

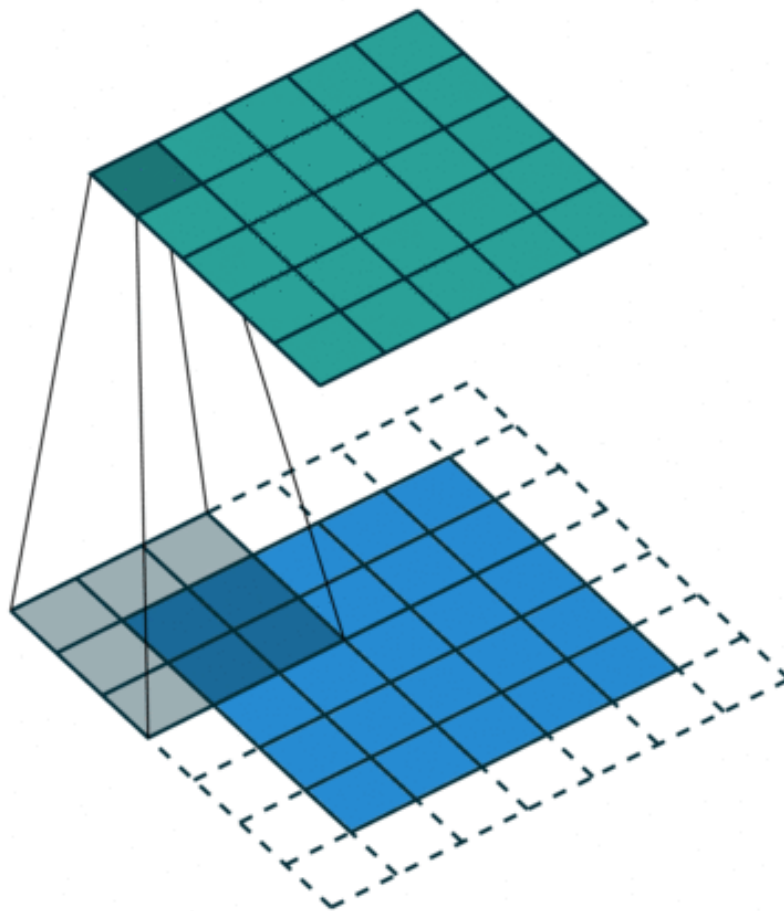
Robótica Computacional

Visão de Alto Nível

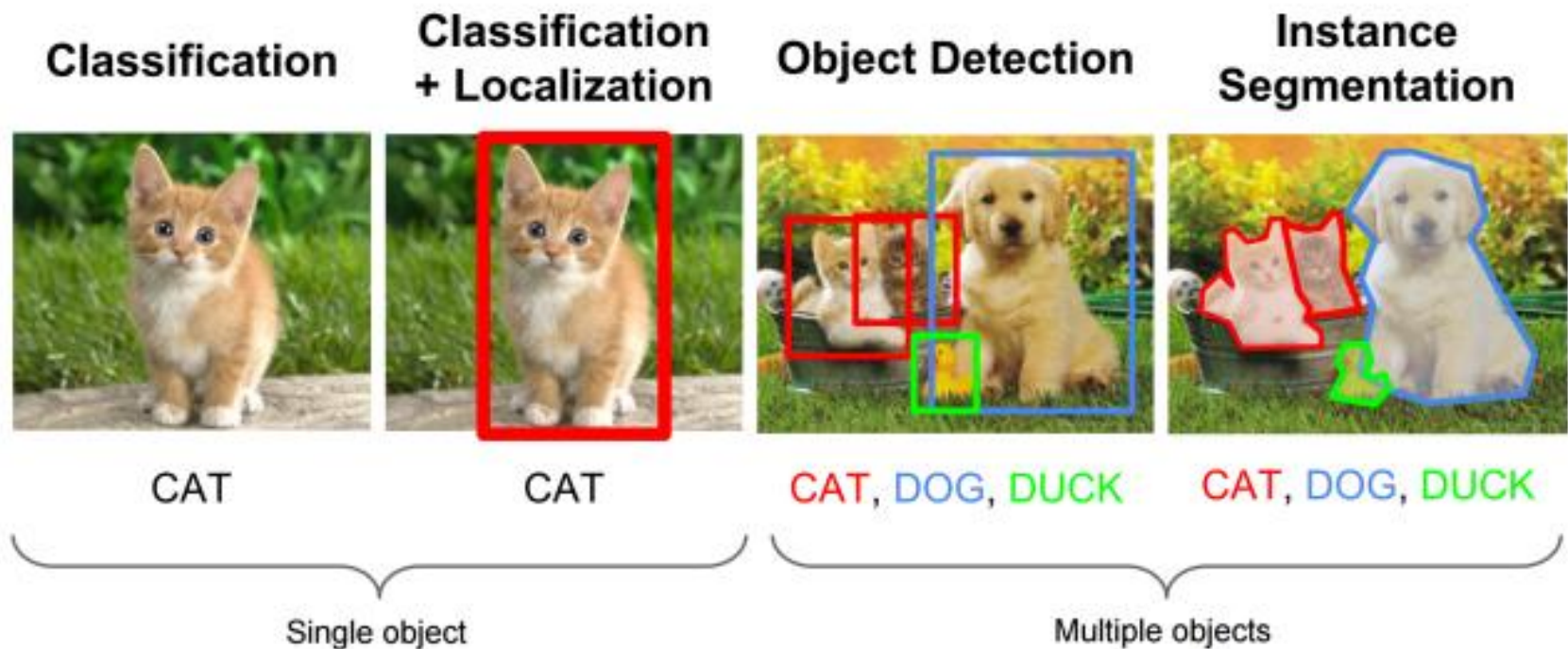


Visão de Alto Nível

Convolução em 2D



Problemas em Visão Computacional



Classificação Binária



Cachorro

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



Não cachorro

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

Classificação multi-classe (categórica)



