

Aprendizagem Profunda para Visão por Computador

(Deep Learning for Computer Vision)

Apresentação da UC

Tomás Brandão – DCTI – ISTA

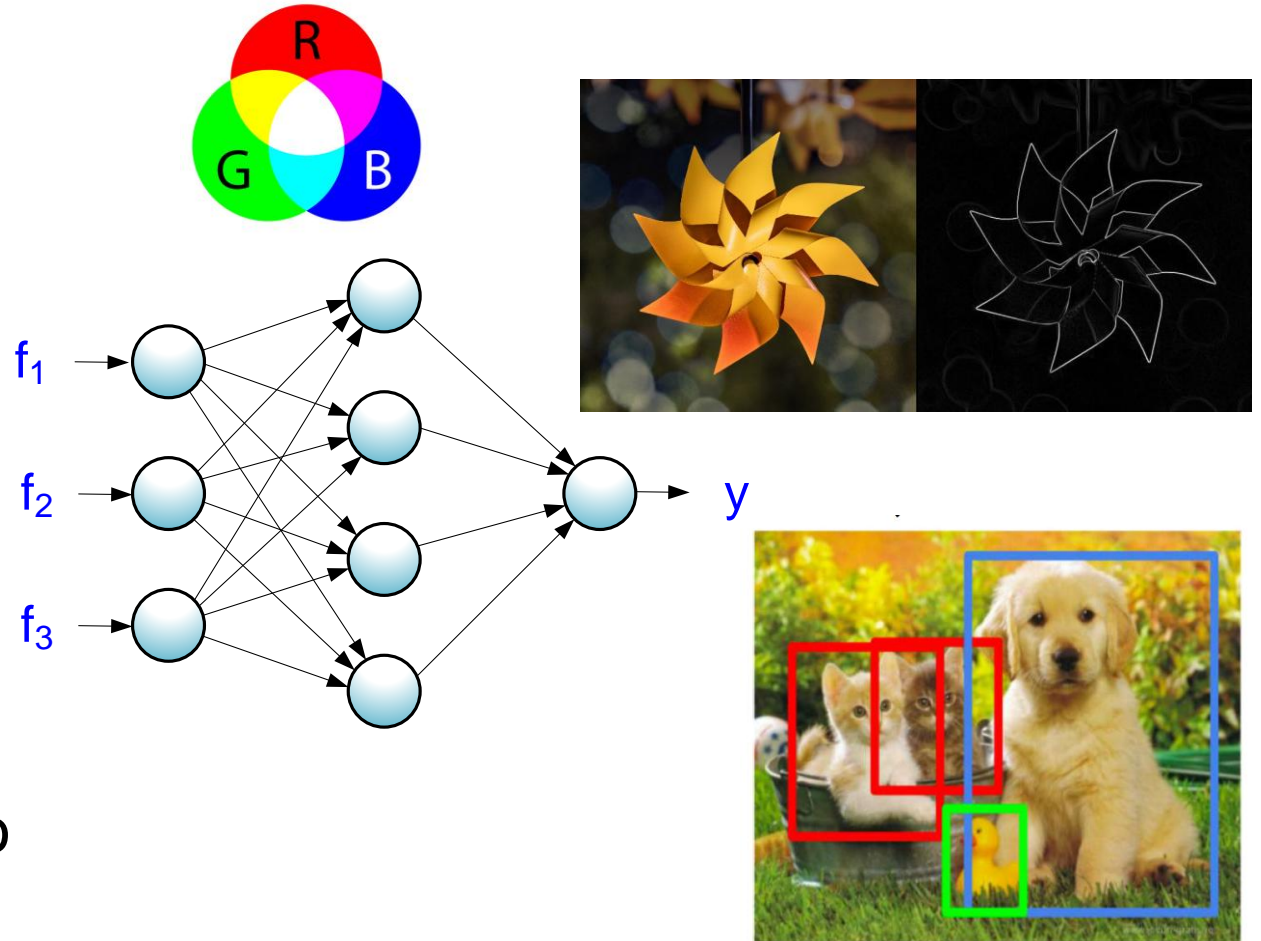
Sumário

- Funcionamento da UC
 - Programa
 - Funcionamento das aulas
 - Avaliação
 - Material
- Introdução à visão por computador e análise de imagens

Funcionamento da UC

Programa

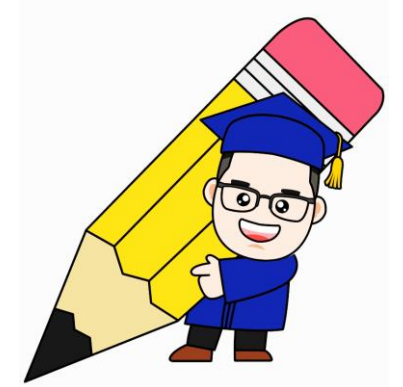
- Representação de imagens
- Operações com imagens
- Redes neuronais “clássicas”
- Redes neuronais convolucionais
- Transferência de conhecimento
- Detecção de objetos e segmentação
- Outras aplicações: geração de conteúdos, super-resolução, ...



