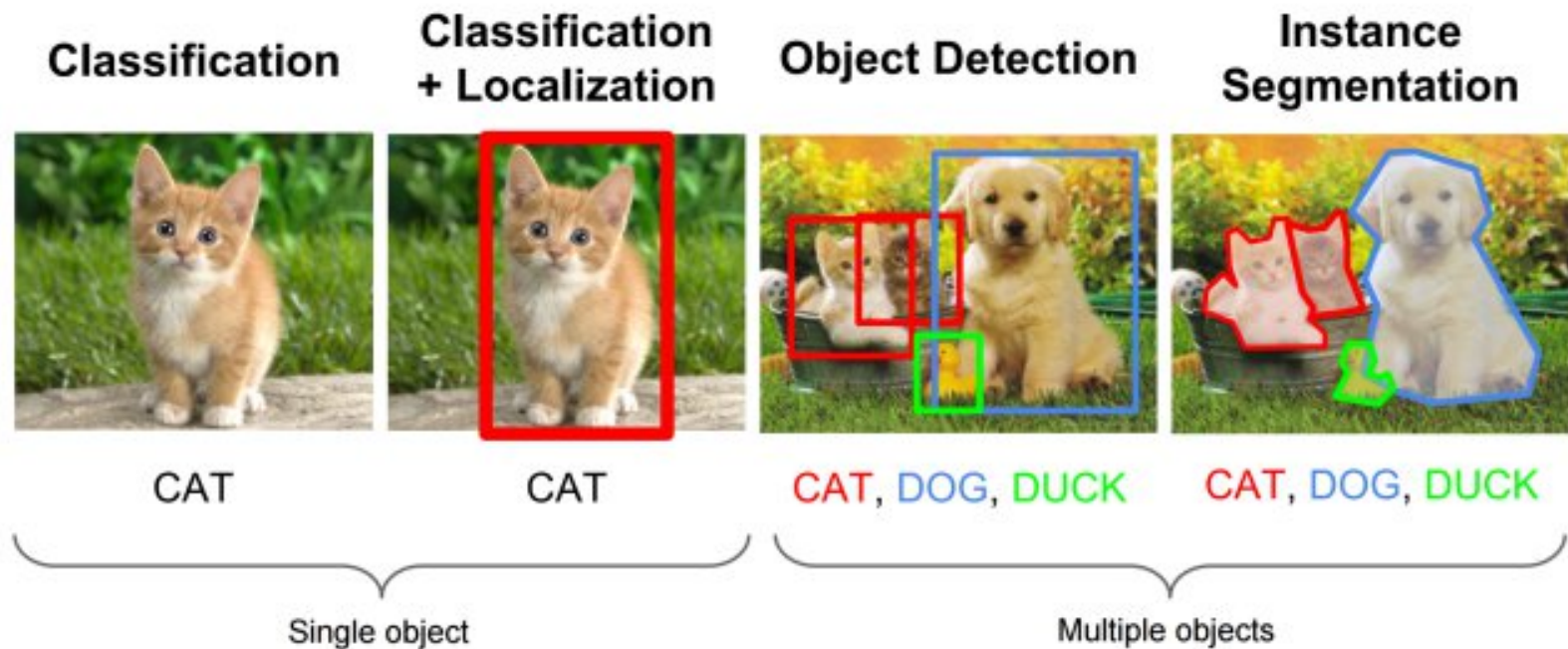


Robótica Computacional

Detecção de retas e circunferências

Visão de alto nível



Classificação binária



Cachorro

1



Não cachorro

-1

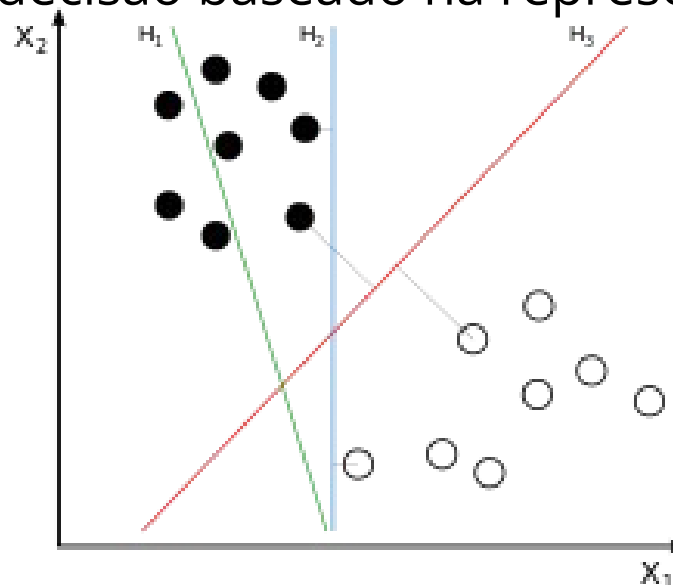
Classificação binária

Entrada:

- Conjunto de imagens, cada uma identificada como pertencendo ou não a uma classe

Saída:

- Critério de decisão baseado na representação das imagens



Classificação binária (probabilidades)



Cachorro

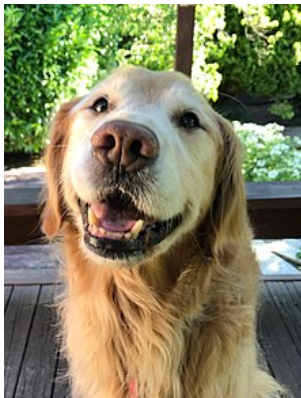
$P(\text{cachorro}) \geq P(\text{pessoa})$



Não cachorro

$P(\text{cachorro}) < P(\text{pessoa})$

Classificação multi-classe (categórica)



Cachorro

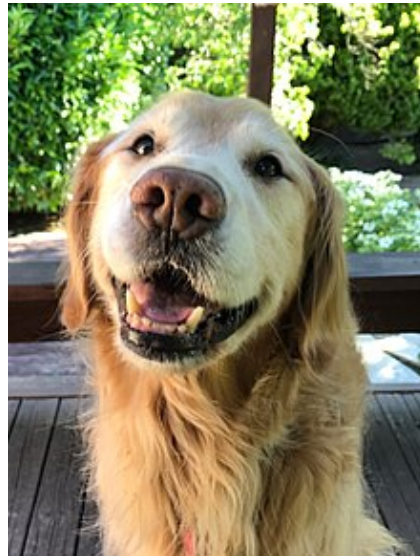


Pessoa



Cavalo

Classificação multi-classe (categórica)



$P(\text{Cachorro}) = 1$

$P(\text{Pessoa}) = 0$

$P(\text{Cavalo}) = 0$

Imagem conhecida

Classificação multi-classe (categórica)



$P(\text{Cachorro}) = 0.7$

$P(\text{Pessoa}) = 0.0$

$P(\text{Cavalo}) = 0.3$

Imagem nova

Classificação multi-classe

Entrada:

- Conjunto de imagens, cada uma identificada como um número correspondente a sua classe

Saída:

- Função densidade de probabilidade → atribui a cada objeto probabilidades de pertencer a cada classe