$P(\text{Cachorro}) = 0.7$
 $P(\text{Pessoa}) = 0$
 $P(\text{Cavalo}) = 0$



$P(\text{Cachorro}) = 0$
 $P(\text{Pessoa}) = 0.9$
 $P(\text{Cavalo}) = 0$

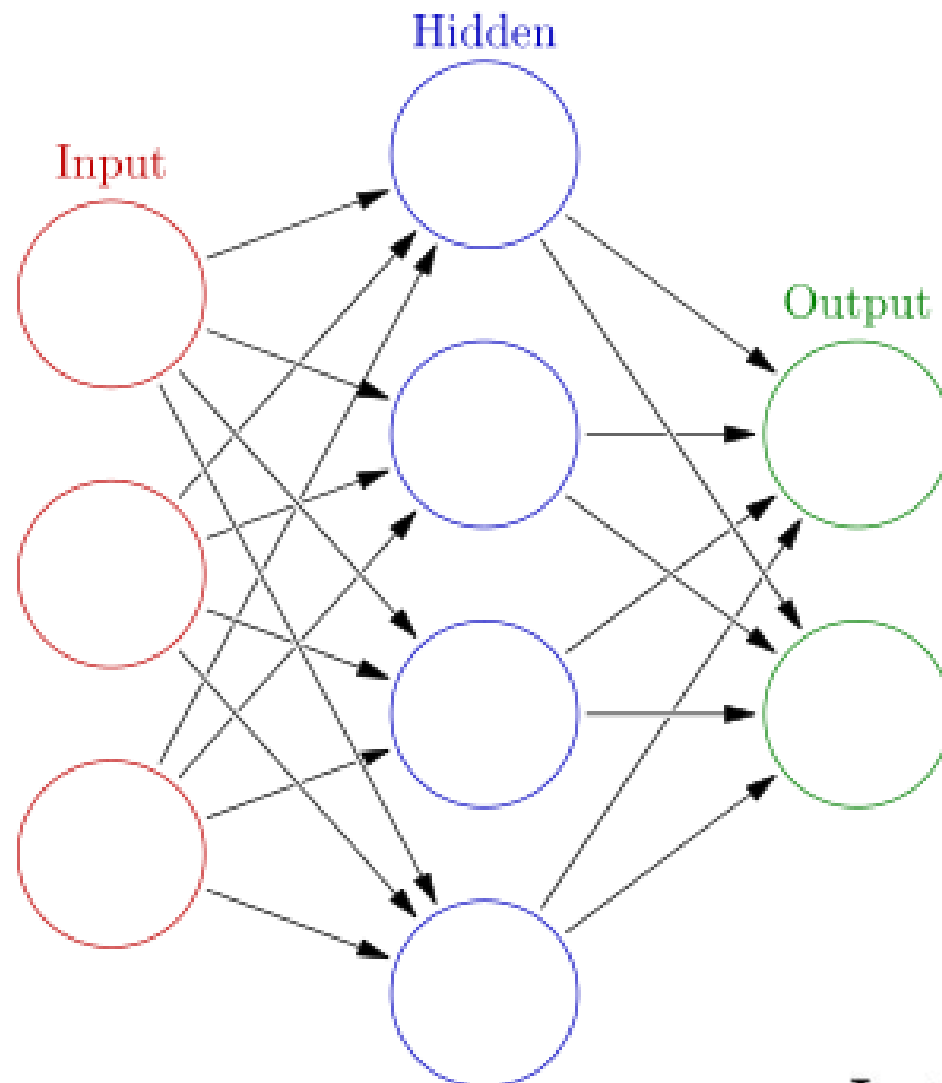


$P(\text{Cachorro}) = 0$
 $P(\text{Pessoa}) = 0$
 $P(\text{Cavalo}) = 0.6$

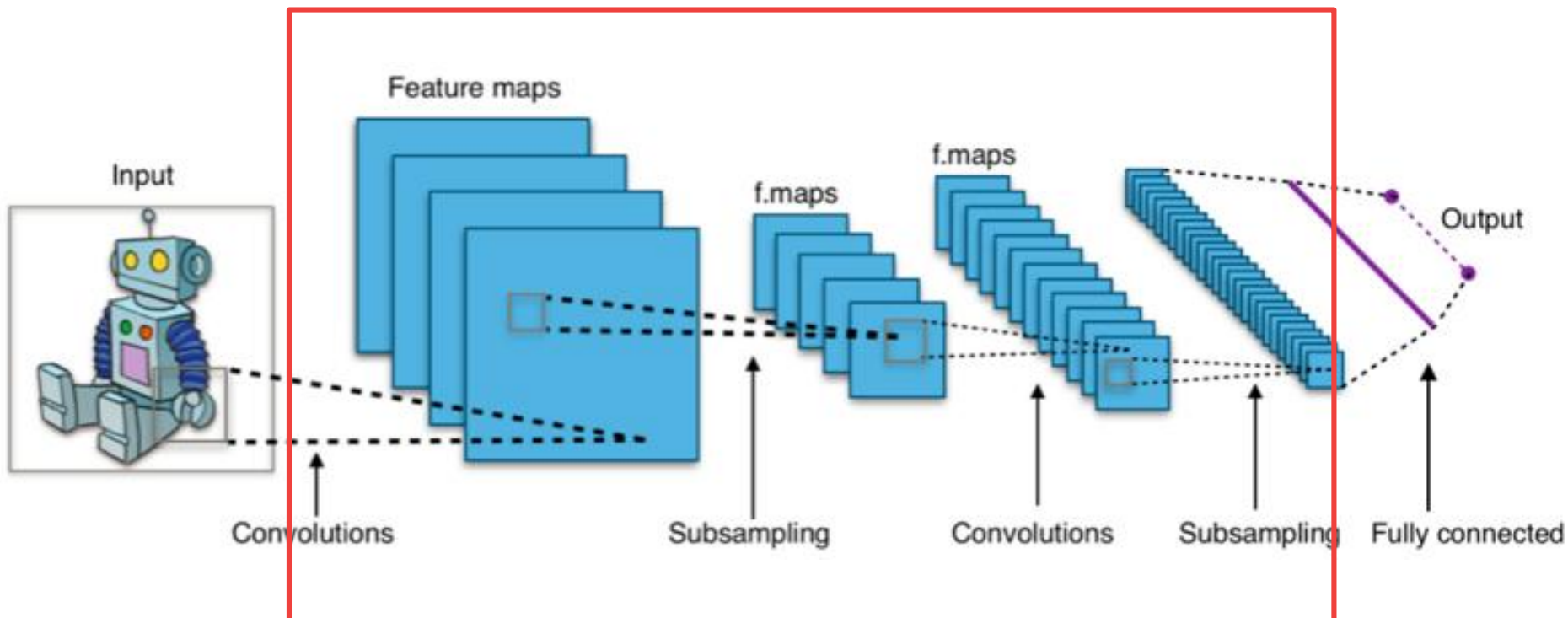
Redes Neurais

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