Avaliação

■ Avaliação ao longo do semestre – não há avaliação por exame!

- Modalidade A (implica presença em pelo menos 60% das aulas)
 - Exercícios (30%): participação em aula (10%) e 2 trabalhos de grupo (10%+10%)
 - Teste (30%)
 - Projeto (40%) – em grupo
- Modalidade B (sem requisitos de assiduidade)
 - Teste (45%)
 - Projeto (55%) – individual ou em grupo
- Obs.
 - O teste realiza-se após o final do período de aulas (29/abril?)
 - O projeto inclui relatório e discussão oral; e tem nota mínima de 10 val.
 - Prazo de entrega do projeto e discussões em datas a combinar (há flexibilidade)



Grupos de 3 ou 4 elementos
Max 10 grupos/turno (será possível?)
Escolha de grupos através do Moodle

Funcionamento das aulas

- Aulas teórico-práticas

- 1ª parte – Introdução aos temas
 - Exposição e exemplos
- 2ª parte – Exercícios (e quizz?)
 - Podem-se resolver em grupo, mas a avaliação da participação é individual
 - É fundamental trazerem portáteis (pelo menos um por grupo)



Material

- Python 3 – versão **3.9** a **3.12**
- Principais pacotes que irão ser utilizados:
 - opencv-python (e numpy)
 - matplotlib
 - pandas
 - tensorflow
 - ultralytics (e pytorch)
- Um IDE para python (à escolha)
 - PyCharm (usado pelo professor nas aulas)
 - Notebooks (Jupyter, Colab)
 - VS Code
 - ...



TPC – Próxima aula

- Instalar/configurar um ambiente de desenvolvimento de Python com o OpenCV
- Criar um pequeno script para ler e mostrar uma imagem, usando o OpenCv.

```
# importar a biblioteca do opencv
import cv2

# ler uma imagem de teste (ajustar nome/caminho)
# garantir que a imagem está no sitio certo
img = cv2.imread(r'd:\imagens\test.jpg')

# mostrar a imagem numa janela
cv2.imshow('Imagem de Teste',img)

# colocar em espera até que se carregue numa tecla
cv2.waitKey(0)

# destruir a janela criada
cv2.destroyAllWindows()
```


Introdução à visão por computador

O que é? Para que serve?

Visão por Computador – o que é?

Área multidisciplinar onde o principal objetivo é conseguir fazer com que um computador consiga extrair informação de imagens e vídeos de forma a compreender o seu conteúdo para realizar tarefas úteis num dado contexto



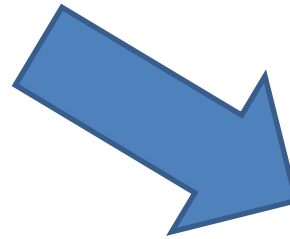
"um gato em cima de um skate no chão da sala"

Do ponto de vista da engenharia, a visão por computador está associada à automatização de funções que muitas vezes os seres humanos também conseguem realizar com base no seu sistema visual

Os computadores “vêm” números...



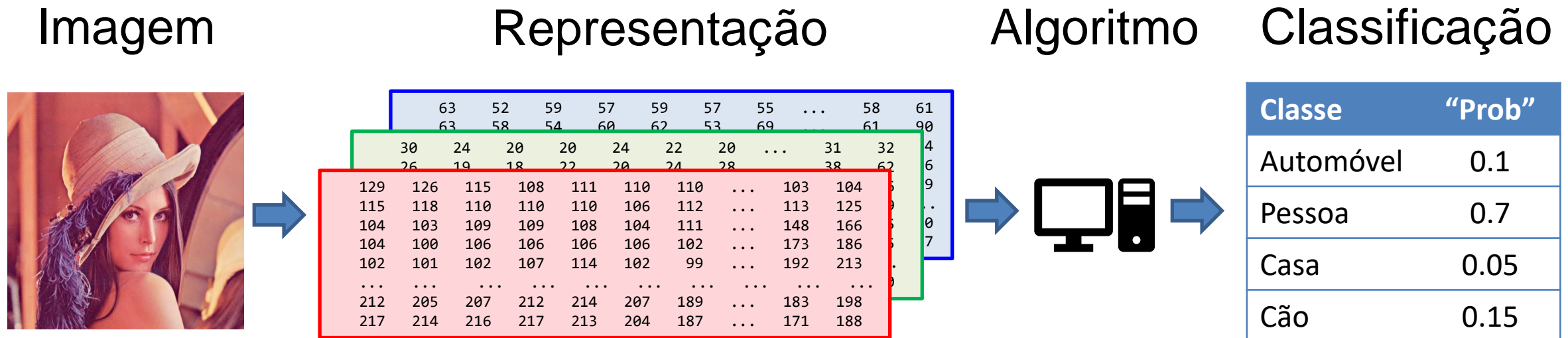
Para nós é fácil de dizer que esta é a imagem de uma pessoa...



Mas e o computador? O que ele “vê” relativamente à imagem tipicamente são três matrizes com valores 0-255...

