

# Introdução à Visão Computacional e Aprendizado Profundo com Python

Como essa  
tecnologia auxilia  
no diagnóstico de  
doenças





## **Clésio Gonçalves**

Mestre em Engenharia Elétrica  
e Doutorando em Ciência da  
Computação - UFPI. Professor  
de Informática do IFSertãoPE -  
Campus Ouricuri

# Repositório do minicurso

- <https://github.com/clesio-goncalves/MinicursoIFPI2024>

# Conteúdo

01

**Inteligência Artificial**

02

**Visão Computacional**

03

**PAVIC**

04

**Aprendizado  
Profundo**

01

# Inteligência Artificial



# Conceito

- A Inteligência Artificial (IA) refere-se à capacidade de máquinas aprenderem e tomarem decisões de maneira autônoma
  - Algoritmos inteligentes;
  - Dados.
  - Uso de GPU para treinamento



**Em julho de 2020**, a OpenAI revelou o GPT-3, o maior modelo de linguagem então conhecido.

O GPT-3 possui **175 bilhões de parâmetros** e foi treinado em **570 gigabytes de texto**. Para efeito de comparação, seu antecessor, **GPT-2**, era **100 vezes menor**, com **1,5 bilhão de parâmetros**.

Fonte: THE AI INDEX REPORT 2024

O GPT-4 da OpenAI usou cerca de **US\$ 78 milhões** em computação para treinar, enquanto o Gemini Ultra do Google custou **US\$ 191 milhões** em computação.

Fonte: THE AI INDEX REPORT 2024



# Aplicações



## Energia

Redirecionamento de energia, inspeções, etc.



## Assistentes de voz

Siri, Google Home, Alexa, etc.



## Aplicativos de streaming

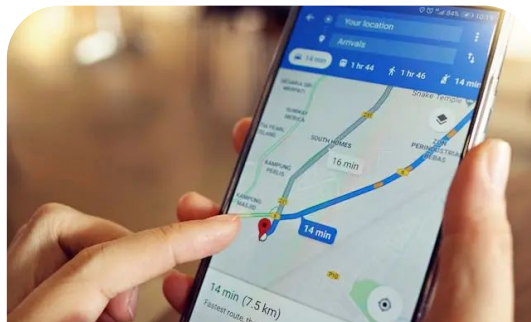
YouTube, Spotify, Netflix, etc.

# Aplicações



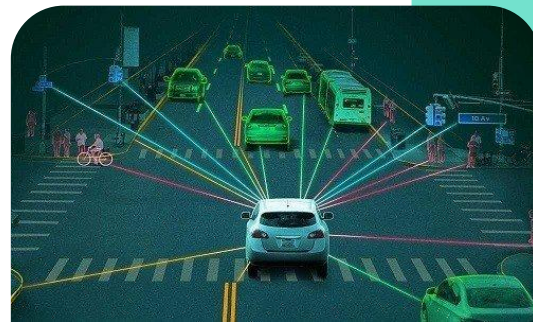
## Marketing personalizado

Ajuda as marcas a alcançar as pessoas certas.



## Navegação e viagem

Exemplos: Google Maps e Waze.



## Veículos autônomos

IA ensinando os veículos a operar de forma independente

# Aplicações



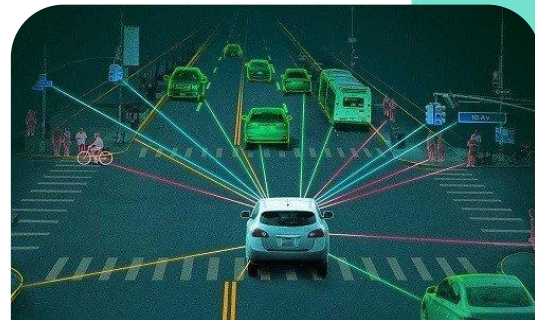
## Reconhecimento facial

Aplicações de segurança, identificação de humor, etc.



## Segurança e Vigilância

Exemplo: monitoramento e verificação de incidentes.



## Internet das Coisas

Dispositivos inteligentes.

# Na saúde

ooo

“Inteligência Artificial em medicina é o uso de computadores que, analisando um **grande volume de dados** e seguindo **algoritmos** definidos por especialistas na matéria, são capazes de propor soluções para problemas médicos.”

— Luiz Carlos Lobo (2017)

02

# Visão Computacional



# Visão Computacional

- A Visão Computacional tem por objetivo final simular o olho humano para realizar análises e aprender padrões a partir de entradas visuais





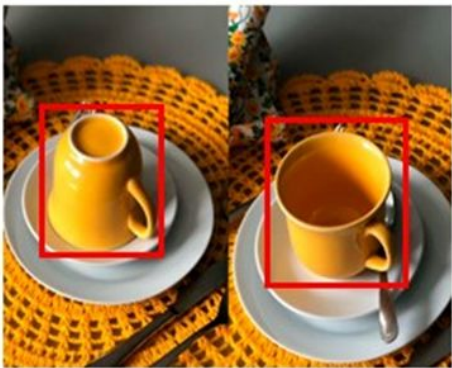
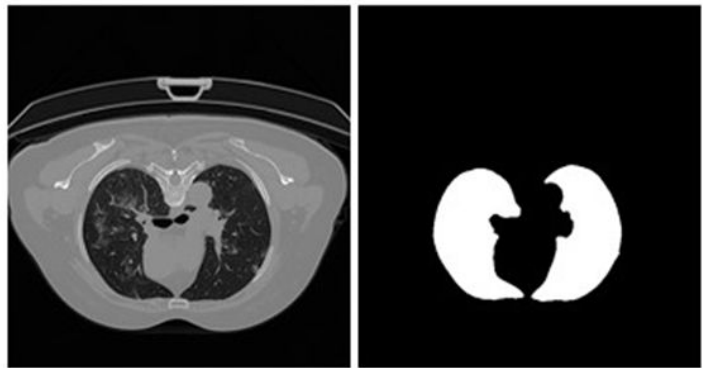
# **Visão Computacional**

Modelagem e  
replicação da  
visão humana  
usando  
software e  
hardware

analisar, interpretar  
e extrair informações  
relevantes de  
imagens e/ou vídeos



# Principais tarefas realizadas utilizando a Visão Computacional

Classificação	Detecção de objetos	Segmentação
<div data-bbox="106 480 193 513">Gato</div>  <div data-bbox="338 480 492 513">Cachorro</div> 		

# Principais desafios



1

**Necessidade de  
grandes volumes de  
dados rotulados para  
treinamento**

2

**Risco de overfitting  
(ajuste excessivo) em  
modelos muito  
complexos**

3

**Alto custo  
computacional  
associado ao  
treinamento de redes  
profundas**

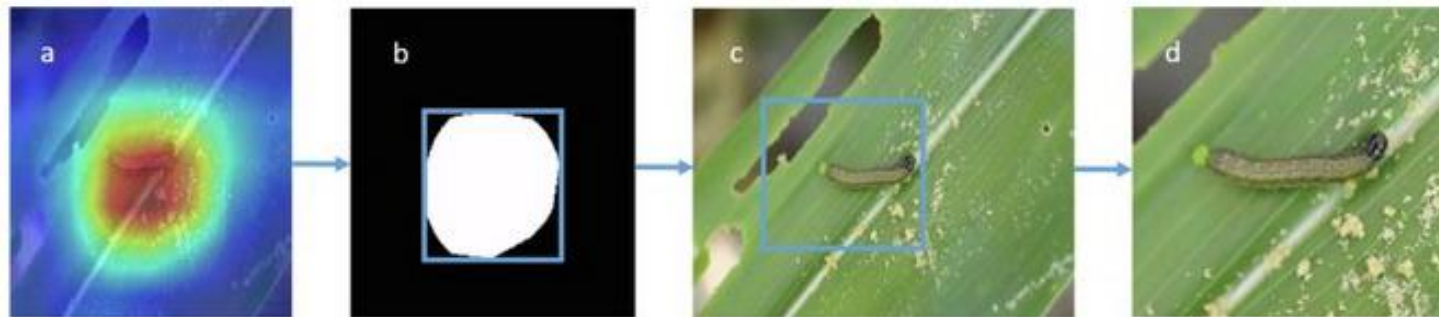
4

# Anotação das imagens para treinamento

5

Falta de  
interpretabilidade em  
modelos complexos  
(redes neurais  
profundas)