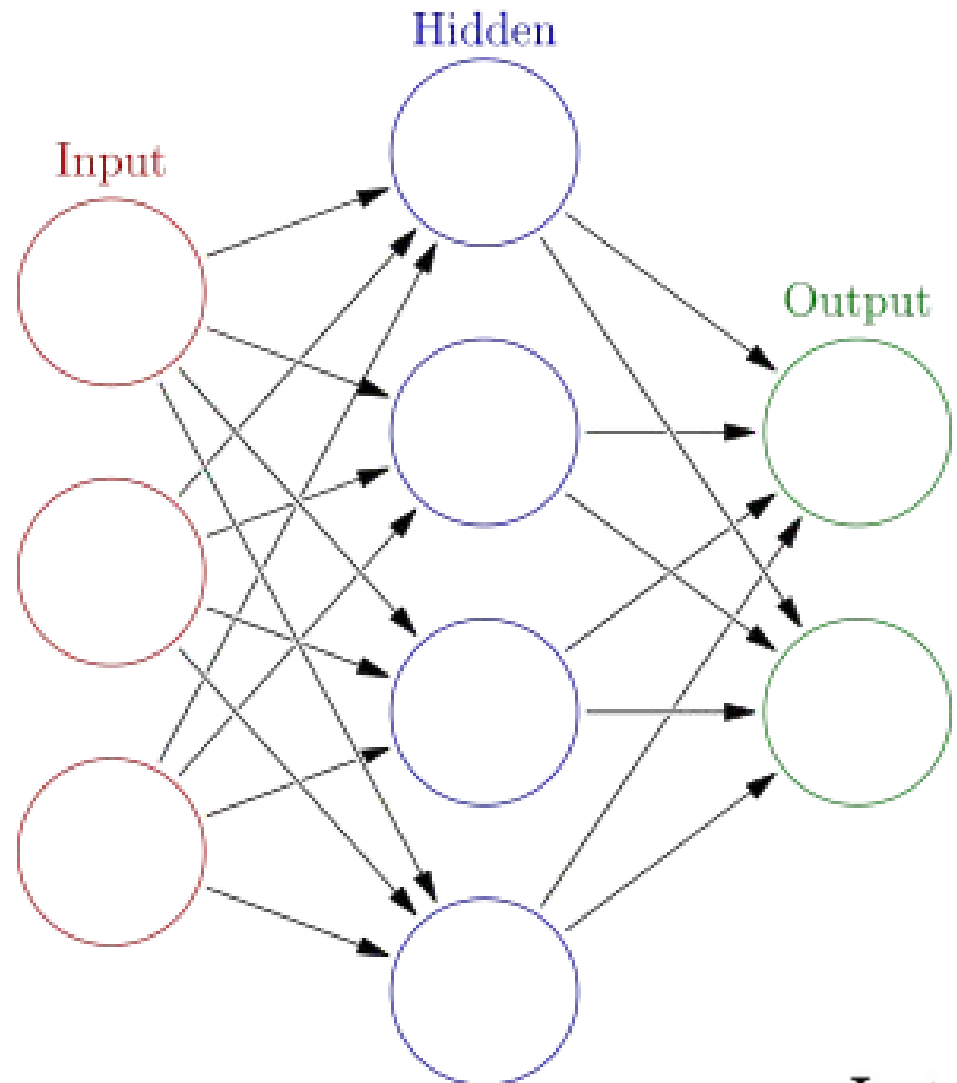
Interpretação:

