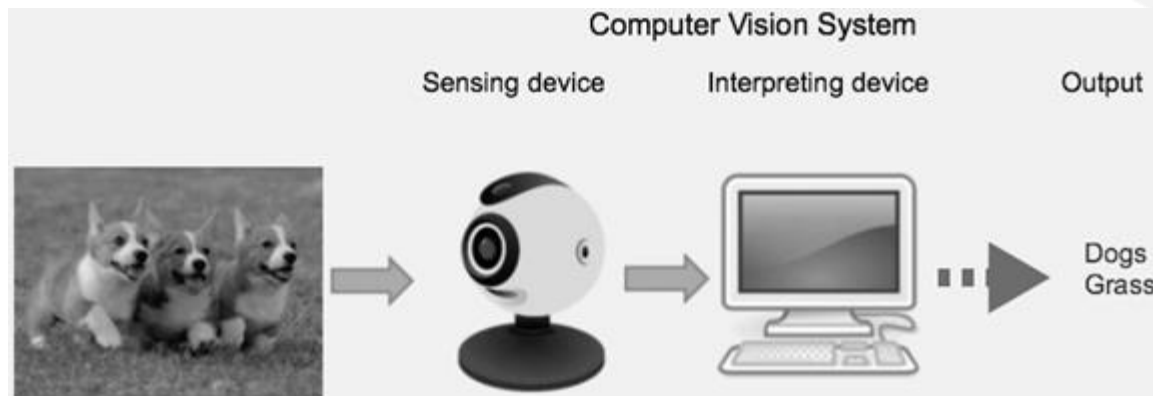
Sistema visual humano



Ref.: Mohamed Elgendy. **Deep Learning for Vision Systems**. 2020

Introdução

Sistema visual por inteligência artificial



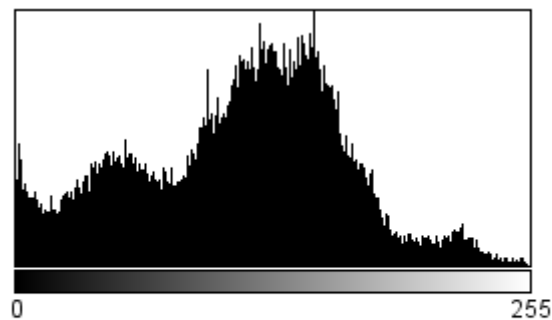
Ref.: Mohamed Elgendy. **Deep Learning for Vision Systems**. 2020

Introdução

Estes métodos procuram processar a imagem gerando um modelo descritivo (texto, números, etc..) que representa o conteúdo visual.



Fonte da imagem: wallpaperlist.com



Count: 50268 Min: 0
Mean: 110.761 Max: 255
StdDev: 52.700 Mode: 148 (509)

Histograma de cor da imagem
(escala de cinza)

Visão Computacional

Visão computacional e Processamento de Imagem



Imagem



Processamento de
Imagens



Imagem Processada



Imagem



Visão Computacional

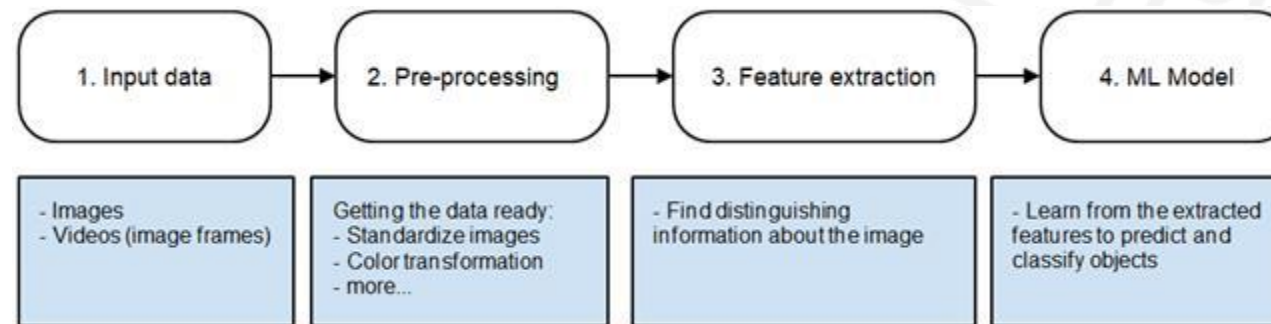


1.2133 1.231 0.123 ...

Modelo descritivo

Pipeline de visão computacional

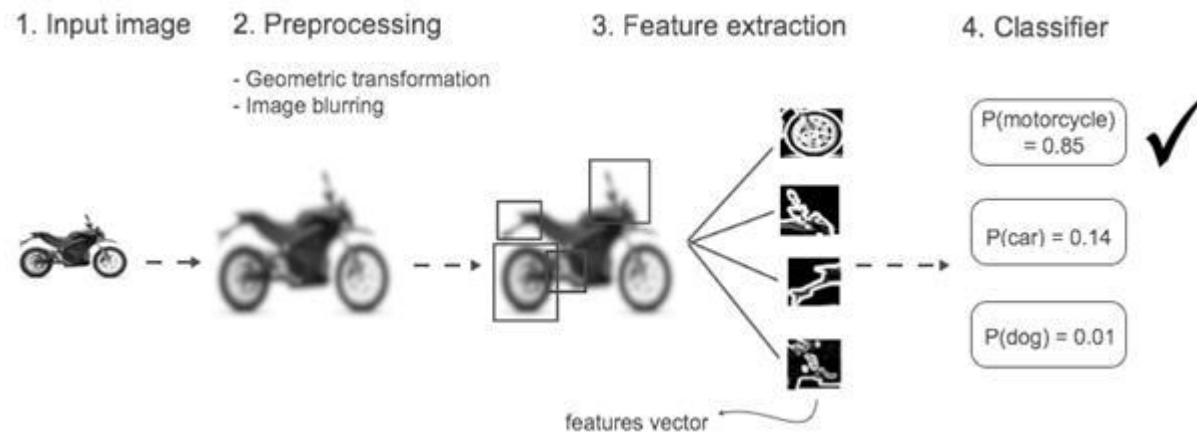
Sistema visual por inteligência artificial