										63	61
										63	90
										30	114
										26	166
										129	179
										115	...
										104	110
										104	117
										102	...
										...	100
										212	...
										217	...

Exemplo – classificação de imagens



- **Classificação** – o resultado será uma classe (ou categoria)
- **Regressão** – o resultado será a previsão de um valor (contínuo)

Características de imagem (*features*)



Nariz, olhos, boca



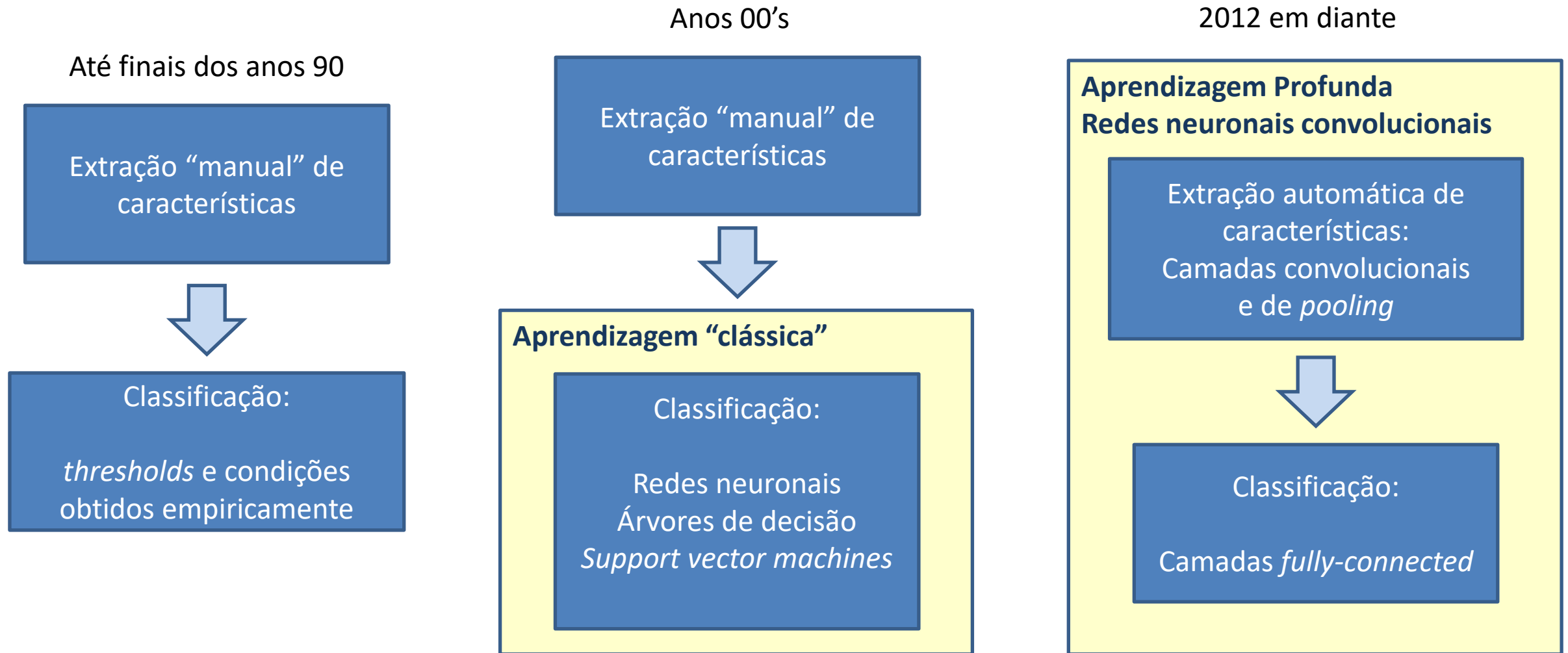
Roda, faróis, para-brisas, grelha



Porta, janelas, paredes, telhado

- As características relevantes variam muito consoante o domínio do problema
- O ser humano tende a classificar usando características de alto nível
- As características de alto nível não são fáceis de obter através de cálculos...

Classificação de imagens: *timeline*



Extração “manual” de *features*

- Variam muito consoante o contexto do problema de classificação em causa
- E há questões que não são fáceis de ultrapassar...

Variações do ponto de vista



Diferenças de escala



Oclusões



Condições de iluminação



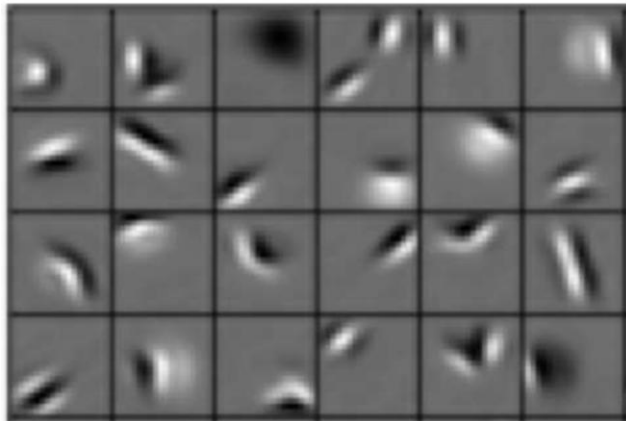
Variações “intra-classe”