# Grad-CAM: aplicação em imagens





03

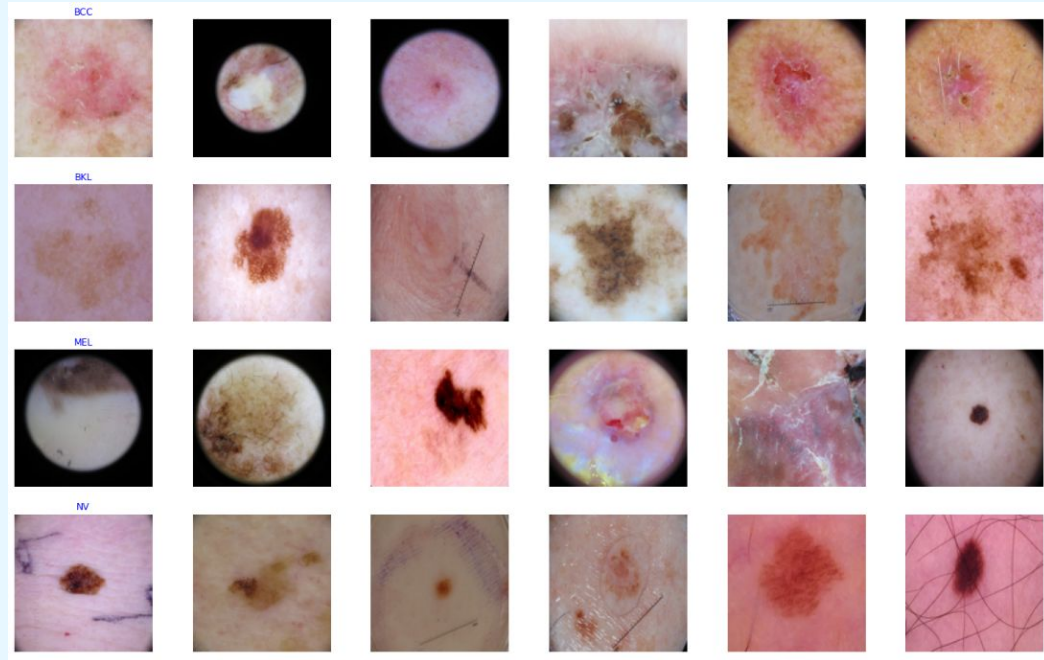
**PAVIC**



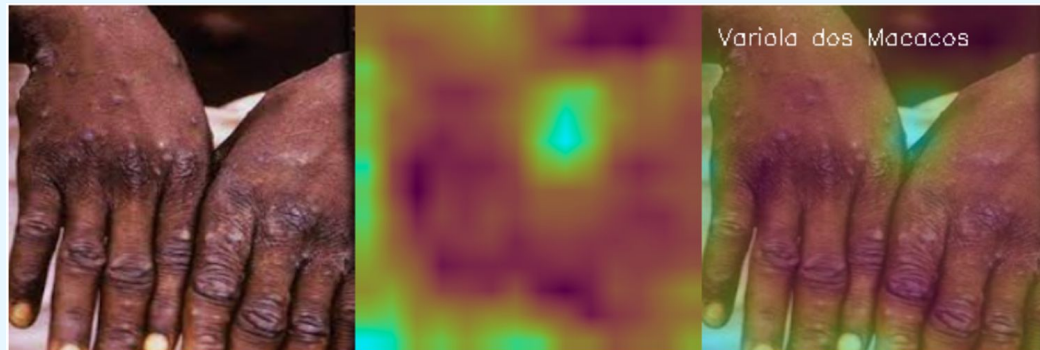
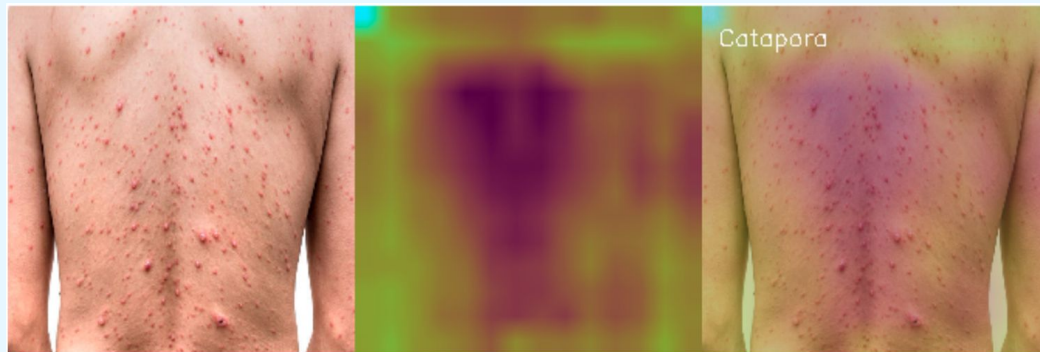
# Laboratório de Pesquisas Aplicadas à Visão e Inteligência Computacional - PAVIC UFPI



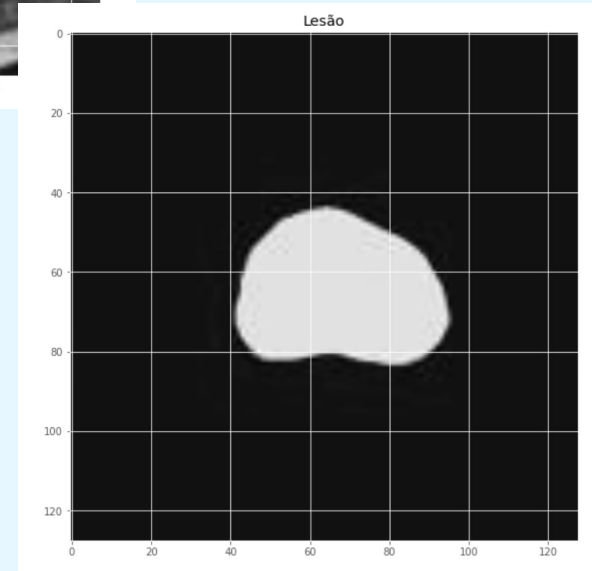
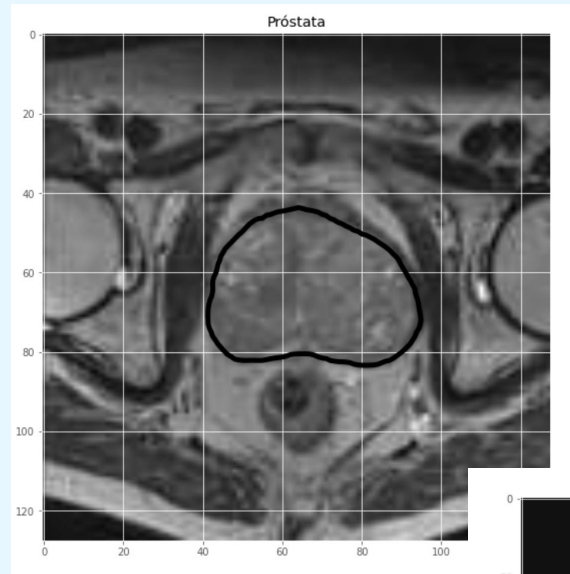
# Detecção de doenças de pele