Ref.: Mohamed Elgendy. **Deep Learning for Vision Systems**. 2020

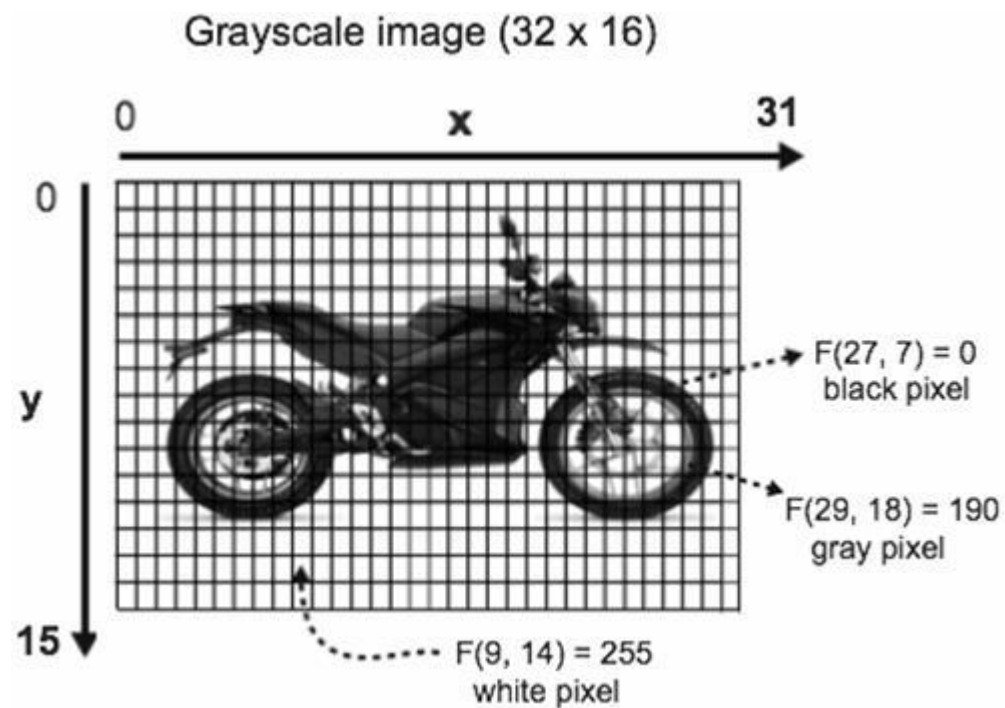
Pipeline de visão computacional

Exemplo de classificação de imagens



Ref.: Mohamed Elgendy. **Deep Learning for Vision Systems**. 2020

Imagem como uma função



Ref.: Mohamed Elgendy. **Deep Learning for Vision Systems**. 2020

Imagem como uma função

Application

Diminuir o brilho

Aumentar o brilho

Move um objeto 150 para baixo

Remover a tonalidade cinza da imagem e transformá-la em preto e branco

Transformation

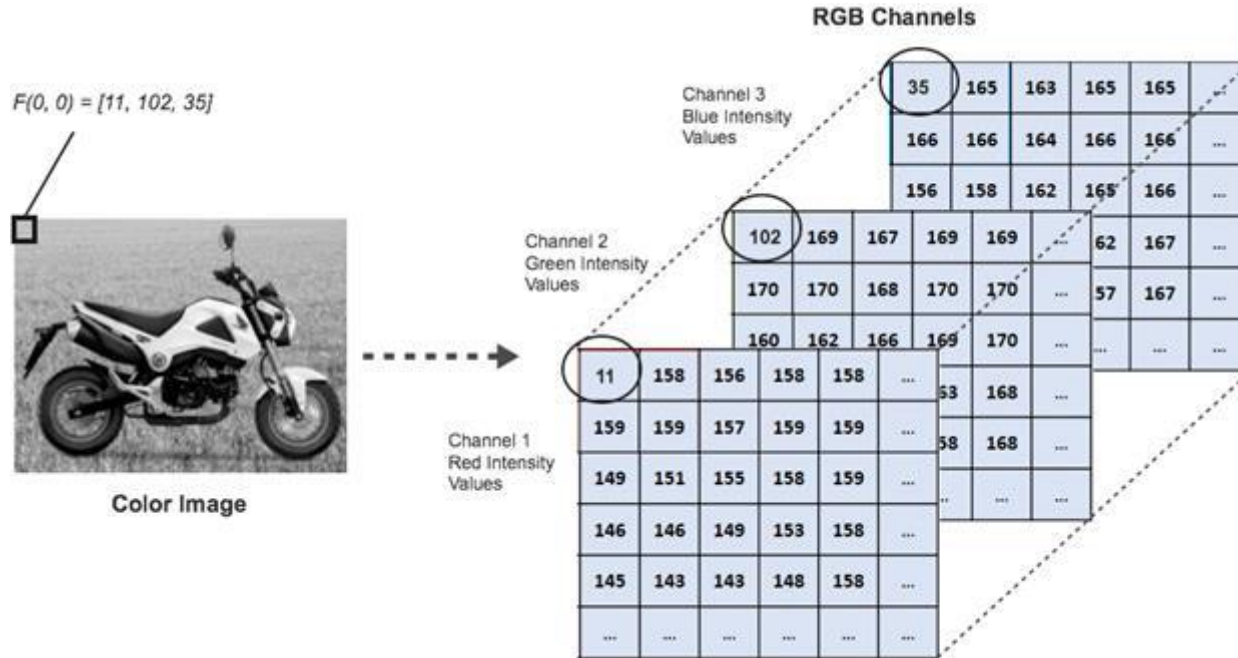
$$G(x, y) = 0.5 * F(x, y)$$

$$G(x, y) = 2 * F(x, y)$$

$$G(x, y) = F(x, y + 150)$$

$$G(x, y) = \{ 0 \text{ if } F(x, y) < 130, 255 \text{ otherwise } \}$$

Como um computador vê uma imagem



Red Green Blue

11 + 102 + 35 = Forest Green
(11, 102, 35)

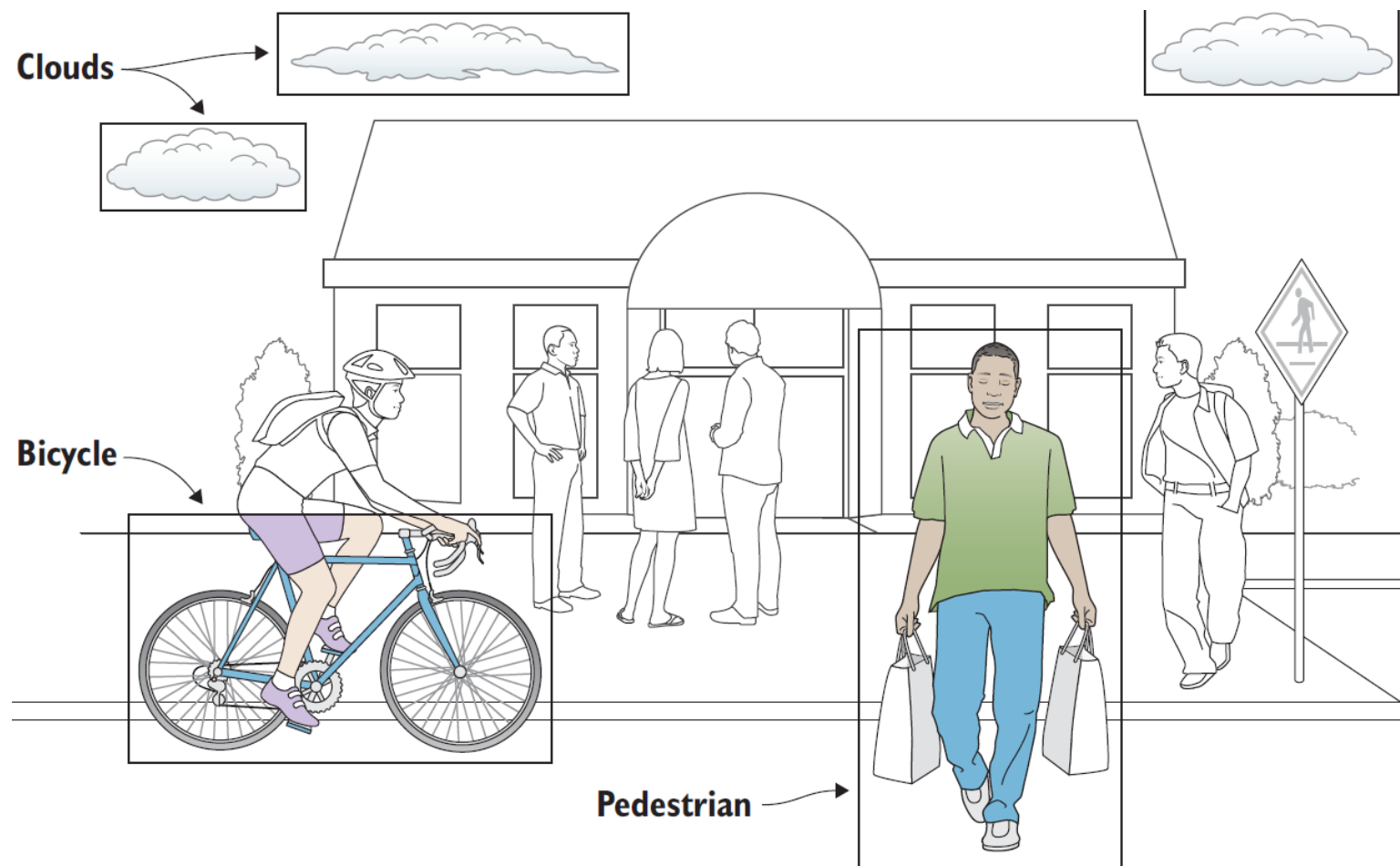
Ref.: Mohamed Elgendy. **Deep Learning for Vision Systems**. 2020