Aprendizagem automática de *features*

Abordagem atual amplamente utilizada: **redes neurais convolucionais**
(será um dos temas mais importantes em APVC)

O computador aprende uma hierarquia de *features* com base em exemplos

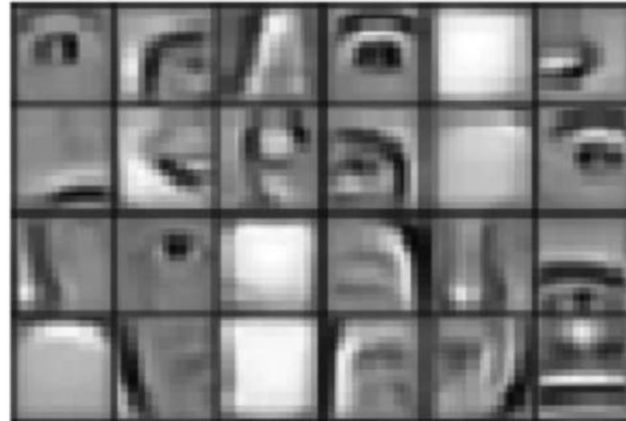


Baixo nível

Manchas

Linhas

Cores



Nível médio

Olhos

Bocas

Narizes



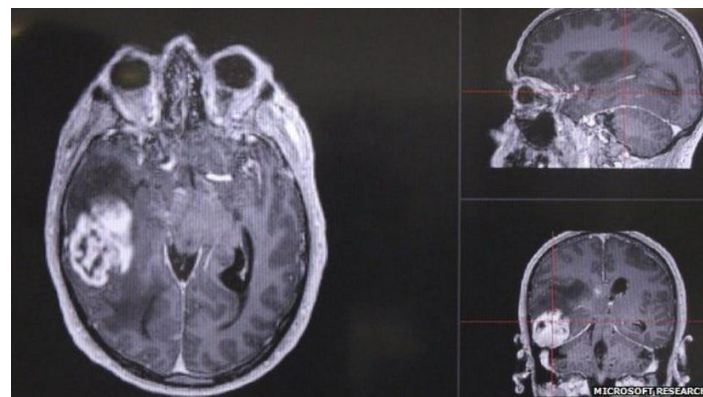
Alto nível

Estrutura facial

Áreas de aplicação



Videovigilância



Medicina



Condução autónoma



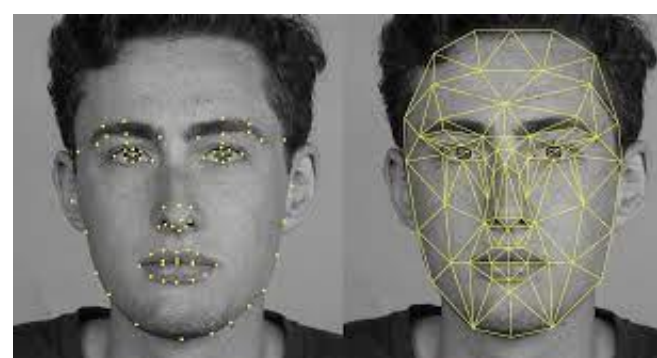
Acesso à informação



Robótica e indústria



Agricultura



Reconhecimento facial

E outras mais...