- Objeto pertence à classe com maior probabilidade

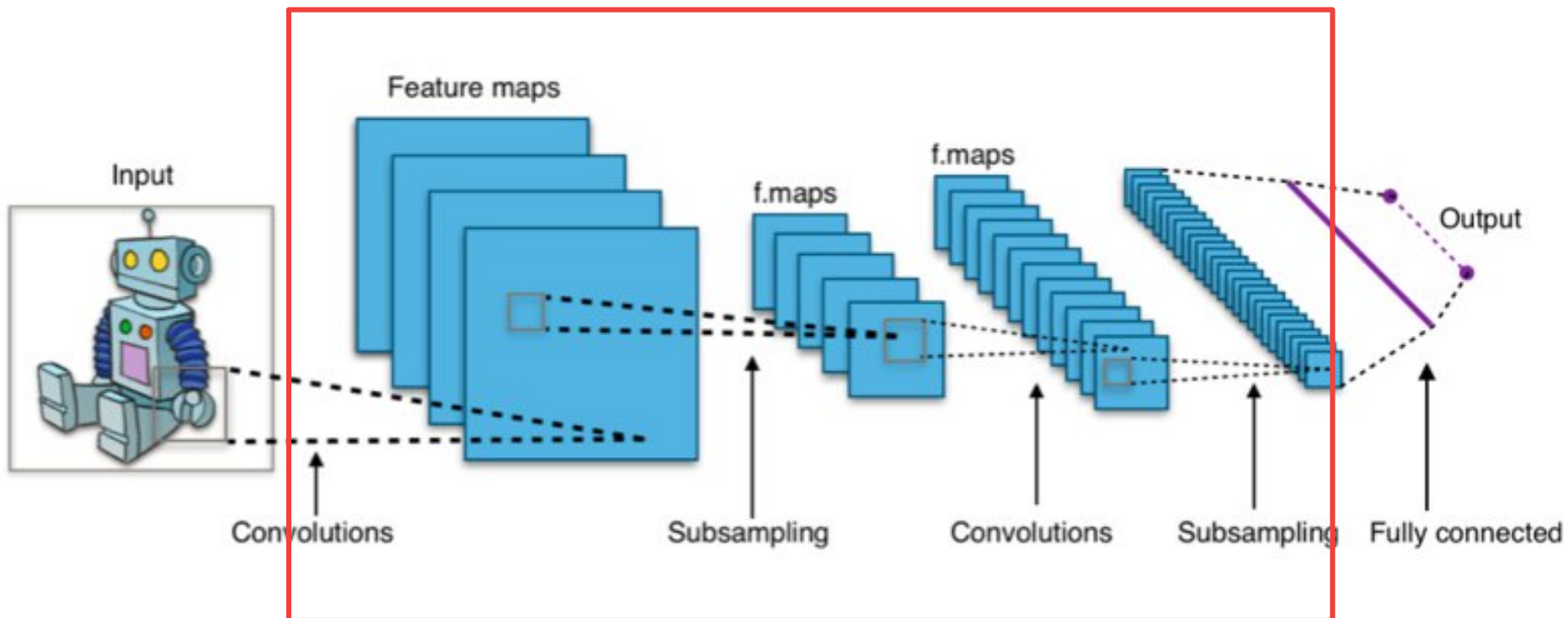
Redes Neurais

Modelo em camadas:

- Cada camada processa a saída da anterior
- Camadas são compostas por nós (ou neurônios) que combinam a saída da camada anterior usando uma função não linear.
- A camada final retorna probabilidade de cada classe



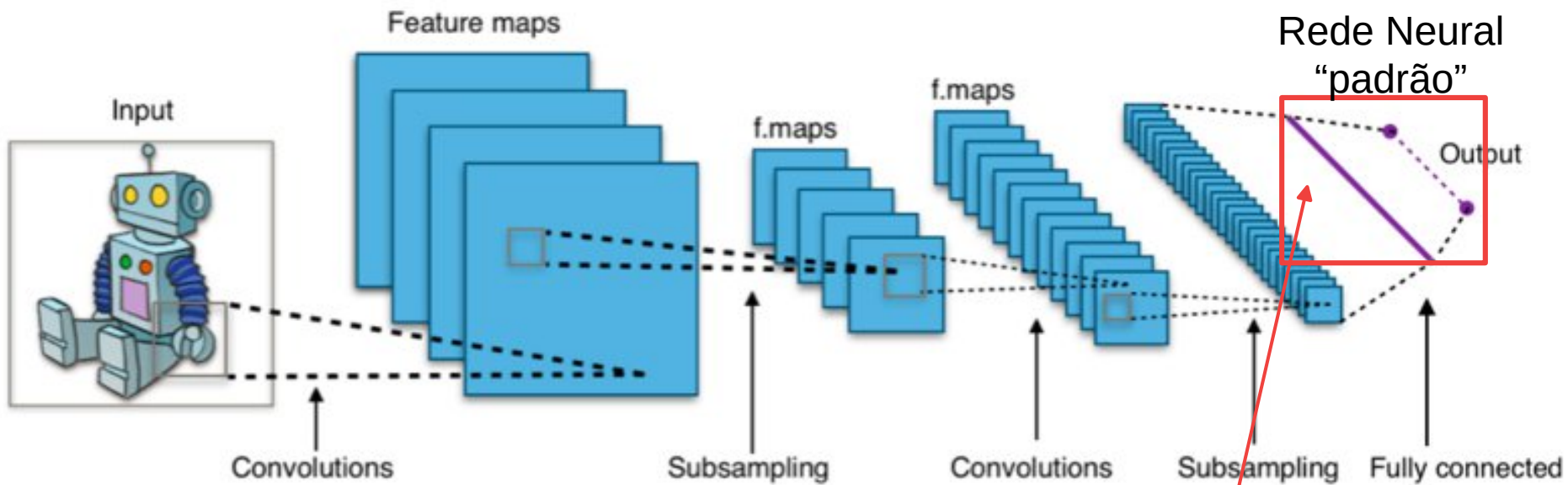
Redes Convolucionais (imagens)



Transformações de
imagens:

1. Convoluções
2. Redimensionamentos

Redes Convolucionais



Representação da
imagem

Redes neurais - desafios

- Treinamento
 - Lento
 - Instável: rodar várias vezes dá resultados diferentes
 - Necessita de grandes quantidades de dados
- Como definir arquitetura?
 - Quantos níveis
 - Quantos nós em cada nível
 - Tamanho e quantidade de convoluções
 - Tamanho e quantidade de redimensionamentos

Redes neurais - reprodutibilidade

- Muitas arquiteturas testadas;
- Treinamento caro, lento e não reprodutível
- Possibilidade de aplicação em diferentes problemas

Trabalhos científicos disponibilizam redes pré-treinadas em bancos de imagens gigantes

Classificação de Imagens

IMAGENET Large Scale Visual Recognition Challenge 2017 (ILSVRC2017)