Pré-processamento de imagens

- Conversão para escala de cinza: em muitos objetos, a cor não é necessária (a quantidade de informação a mais das cores adiciona uma complexidade desnecessárias em alguns cenários, como detecção de objetos)

Ref.: Mohamed Elgendy. **Deep Learning for Vision Systems**. 2020

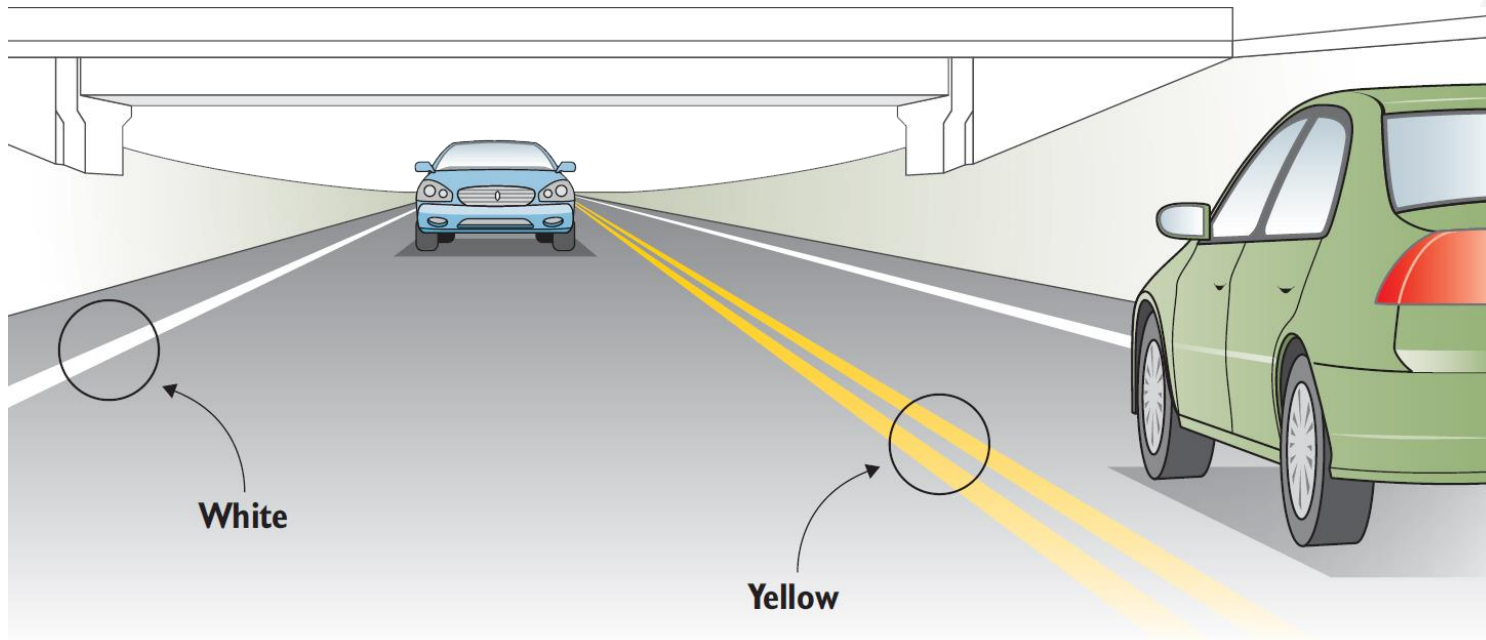
Conversão para escala de cinza



A forma do objeto
pode ser mais
relevante

Pré-processamento de imagens

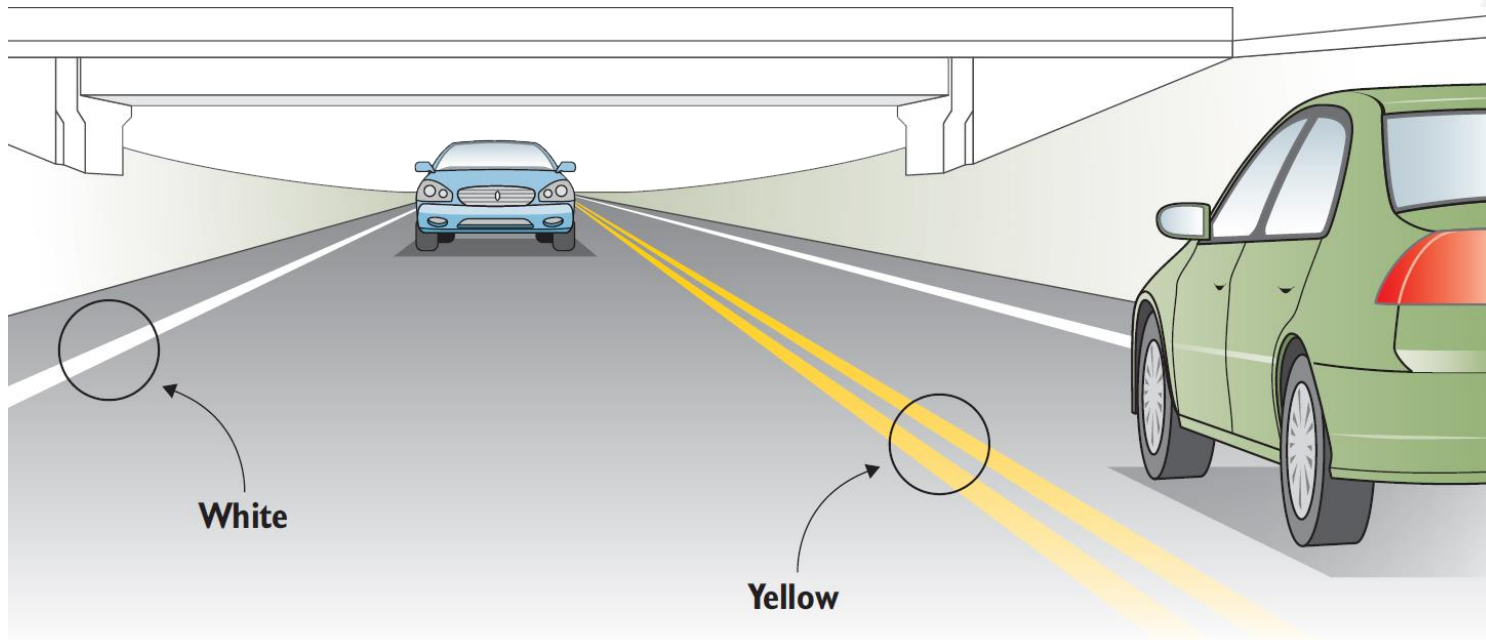
- Mas quando a cor é importante?