Tarefas mais habituais

Segmentação
semântica



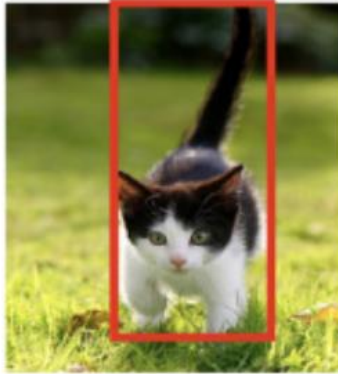
Pixels

Classificação
de imagem

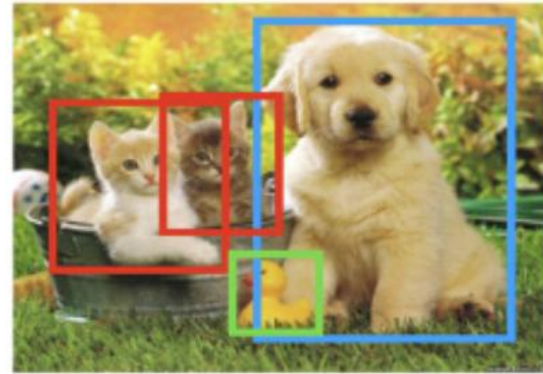


Objeto único

Localização e
identificação



Localização e identificação
de múltiplos objetos

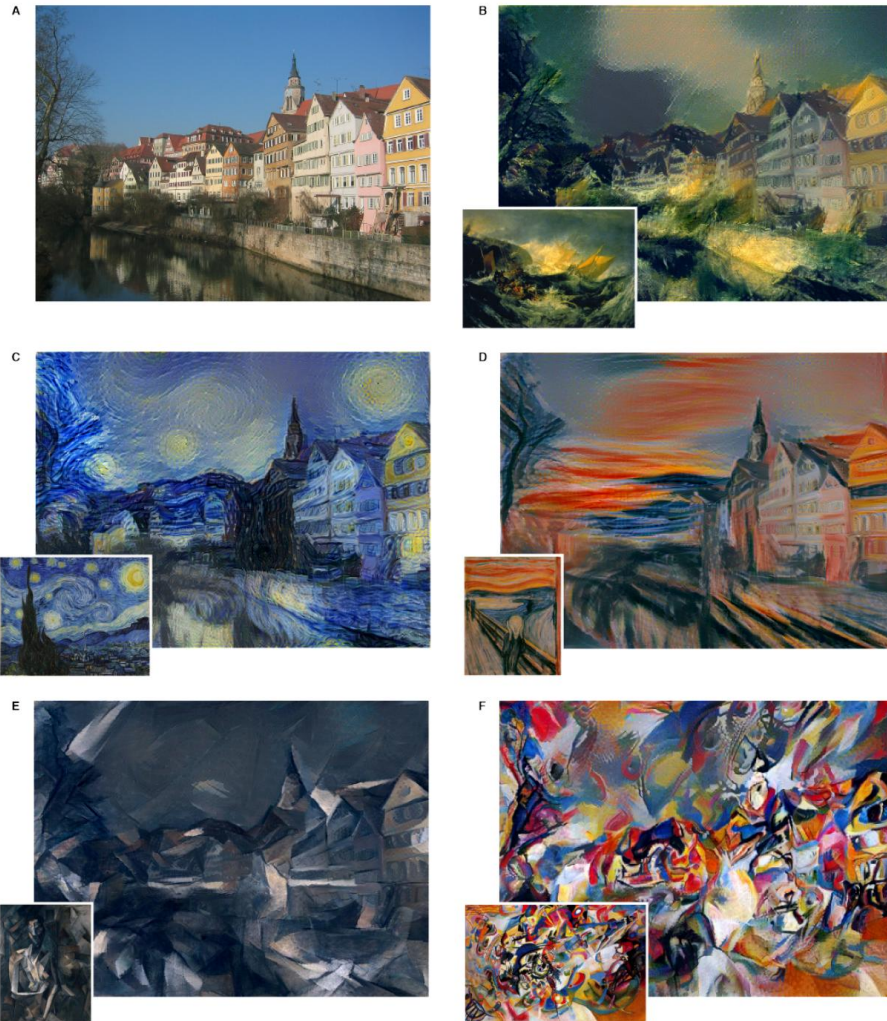


Múltiplos objetos

Segmentação de
de múltiplos objetos

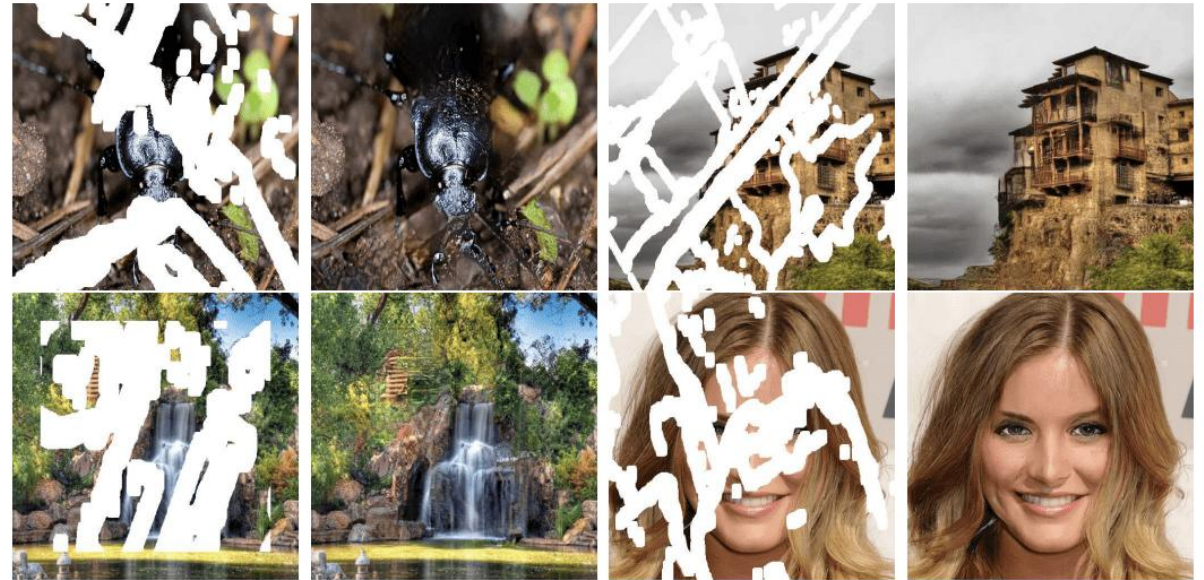


Outras tarefas mais específicas



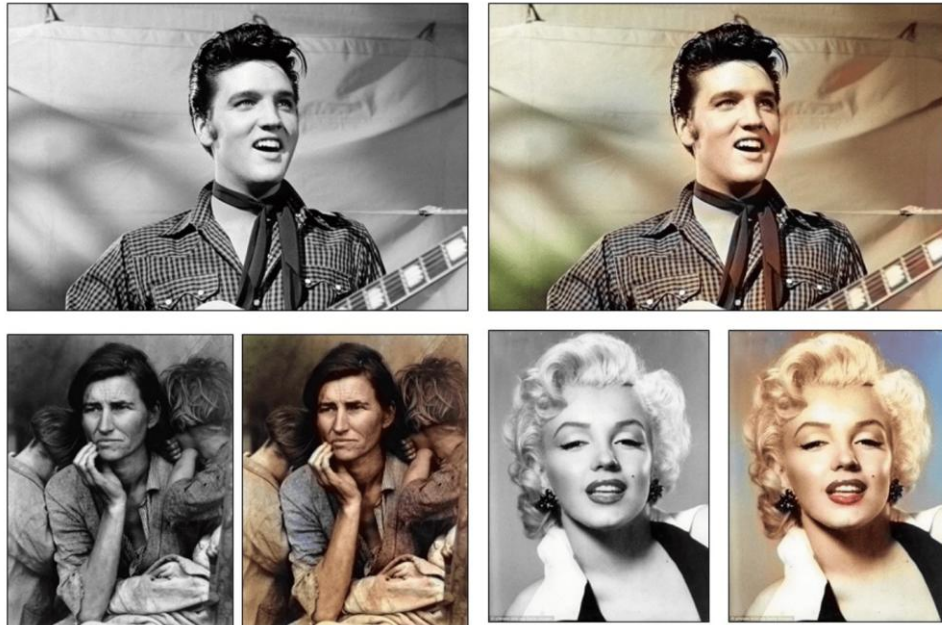
Transferência
de estilo

Restauração
(*Inpainting*)



Outras tarefas mais específicas

Coloração automática



Geração de conteúdos



Alguns exemplos de trabalhos desenvolvidos no ISCTE

Identificação de marca, modelo e ano de veículos

João Cruz,

Deep learning for large-scale fine-grained recognition of cars

Dissertação MEI, orientação de Luís Nunes e Tomás Brandão, 2018

- **Objetivo**