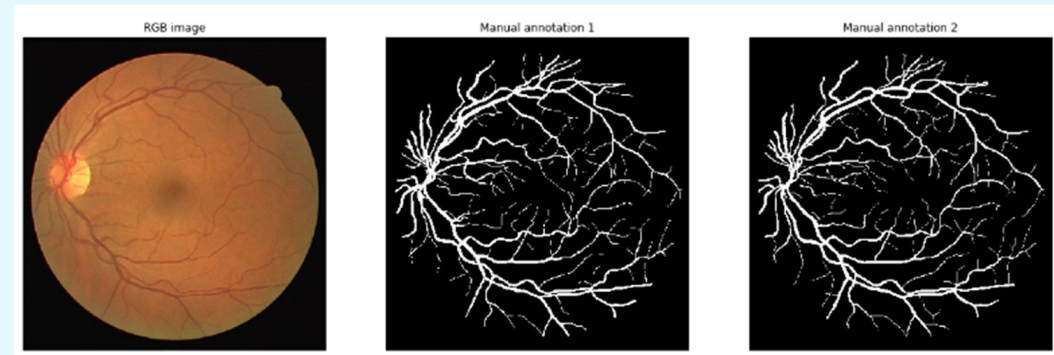
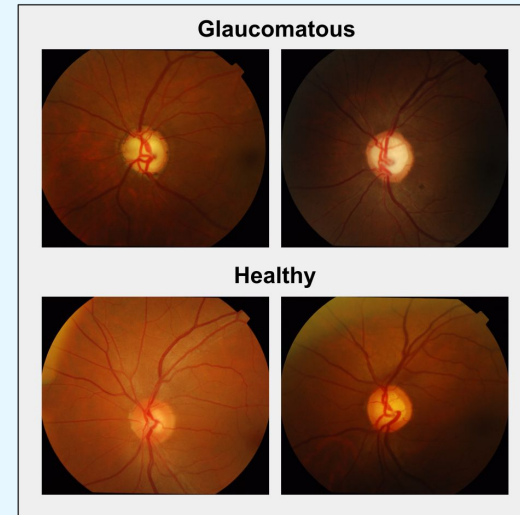
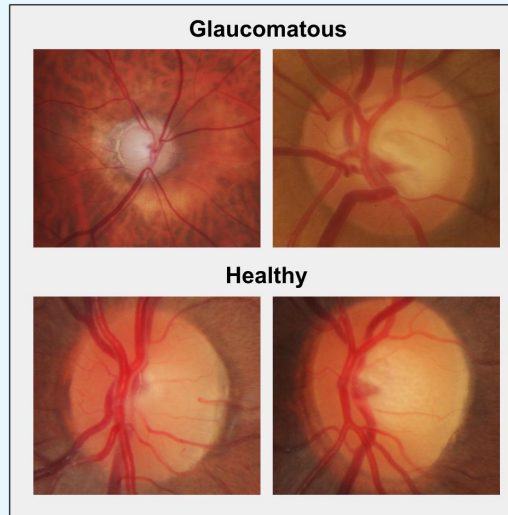
# Classificação das doenças pelas manifestações na pele