Ref.: Mohamed Elgendy. **Deep Learning for Vision Systems**. 2020

Pré-processamento de imagens

- Mas quando a cor é importante?



Ref.: Mohamed Elgendy. **Deep Learning for Vision Systems**. 2020

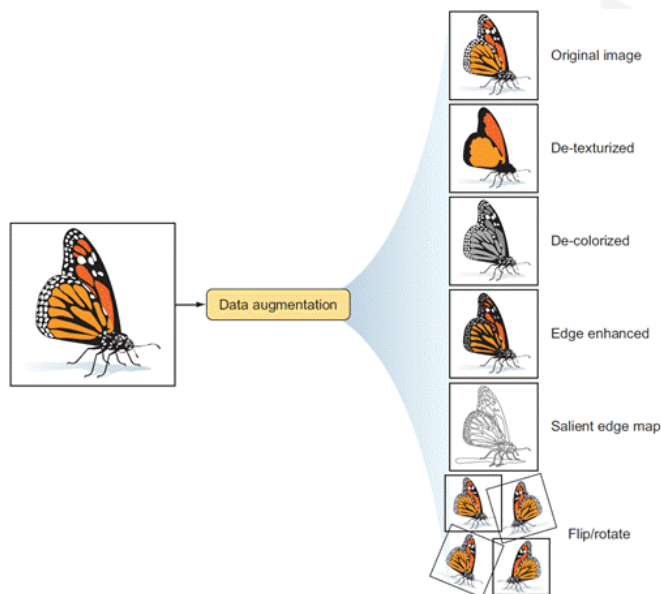
Pré-processamento de imagens

- Padronizar imagens: uma restrição importante que existe em alguns algoritmos de aprendizado de máquina é a necessidade de redimensionar as imagens no seu conjunto de dados para uma dimensão unificada. Isso implica que nossas imagens devem ser pré-processadas e dimensionadas para ter larguras e alturas idênticas antes de serem alimentadas pelo algoritmo de aprendizado.

Ref.: Mohamed Elgendy. **Deep Learning for Vision Systems**. 2020

Pré-processamento de imagens

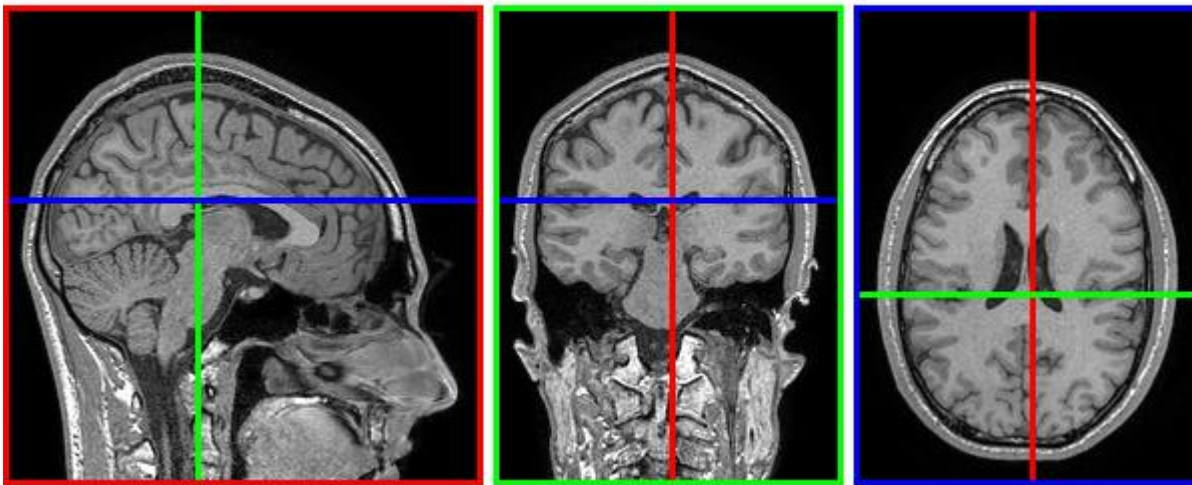
- Data augmentation: aumento do conjunto de dados com versões “perturbadas” das imagens existentes.



Ref.: Mohamed Elgendy. **Deep Learning for Vision Systems**. 2020

Imagens médicas

Imagens médicas



Realce (melhoramento) de imagens para auxiliar médicos no diagnóstico

Fonte da imagem: <https://www.aitrends.com/healthcare/machine-learning-in-medical-imaging-and-analysis/>

Imagens médicas

Hand X-Ray Search

http://zeus.robots.ox.ac.uk/hand_retrieval/query?id=34

Imagens médicas

Medical Image Search Engine

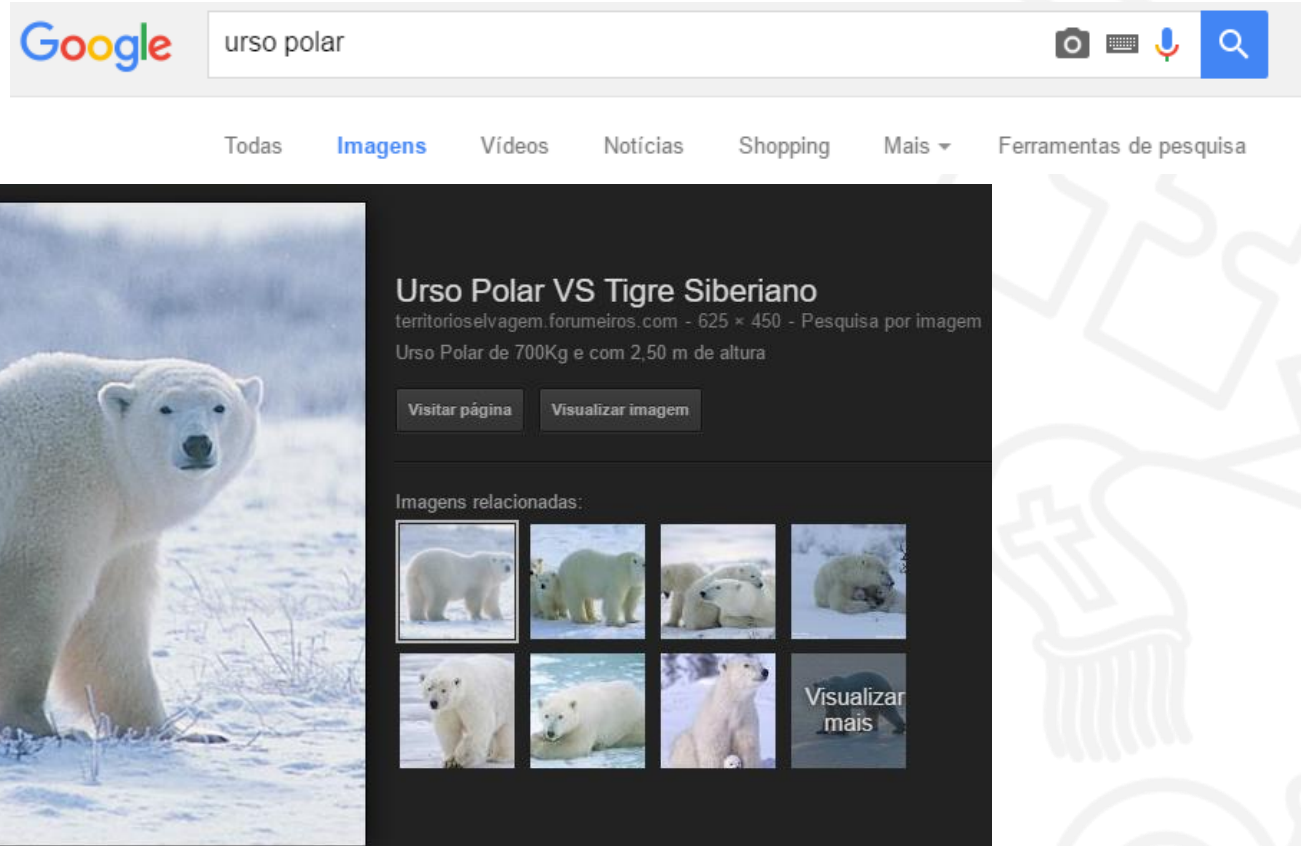
http://zeus.robots.ox.ac.uk/adni_retrieval/query?id=ad_1004