Facilitar a introdução de dados de veículos em stands online, identificando automaticamente a marca, modelo e ano de fabrico de automóveis com base nas imagens submetidas
- Sistema baseado em *transfer learning*
- Principais resultados
 - Construção de um dataset (OTO-970) com cerca de 1.2 milhões de imagens e 790 classes
 - Taxa de acertos global na ordem dos 92.7%
 - Top-3 e Top-5 na ordem dos 97,8% e 98,4%

opel_corsa_2006 alfa_romeo_147_na vw_passat_2000

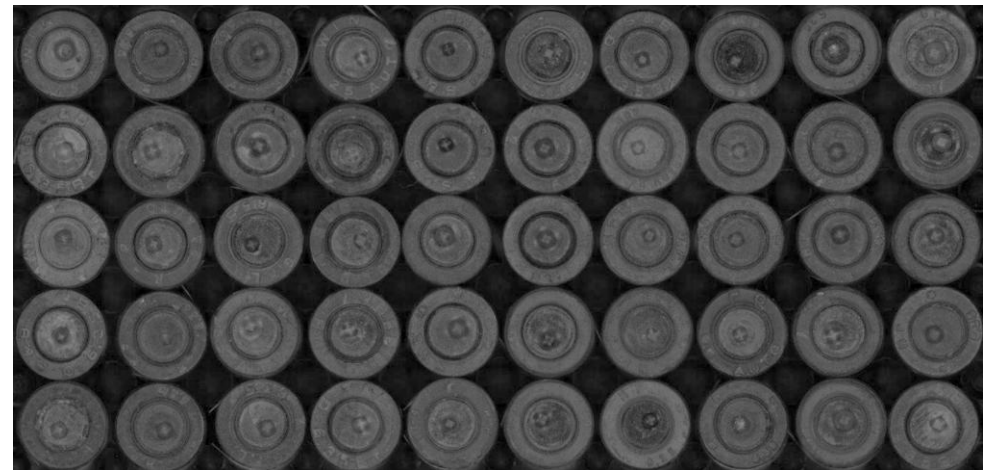
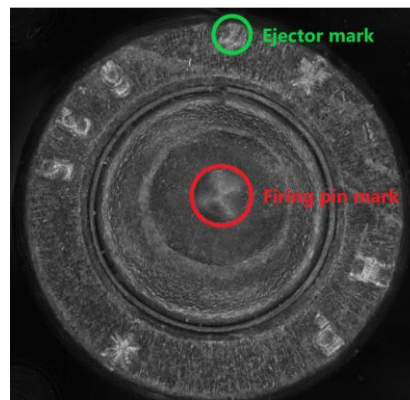


peugeot_partner_2008 audi_a4_1995 seat_leon_2012

Identificação de modelo de arma

S. Valentim, T. Fonseca, T. Brandão, J. Ferreira, R. Ribeiro e S. Nae,
Gun model classification based on fired cartridge case head images with siamese networks,
Lecture Notes in Networks and Systems – Intelligent Systems Design and Applications, 2021