[Introduction](#) [News](#) [History](#) [Timetable](#) [Challenges](#) [FAQ](#) [Citation](#) [Contact](#)

Introduction

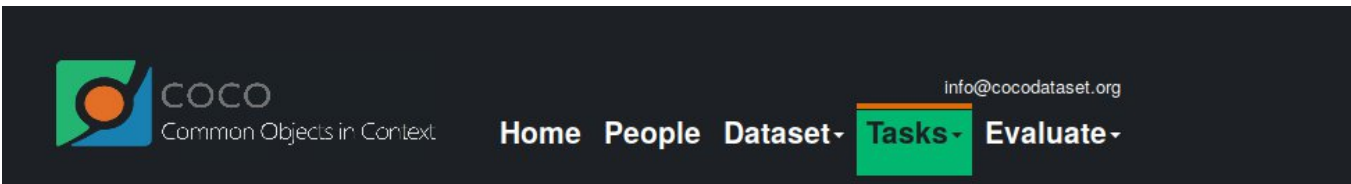
This challenge evaluates algorithms for object localization/detection from images/videos at scale. Most successful and innovative teams will be invited to present at [CVPR 2017 workshop](#).

- I. [Object localization](#) for 1000 categories.
- II. [Object detection](#) for 200 fully labeled categories.
- III. [Object detection from video](#) for 30 fully labeled categories.

News

- Jul 26, 2017: We are passing the baton to [Kaggle](#). From now on, all three challenges(LOC-CLS, DET, VID) will be hosted on Kaggle!
- Jul 17, 2017: [Results](#) announced.
- Jun 25, 2017: [Submission server](#) for VID is open, new additional train/val/test images for VID is available now, deadline for VID is extended to July 7, 2017 5pm PDT.
- Jun 18, 2017: [Submission server](#) for CLS-LOC and DET is open.
- Jun 15, 2017: Taster challenges with [amazon bin image dataset](#) will not be held. There were some issues on final dataset release. We sincerely apologize to the teams that have been working on this challenge.
- Jun 12, 2017: New additional test set(5,500 images) for object detection is available now.
- Mar 31, 2017: Register your team and download data at [here](#).
- Mar 31, 2017: Tentative time table is announced.
- Mar 13, 2017: Stay tuned. Coming soon.

Detecção de objetos



COCO 2018 Object Detection Task



1. Overview

The COCO Object Detection Task is designed to push the state of the art in object detection forward. COCO features two object detection tasks: using either bounding box output or object segmentation output (the latter is also known as instance segmentation). For full details of this task please see the [detection evaluation](#) page. Note: **only the detection task with object segmentation output will be featured at the COCO 2018 challenge** (more details follow below).

This task is part of the [Joint COCO and Mapillary Recognition Challenge Workshop](#) at ECCV 2018. For further details about the joint workshop please visit the workshop page. Researchers are encouraged to participate in both the COCO and Mapillary Object Detection Tasks (the tasks share identical data formats and evaluation metrics). Please also see the related COCO [keypoint](#), [stuff](#), and [panoptic](#) tasks. Whereas the detection task addresses *thing* classes (person, car, elephant), the [stuff](#) task focuses on *stuff* classes (grass, wall, sky) and the newly introduced [panoptic](#) task addresses both simultaneously.

The COCO train, validation, and test sets, containing more than 200,000 images and 80 object categories, are available on the [download](#) page. All object instances are annotated with a detailed segmentation mask. Annotations on the training and validation sets (with over 500,000 object instances segmented) are publicly available.

APS 03 – Parte 2

Usaremos redes pré-treinadas para detectar objetos em imagens.

- MobileNet é o nome da rede que usaremos
- Caffe é um formato de dados para redes pré-treinadas
- O módulo `cv2.dnn` tem funções para trabalhar com redes neurais aplicadas a visão. Buscar em sua documentação é muito útil
- Ler recursos apontados no página do módulo atual

Fontes

- Rede Convolutcional – By Aphex34 [CC BY-SA 4.0 (<https://creativecommons.org/licenses/by-sa/4.0>)], from Wikimedia Commons
- Rede neural – By Glosser.ca [CC BY-SA 3.0 (<https://creativecommons.org/licenses/by-sa/3.0>)], from Wikimedia Commons
- Outras imagens nos outros slides

Insper

www.insper.edu.br