Recuperação de informação visual

Recuperação de informação visual (Web)

- Como é feito hoje?
- Indexação manual:
 - definição manual de palavras-chave que descrevem a imagem.
- Indexação automática:
 - utilização de: metadados semânticos associados à imagem; nome do arquivo da imagem; título do documento onde a imagem se encontra; texto próximo à imagem; etc.

Recuperação de informação visual (Web)



Fonte da imagem: Google Images

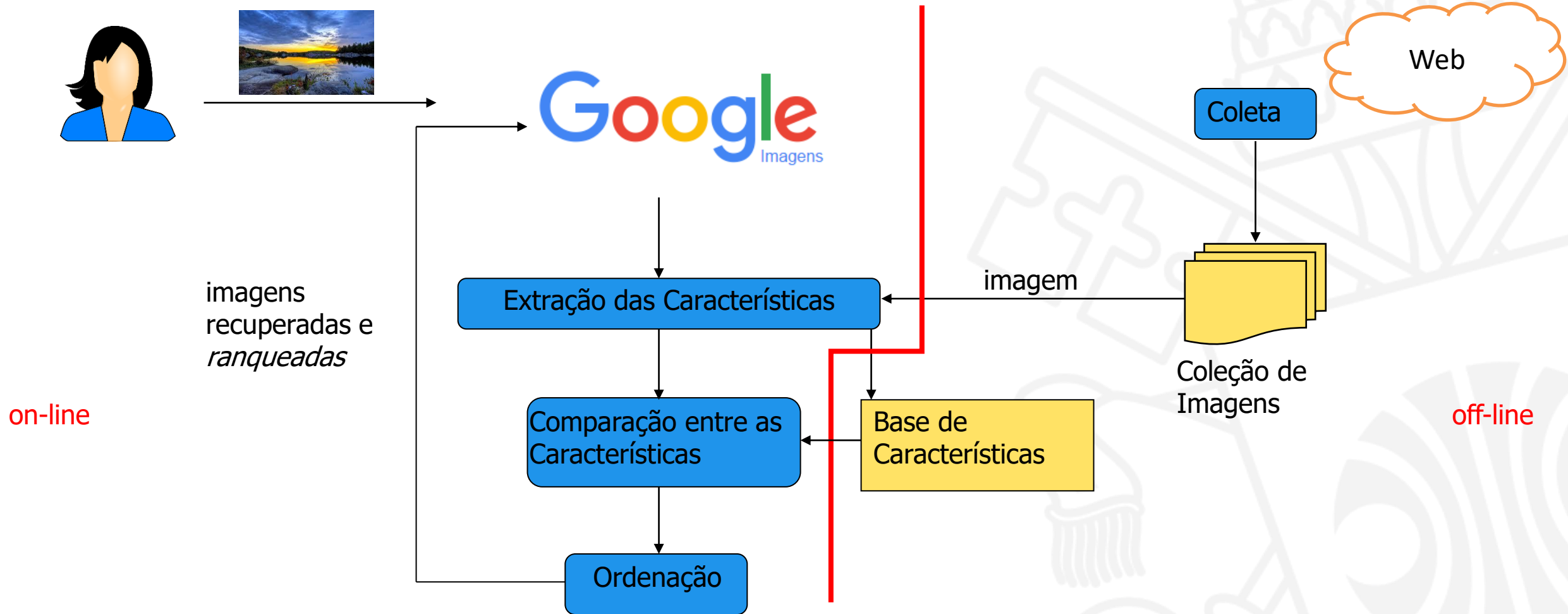
Recuperação de informação visual (Web)

- CBIR – Content-Based Image retrieval
- Representação e recuperação das imagens;
 - através de suas características primitivas, como:
 - cor;
 - forma dos objetos;
 - relação espacial entre as cores e objetos da imagem; etc.

Recuperação de informação visual (Web)

- Problema:
 - Recuperar imagens que casam (matching);
 - total ou apenas parcialmente com uma imagem dada na consulta.

Recuperação de informação visual (Web)



Recuperação de informação visual (Web)

- BBC News Search:

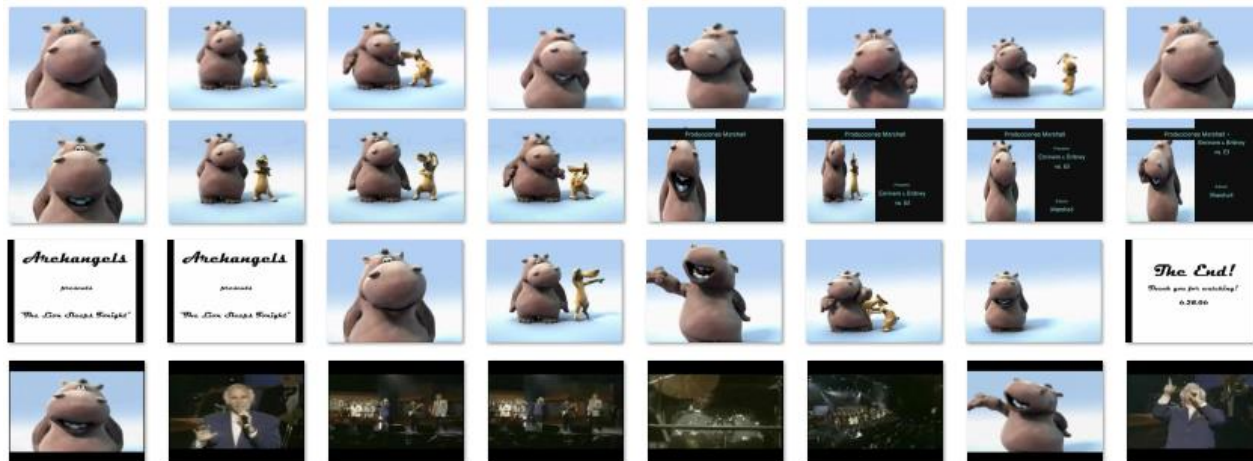
http://zeus.robots.ox.ac.uk/bbc_search/?engine=objects