# Segmentação de deformações na próstata

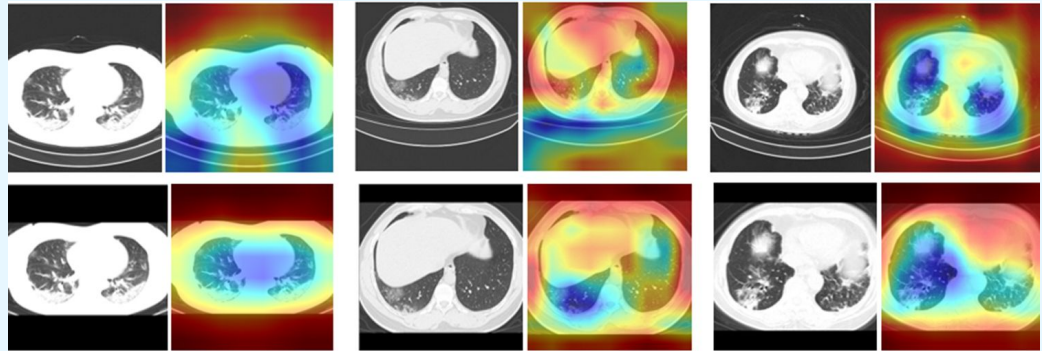


# Detecção de doenças oculares

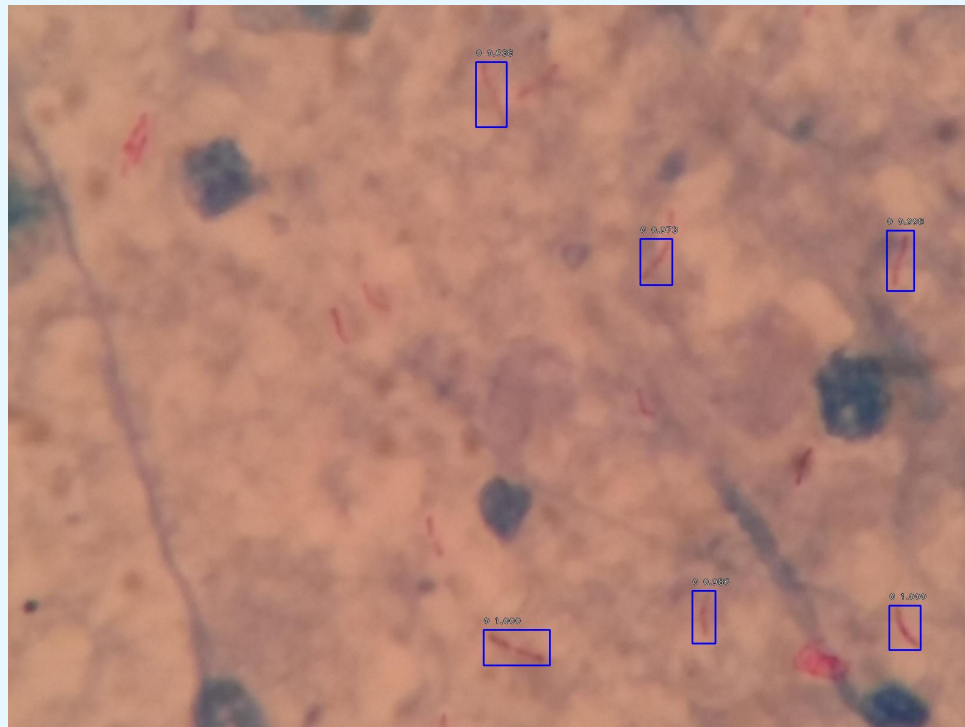




# Detecção de doenças pulmonares

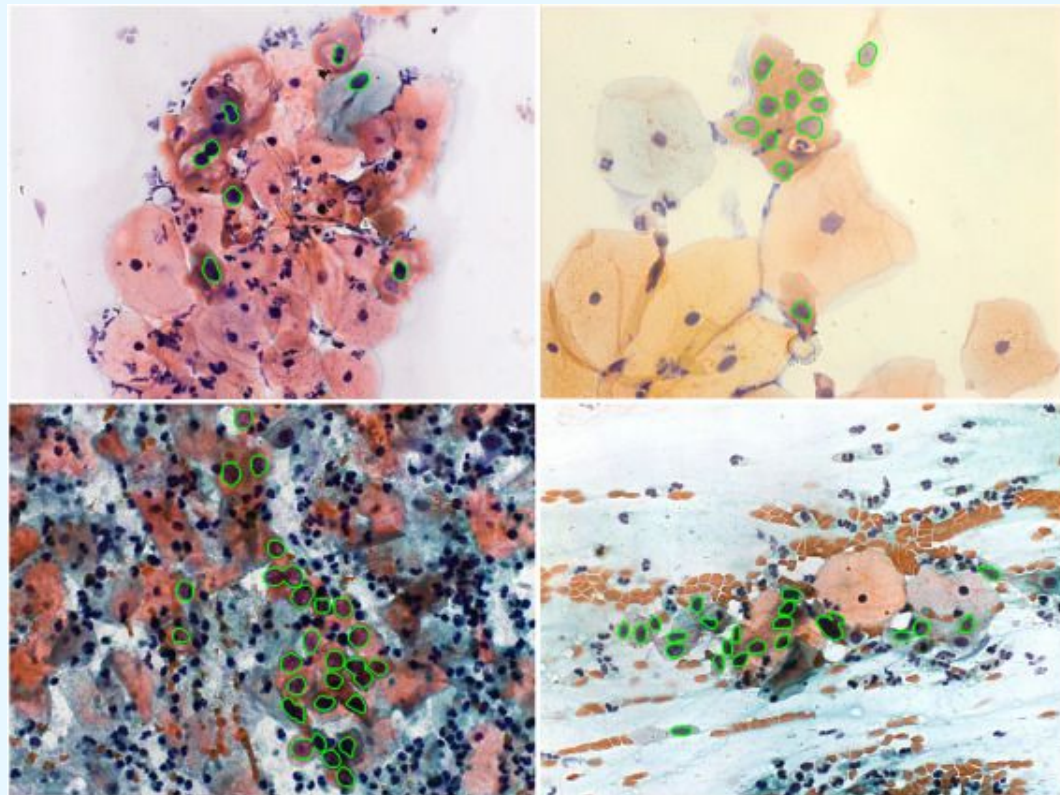


# Detecção de bacilos da Tuberculose

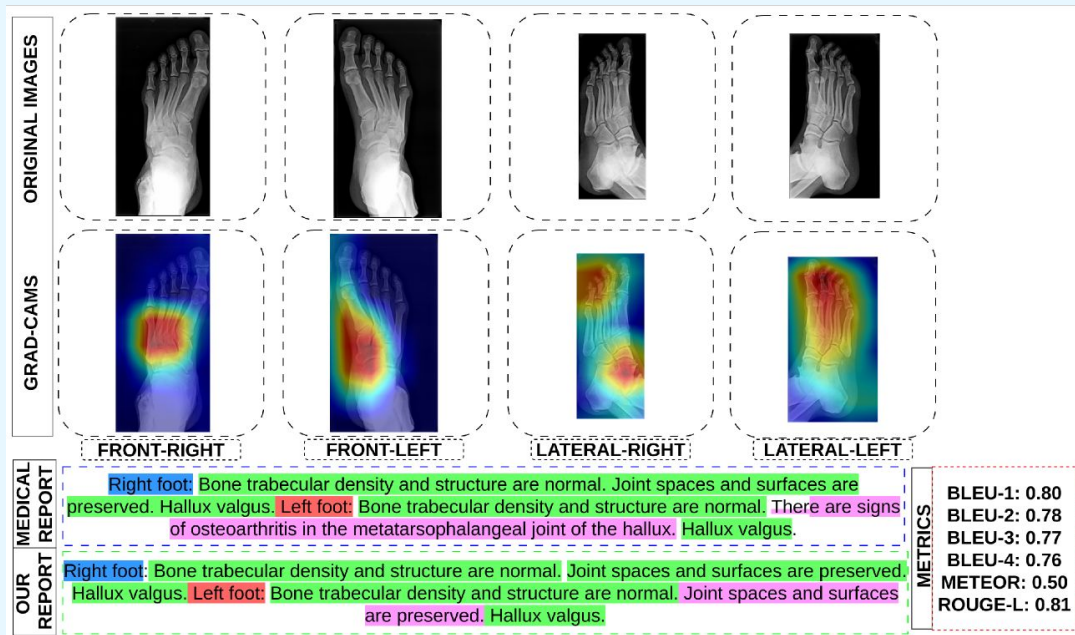




# Segmentação de células cervicais



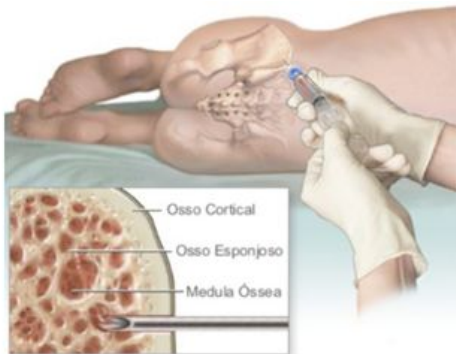
# Geração Automatizada de Relatórios Médicos a partir de Imagens de Raio-X



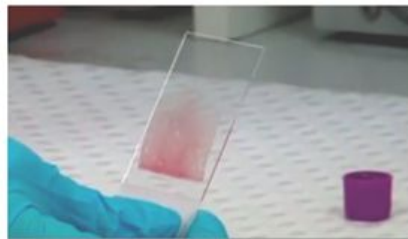
**Deteccção de Leishmaniose  
Visceral Humana Utilizando  
Aprendizado Profundo em  
Imagens de Microscopia do  
Exame Parasitológico da  
Medula Óssea**



# Aquisição das Imagens



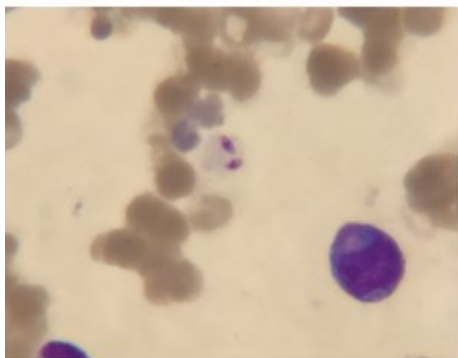
(a)



(b)

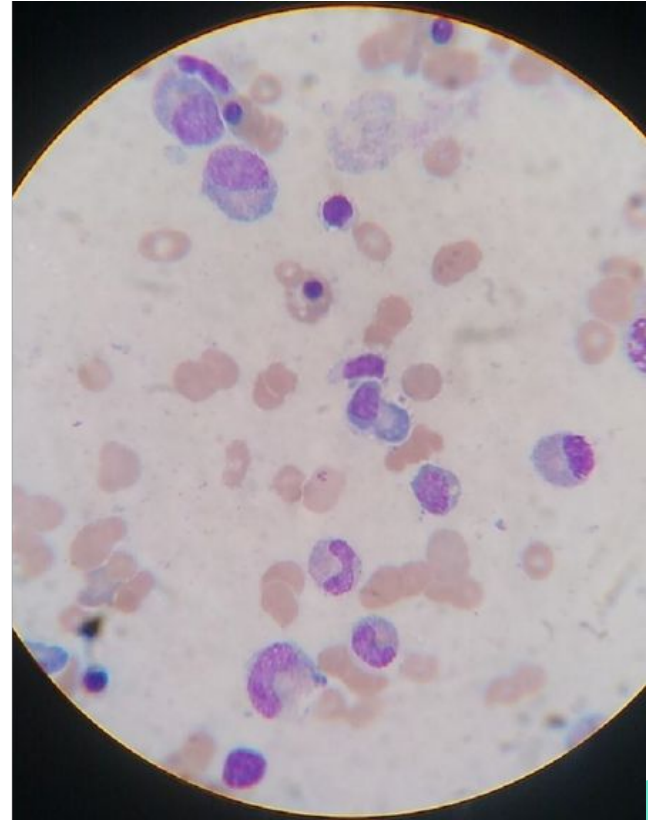
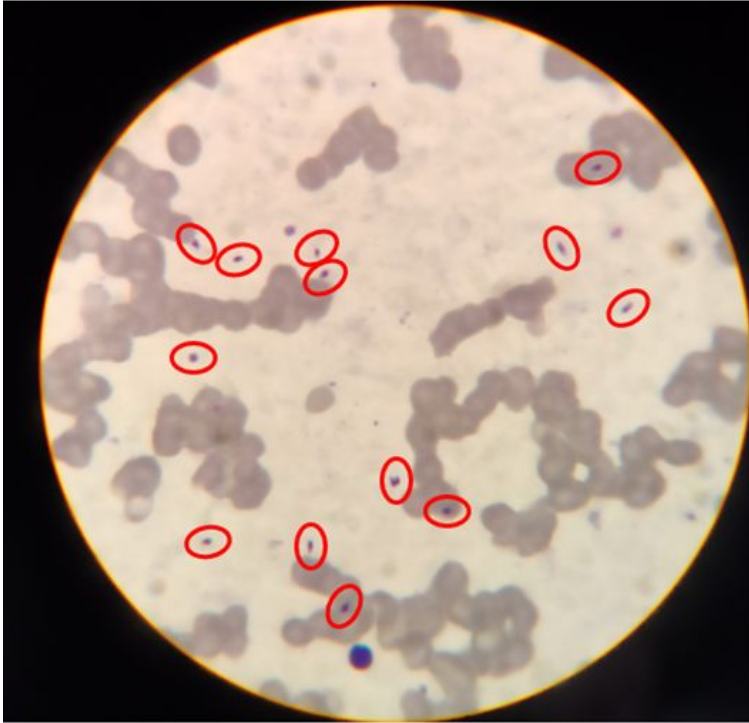