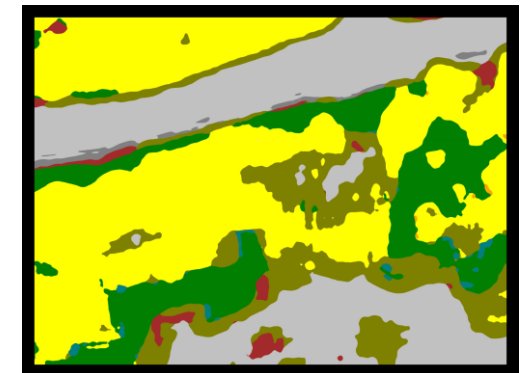
- Objetivo – Identificar automaticamente o modelo de arma que disparou um projétil com base em imagens dos cartuchos ejetados (recolhidos na cena do crime)
- Solução baseada em redes CNN siamesas
- Acertos Top-1/Top-2: 100%/100% – dataset NBTRD (USA); 57%/81% – dataset experimental PJ



Deteção da espécie invasora *Acacia longifolia*

C. Gonçalves, P. Santana, T. Brandão, M. Guedes,
Automatic detection of Acacia longifolia invasive species based on UAV-acquired aerial imagery,
Information Processing in Agriculture, 2022

- Objetivo
Identificar a presença da espécie invasora *Acacia longifolia*, a partir de imagens adquiridas por um drone
- Sistema baseado em CNN que pode classificar pixel-a-pixel ou blocos 100x100
- Resultados:
 - Taxa de acertos de 98.7% (CNN binary)
 - Taxa de acertos de 98.5% (CNN multiclass)



Deteção e classificação de graffitis

J. Fogaça, T. Brandão e J. Ferreira,

Deep learning-based graffiti detection: a study using images from the streets of Lisbon,
Applied Sciences, 2023

- **Objetivo**
Localizar graffitis em imagens e classificar segundo *street art* / *graffiti ilegal*
- Solução baseada em deteção de objetos e classificação com redes neuronais convolucionais
- **Resultados:**
 - Deteção: IoU: 0.70
 - Accuracy (street art/ilegal/sem grafiti): 0.81



True: street art
Predicted: street art



True: illegal graffiti
Predicted: illegal graffiti



True: without graffiti
Predicted: without graffiti

Deteção de pessoas e análise de trajetórias

S. Correia, D. Mendes, P. Jorge, T. Brandão, P. Arriaga e L. Nunes, *Occlusion-aware pedestrian detection and tracking*, International Conference on Systems, Signals and Image Processing, North Macedonia, 2023

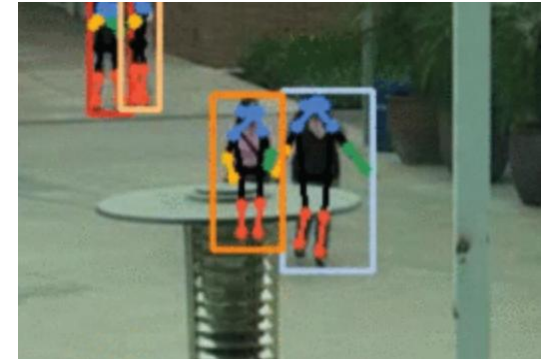
D. Mendes, S. Correia, P. Jorge, T. Brandão, P. Arriaga e L. Nunes, *Multi-camera person re-identification based on trajectory data*, Applied Sciences, 2023

■ Objetivos

- Deteção de pessoas/grupos
- Reconhecimento de ações
- Análise de trajetórias

■ Técnicas usadas

- Deteção de objetos
- Processamento de sinais/imagem
- Extração de esqueletos usando técnicas de deep-learning