Recuperação de informação visual (Web)

- Video Google Demo:

<https://www.robots.ox.ac.uk/~vgg/research/vgoogle/index.html>

Near-duplicate Video Detection



Detecção de duplicatas
similares de vídeos
(aplicações em detecção
de conteúdo com direitos
autorais)

Fonte da imagem: Near-Duplicate Video Detection Based on an Approximate Similarity Self-Join Strategy

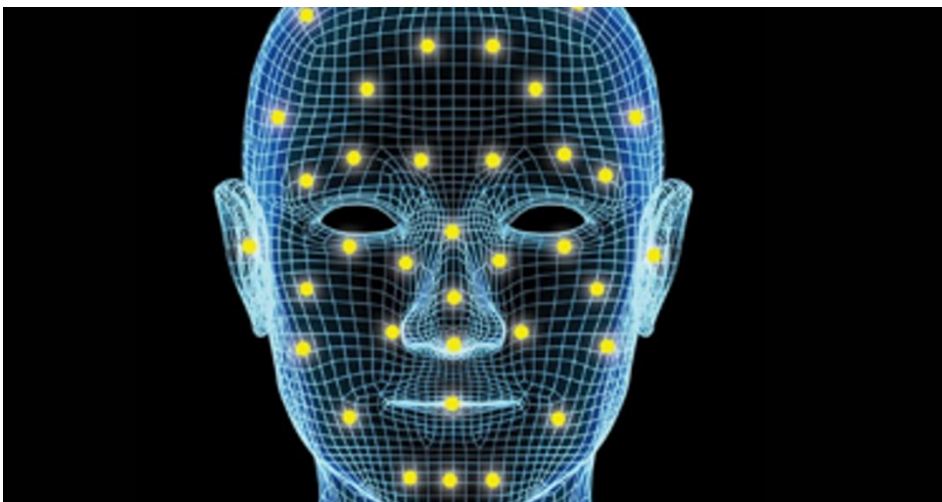
CBIR em Grandes datasets

Web-scale particular object search

<https://www.robots.ox.ac.uk/~vgg/research/oxbuildings/index.html>

Reconhecimento de faces e textos

Reconhecimento de faces



Segurança: smartphones (bloqueio de acesso). Locais públicos e localização de pessoas procuradas.

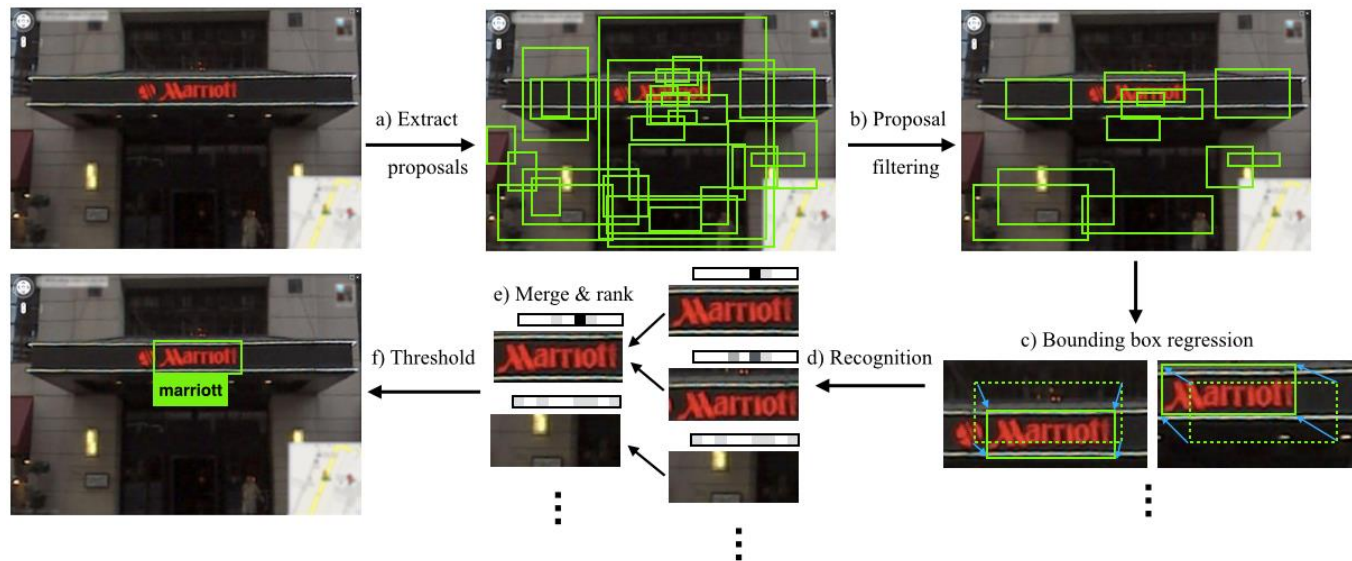
Fonte da imagem: <https://spectrum.ieee.org/transportation/sensors/how-3d-sensing-enables-mobile-face-recognition>

Reconhecimento de faces

Video Google Faces Demo:

<https://www.robots.ox.ac.uk/~vgg/research/fgoogle/index.html>

Descrição de texto em imagens



Extração de texto
em imagens e
vídeos

Fonte da imagem: <https://www.robots.ox.ac.uk/~vgg/research/text/index.html>

Descrição de texto em imagens

BCC News Text Search:

<http://zeus.robots.ox.ac.uk/textsearch/#/search/>

Demais aplicações em visão computacional

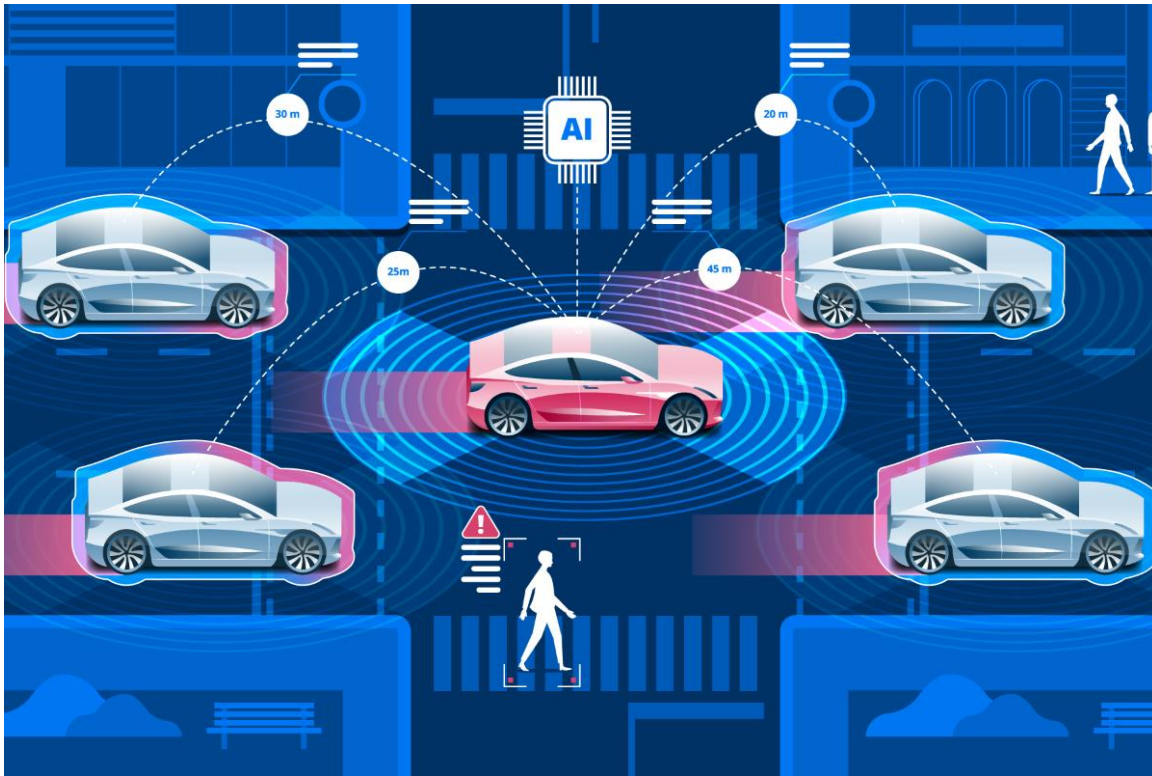
Imagens de satélites



Fonte da imagem: <https://venturebeat.com/2019/04/09/facebook-ai-researchers-want-to-map-earths-population-using-computer-vision-and-satellite-imagery/>

Mapeamento de população. Detecção de pragas em vegetação, etc..

Veículos autônomos



Fonte da imagem: <https://venturebeat.com/2019/04/09/facebook-ai-researchers-want-to-map-earths-population-using-computer-vision-and-satellite-imagery/>

Tomada de decisão da
AI de veículos
autônomos.

Reconhecimento de armas de fogo



Segurança: detecção de pessoas portando armas de fogo.

Fonte da imagem: <https://phys.org/news/2017-07-artificial-intelligence-based-gun-video.html>

Re-identificação de pessoas

Positive Image Pairs

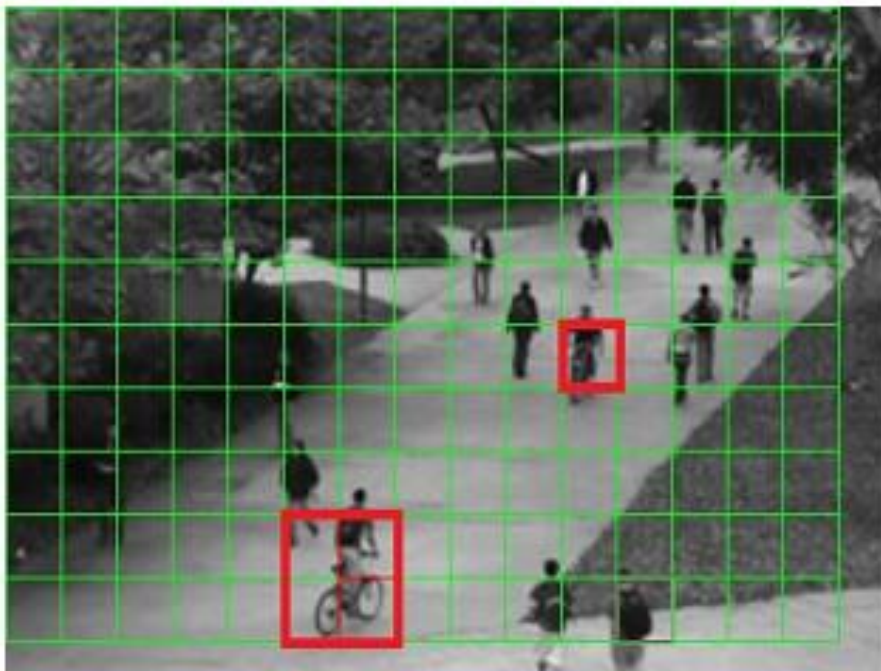


Negative Image Pairs



Fonte da imagem: <http://sites.bu.edu/data/reid/>

Anomalous Event Detection



Detecção de eventos que desviam dos eventos naturais para aquela região.

Fonte da imagem: <http://www.smartsenselab.dcc.ufmg.br/anomalous-event-detection/>

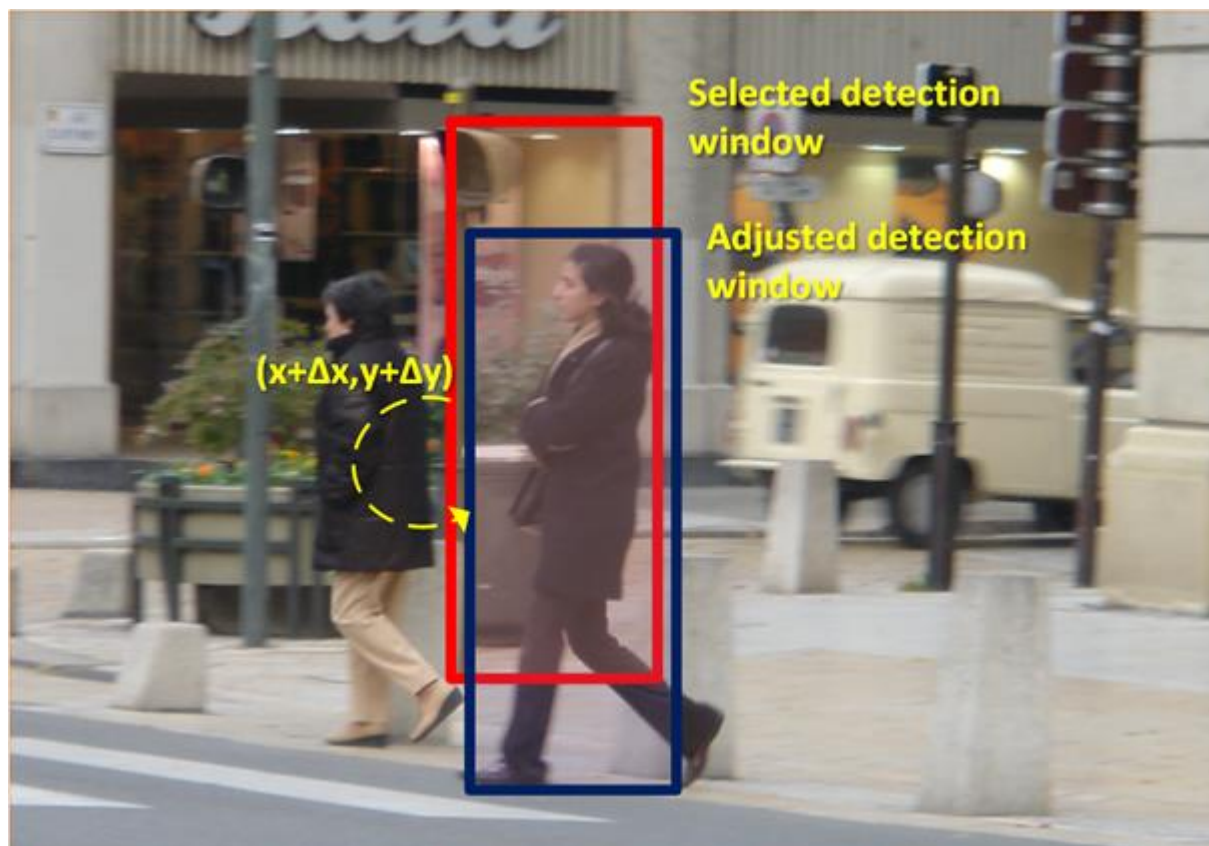
License Plate Recognition (LPR)