(c)



(d)

# Aquisição das Imagens



# Resultados

**97,3%**

## **Classificação**

Kappa superior a 97%  
e F1-Score superior a  
98% utilizando os  
modelos InceptionV3 e  
InceptionResNetV2

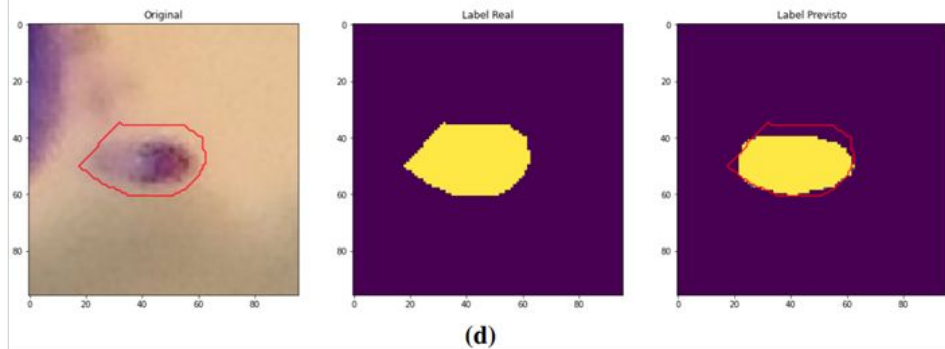
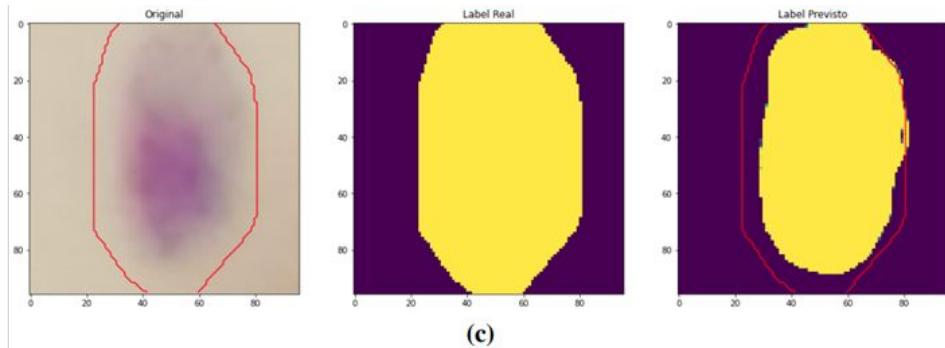
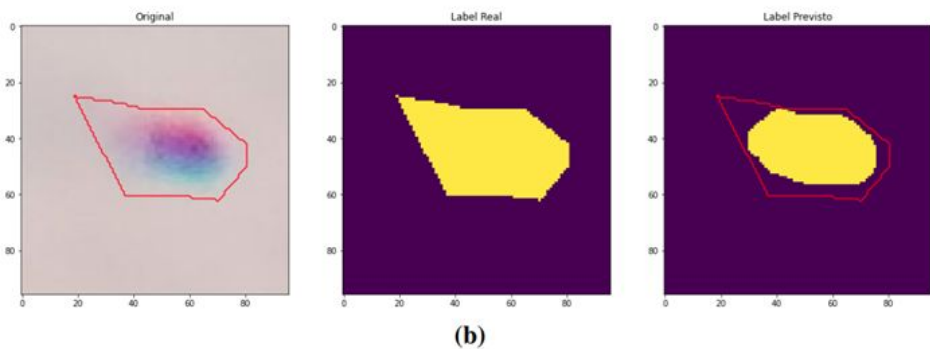
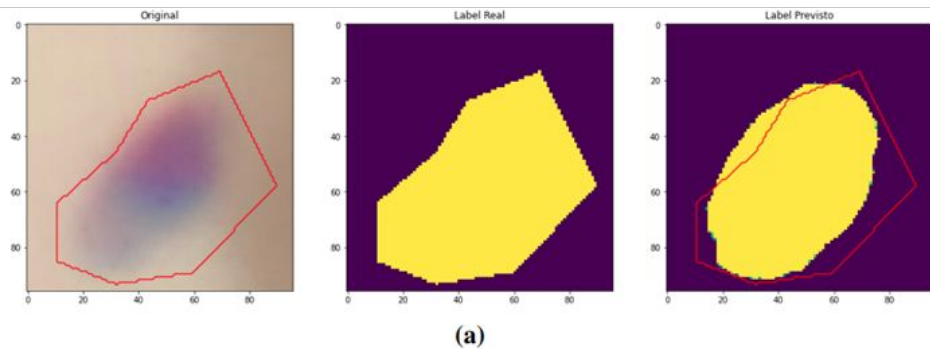
**80,4**

**%**

## **Segmentação**

Dice superior a 80%,  
utilizando recortes  
RGB com dimensões  
96x96 em uma U-Net