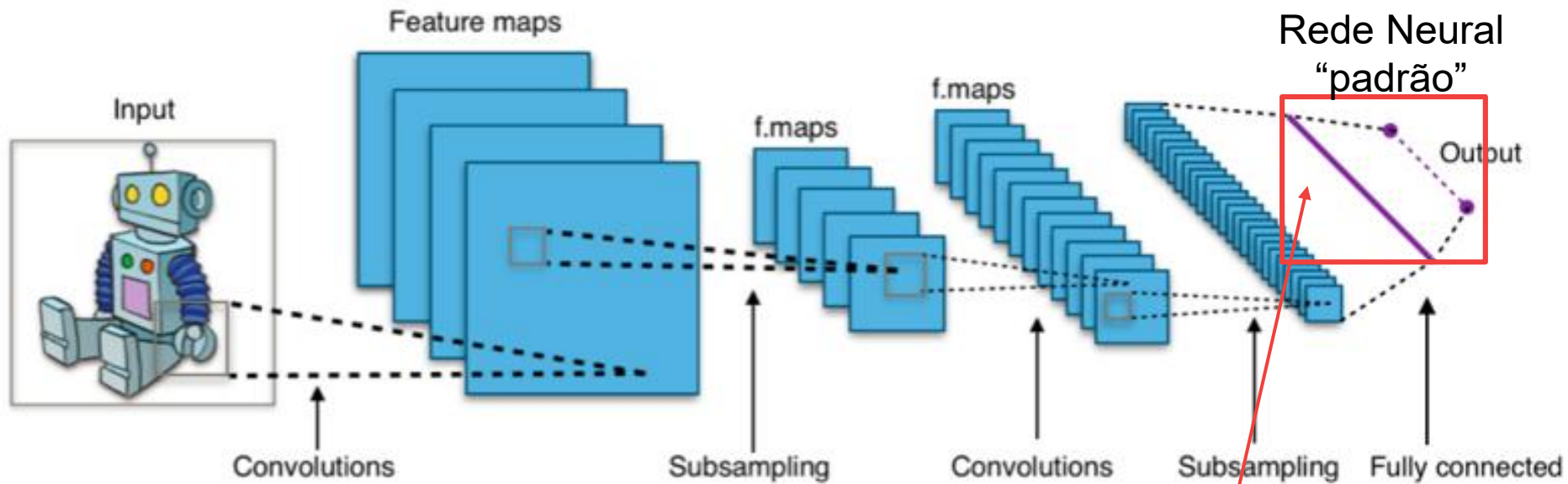
Redes Convolucionais (imagens)



Transformações de
imagens:

1. Convoluções
2. Redimensionamentos

Redes Convolucionais



Representação da
imagem

Redes Neurais – Problemas (Tamanho)

VGG19

- Ano: 2014
- 143 milhões de parâmetros
- Cerca de 93% de acurácia top-5 na ImageNet

ResNet-152

- Ano: 2015
- 60 milhões de parâmetros
- Cerca de 96% de acurácia top-5 na ImageNet

EfficientNet-B7

- Ano 2019
- 66 milhões de parâmetros
- Cerca de 97% de acurácia top-5 na ImageNet

MobileNetV2

- 2017
- 4.2 milhões de parâmetros
- Cerca de 95% de acurácia top-5 na ImageNet

Aplicações



Mask RCNN - Segmentação de objetos - COCO
[Video](#) - [artigo](#)