Detecção de placas de
veículos de
reconhecimento de
caracteres das placas

Fonte da imagem: <http://www.smartsenselab.dcc.ufmg.br/automatic-license-plate-recognition-alpr/>

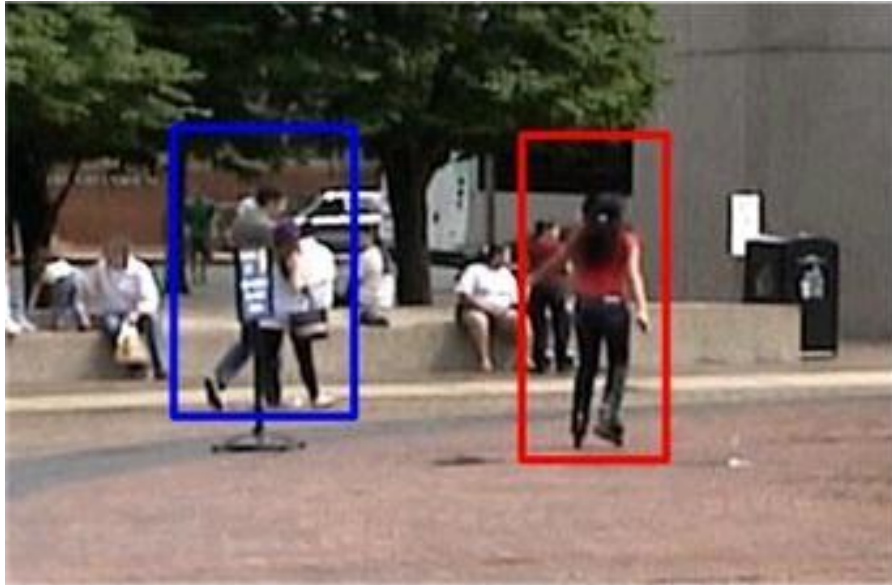
Detecção de pedestre



Fonte da imagem: <http://www.smartsenselab.dcc.ufmg.br/pedestrian-detection/>

Detecção de pedestre:
tracking de pessoas
em vias públicas

Detecção de ações em vídeos



walk

skate

Detecção de ações
realizados por
humanos em vídeo

Fonte da imagem: <http://www.bu.edu/ids/research-projects/action-recognition/>

Demonstrações práticas

YouTube

*“people around the world are now watching **a billion hours of YouTube’s incredible content every single day!**”*

- YouTube Official Blog, February 27, 2017

Fonte: <https://youtube.googleblog.com/2017/02/you-know-whats-cool-billion-hours.html>

Em **2012** esta marca era de **86.400 horas por dias.**

Fonte: <https://youtube.googleblog.com/2012/01/holy-nyans-60-hours-per-minute-and-4.html>

Demonstrações

Microsoft Caption Bot

<https://www.captionbot.ai/>

Veja também Microsoft Computer Vision API

<https://www.microsoft.com/cognitive-services/en-us/computer-vision-api>

<https://azure.microsoft.com/pt-br/services/cognitive-services/face/#demo>

Microsoft Emotion API

<https://www.microsoft.com/cognitive-services/en-us/emotion-api#detection>

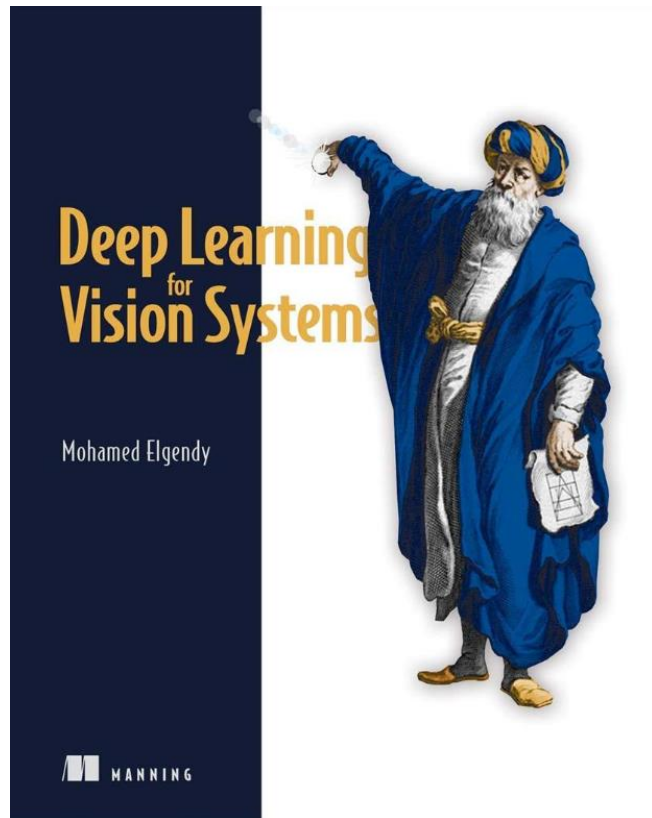
Use Deep Learning to Automatically Colorize Black and White Photos

<http://demos.algorithmia.com/colorize-photos/>

Web Demos

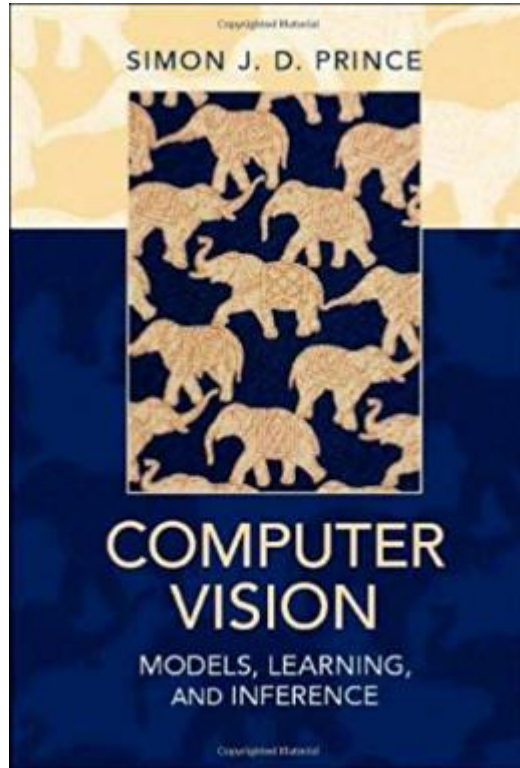
<https://www.robots.ox.ac.uk/~vgg/demo/>

Principais Referências



Mohamed Elgandy. **Deep Learning for Vision Systems**. 2020
(estimated)
ISBN 9781617296192 410 pages

Principais Referências



Simon J. D. Prince. **Computer Vision: Models, Learning, and Inference 1st Edition**