# Anotação Manual **vs** Modelo Desenvolvido



# **Nova parceria com o LAPEDONE**





## **LV Canina**

### **30 animais**

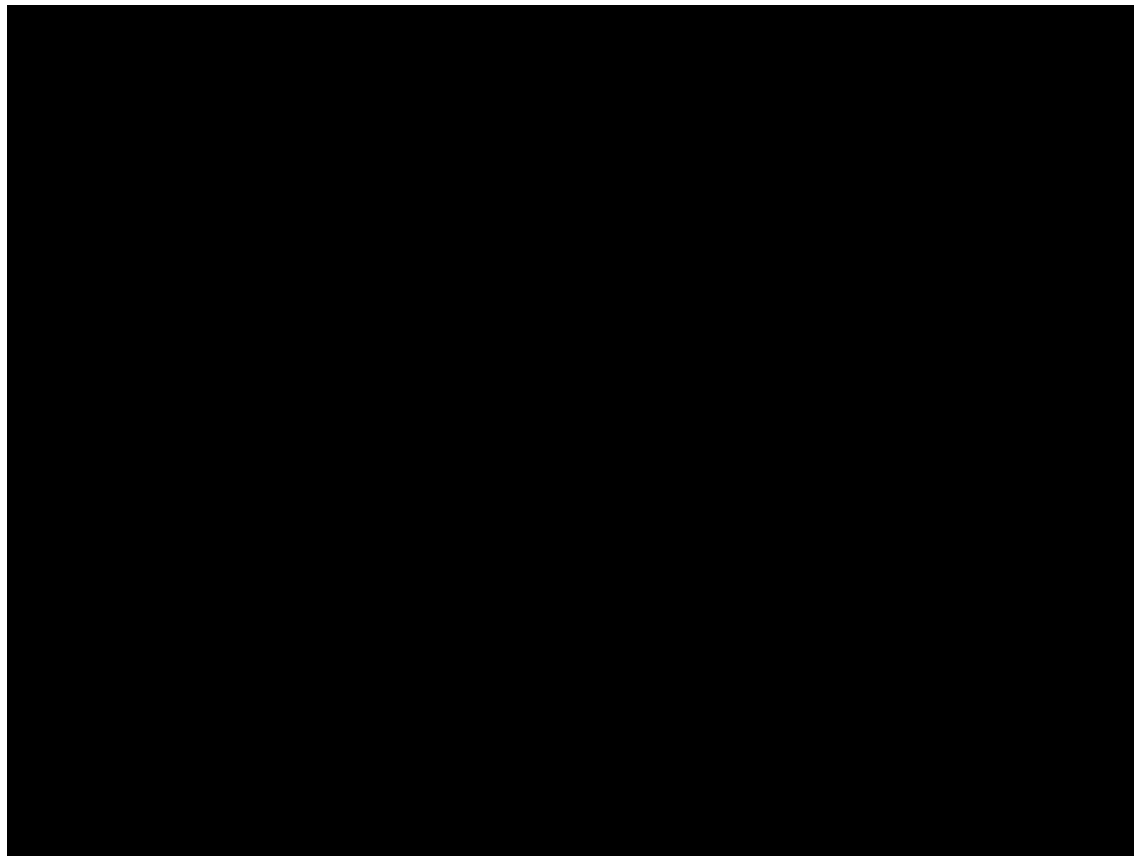
**crista  
ilíaca**

**úmero**

**esterno**

**fêmur**

# Automação na captura das imagens



**04**

# **Aprendizado Profundo**



# O que é Deep Learning ou Aprendizado Profundo?



# **Inteligência Artificial**

**Qualquer técnica que permita aos computadores imitar a inteligência humana**

## **Aprendizado de Máquina**

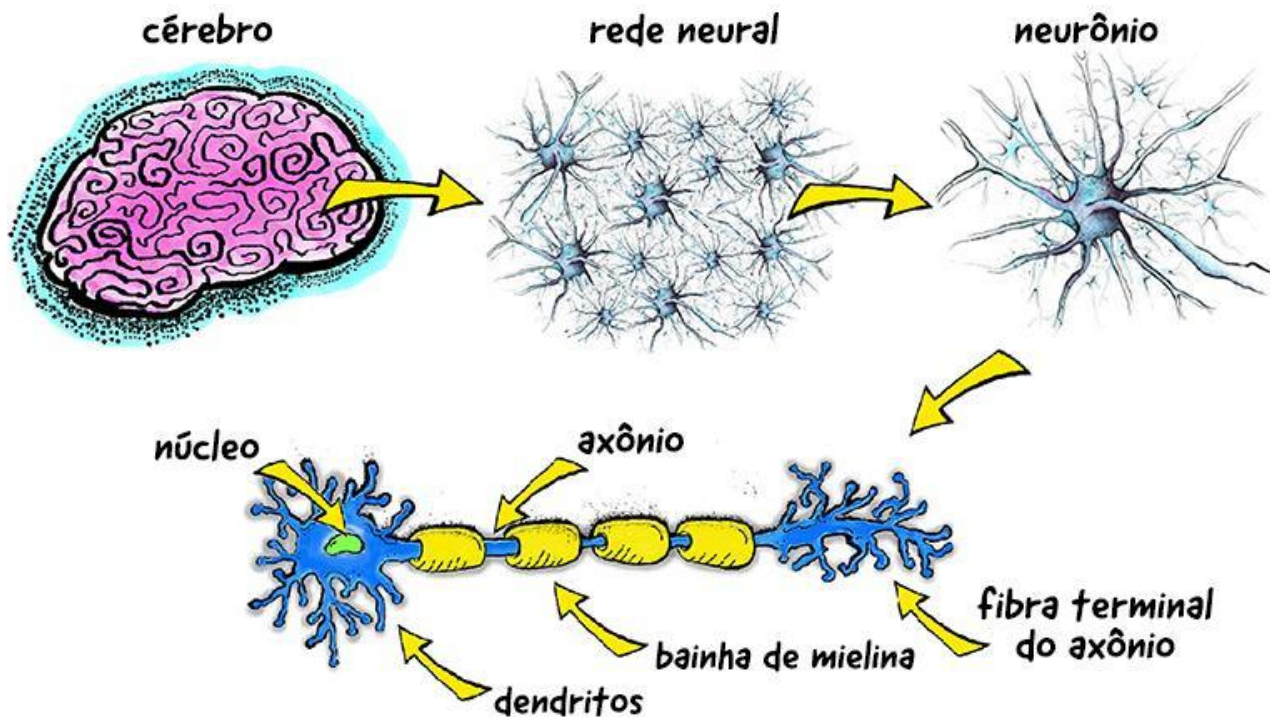
**Sistema ou algoritmos que são projetados para aprender estruturas, para prever resultados futuros**

## **Aprendizado Profundo**

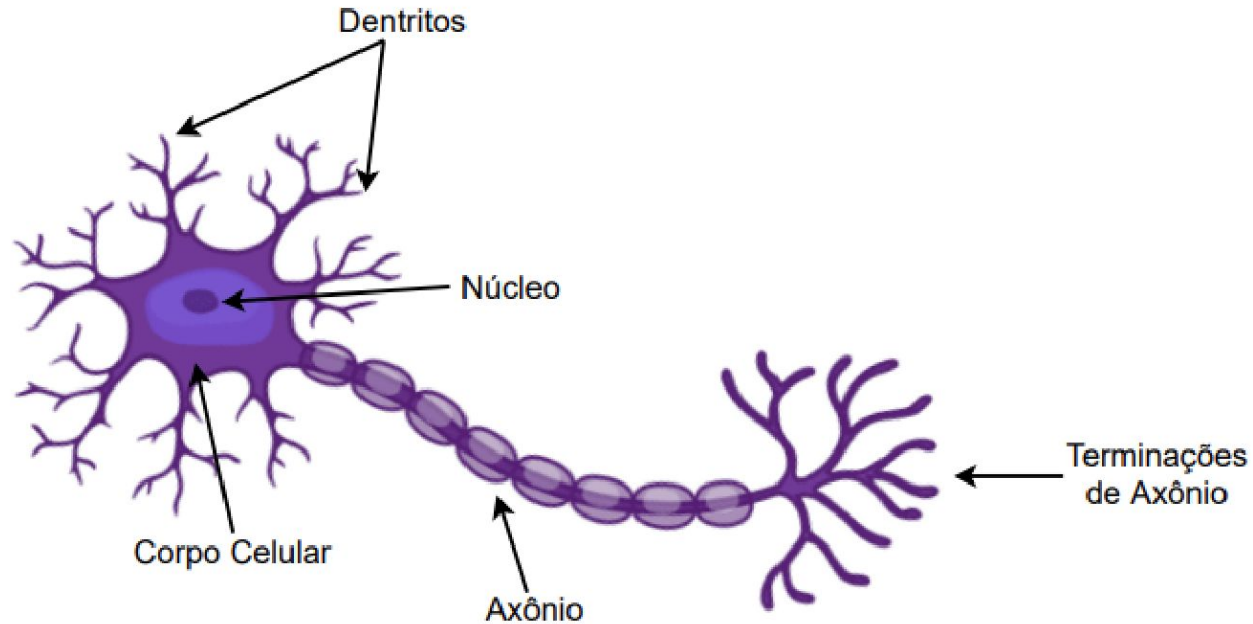
**Subcategoria de aprendizado de máquina adequada para algoritmos de autotreinamento e extração de recursos**

É baseada em sistemas  
computacionais que simulam  
a maneira como o cérebro  
humano funciona