Deteção automática de incêndios

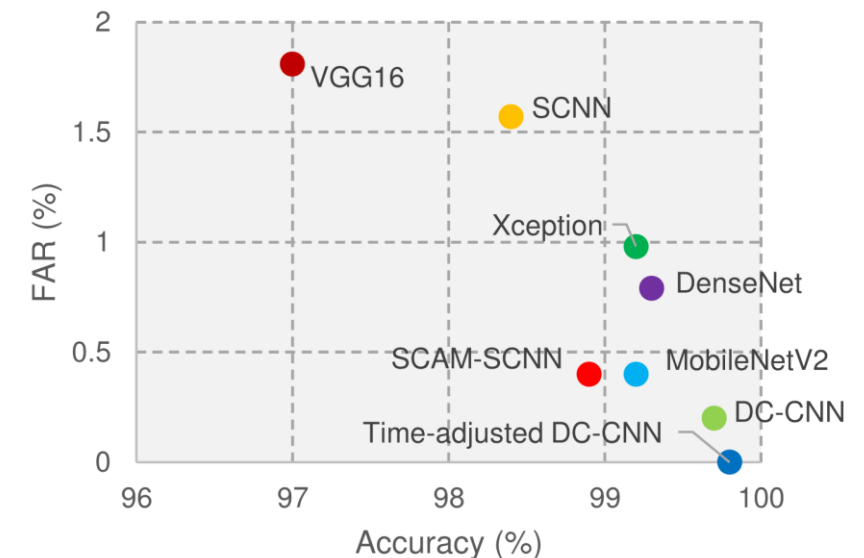
A. M. Gonçalves, T. Brandão, J. C. Ferreira,
Wildfire Detection With Deep Learning — A Case Study for the CICLOPE Project,
IEEE Access, 2024

- Goal

- Melhorar a taxa de falsos alarmes de um sistema de deteção automática de incêndios florestais (CICLOPE)

- Técnicas usadas

- *Dual-channel* CNNs
- Modelos de atenção



Diagnóstico de doenças cardiovasculares

L. B. Elvas, S. Gomes, J. C. Ferreira., L. Rosário, T. Brandão,
Deep learning for automatic calcium detection in echocardiography,
BioData Mining, 2024

■ Objetivos

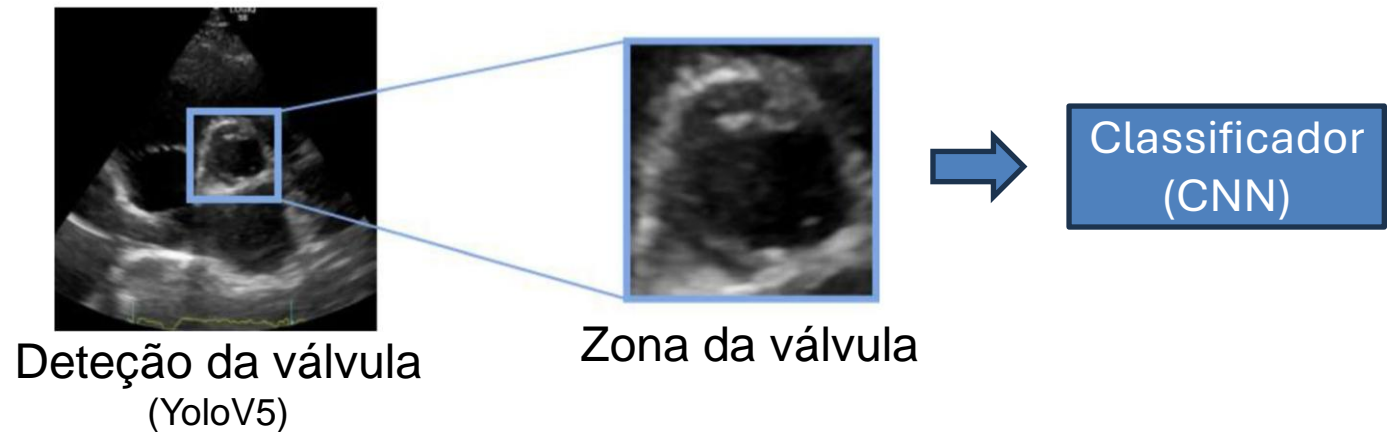
- Auxiliar o diagnóstico de uma doença do coração, a calcificação da válvula aórtica, através de ecocardiografias
- Detetar automaticamente a válvula aórtica
- Classificar a presença de cálcio

■ Técnicas usadas

- *Deteção de objetos*
- *Classificação de imagens*

■ Resultados (preliminares)

- Deteção/localização da válvula: F1-Score de 98%
- Classificação de cálcio: 96% de acertos (MobileNetV3Large com *transfer learning*)



E muitos outros trabalhos...

- Matilde Saraiva, *Detection of hardly-visible road networks in low-resolution satellite imagery*, 2024, MCD
- João Pedro Ferreira, *Deep learning for personal protection equipment (PPE) detection in real-life scenarios*, 2024, MCD
- Rita Bairros, *Calcium detection and scoring in 3D transoesophageal echocardiography based on computer vision*, 2024, MSIAD
- Sara Gomes, *Deep learning-based automatic calcium detection in echocardiography*, 2023, MSIAD
- Francisca Guedes, *Activity detection and classification in public sports spaces*, 2023, MIG
- Sérgio Santos, *Land cover automatic classification using deep learning techniques applied to satellite imagery*, 2023, METI
- Simão Correia, *Extraction of pose characteristics from customer images in large retail environments*, 2023, MEI
- Diogo Mendes, *Reidentificação de pessoas em ambientes comerciais multicâmara com base no seu percurso*, 2023, MEI
- Diana Mendes, *A Importância das imagens no alojamento local: uma abordagem data driven*, 2022, MMAG
- Afonso Gonçalves, *Wildfire detection with deep learning – a case study for the CICLOPE project*, 2022, MSIAD
- ...