# Representação do neurônio biológico

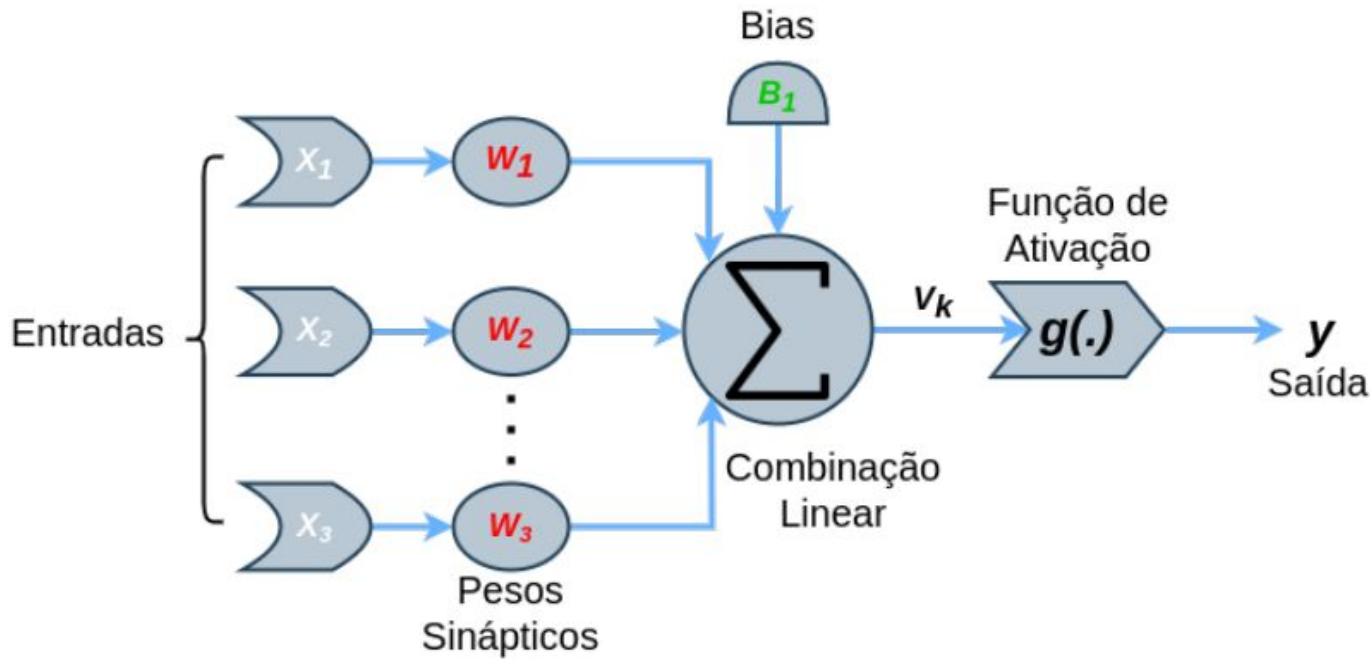


# Representação do neurônio biológico

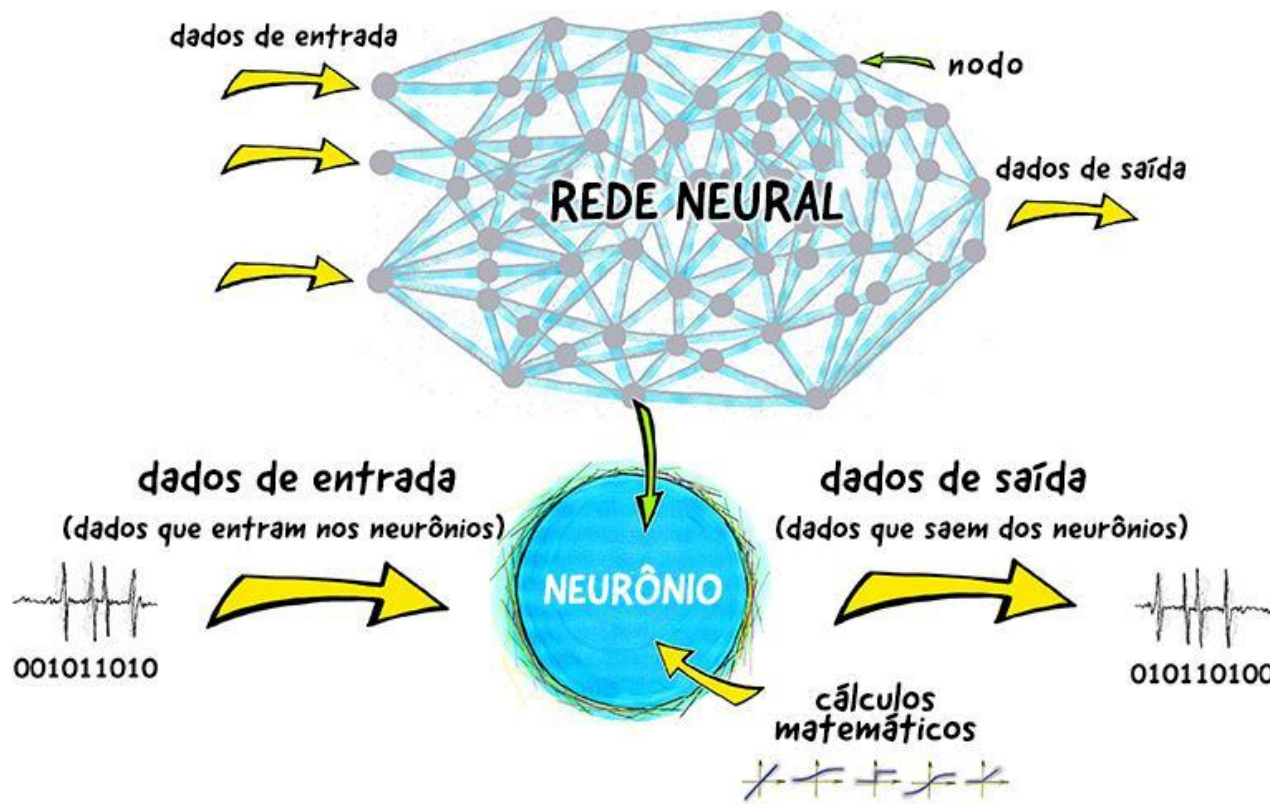




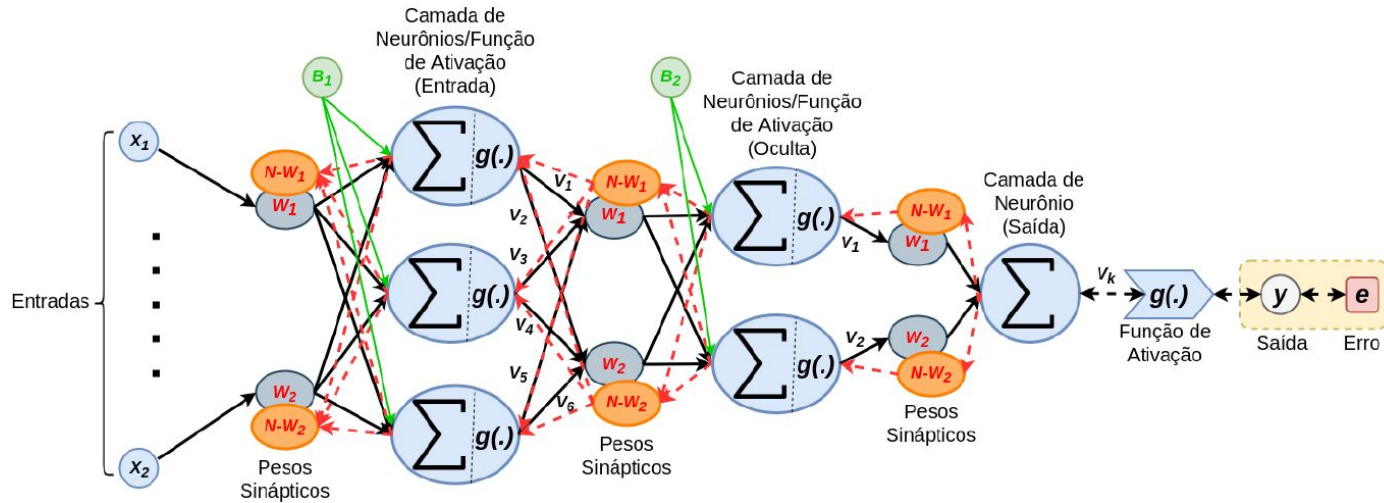
# Representação do neurônio artificial



# Representação de uma rede neural artificial



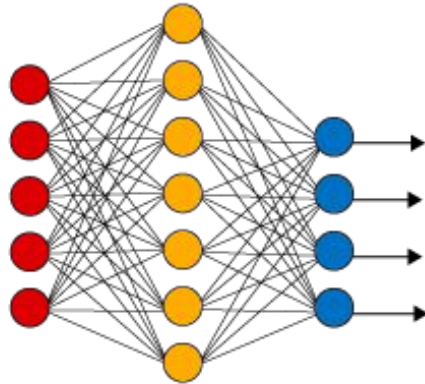
# Representação de uma rede neural artificial



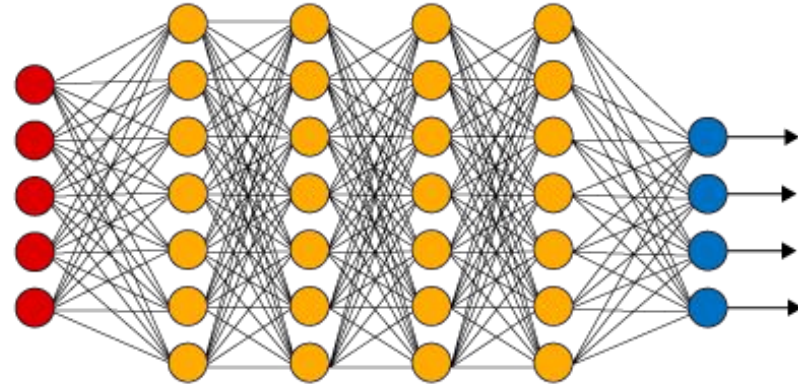
AP usa múltiplas camadas  
de neurônios  
matemáticos para processar  
dados, compreender a fala e  
reconhecer objetos visuais.

# Rede Neural Simples e Rede Neural Profunda (Deep Learning)

Simple Neural Network



Deep Learning Neural Network

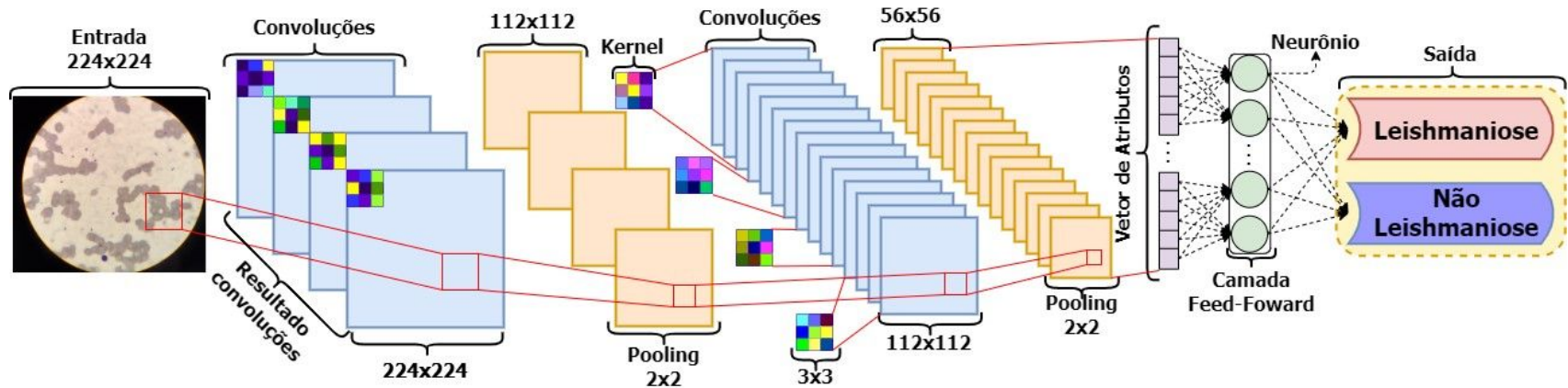


● Input Layer

● Hidden Layer

● Output Layer

# Exemplo de uma rede neural profunda aplicadas à análise de imagens.



# CNN

- Quando se trabalha com imagens envolvendo o aprendizado profundo, as redes neurais convolucionais (CNN) são amplamente utilizadas dentro da literatura.
- São tipos de redes neurais artificiais que utilizam multicamadas complexas e profundas, que são capazes de reconhecer elementos em imagens

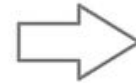
# **Como Funcionam as Redes Neurais Convolucionais?**



## Entrada da rede

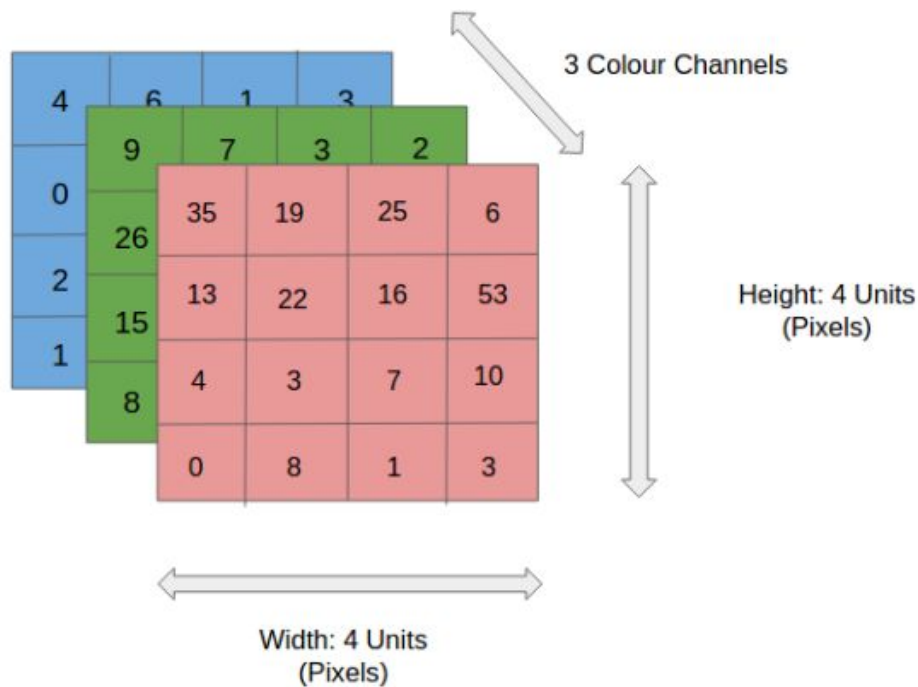
- Uma imagem nada mais é do que uma matriz de valores de pixels, certo? Então, por que não achatar a imagem (por exemplo, matriz de imagem 3x3 em um vetor 9x1) e alimentá-la em uma rede neural? Sim, é isso que fazemos.

1	1	0
4	2	1
0	2	1

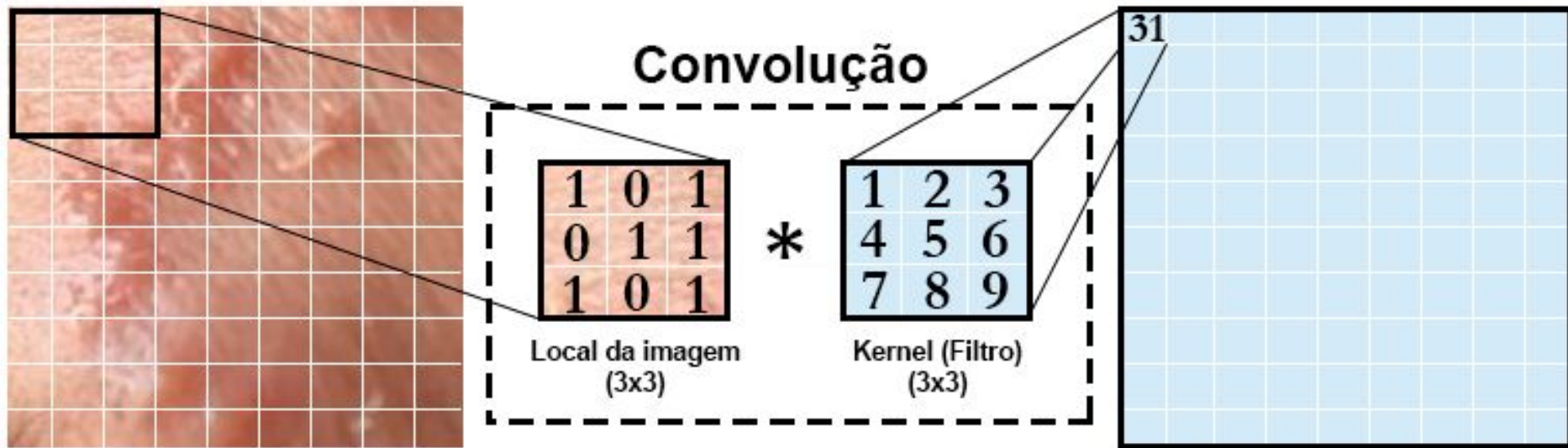


1
1
0
4
2
1
0
2
1

# Representação de uma imagem em RGB



# Convolução



# Convolução em imagem escala de cinza

1 <sub>x1</sub>	1 <sub>x0</sub>	1 <sub>x1</sub>	0	0
0 <sub>x0</sub>	1 <sub>x1</sub>	1 <sub>x0</sub>	1	0
0 <sub>x1</sub>	0 <sub>x0</sub>	1 <sub>x1</sub>	1	1
0	0	1	1	0
0	1	1	0	0

Image

4		

Convolved  
Feature

0	0	0	0	0	0	...
0	156	155	156	158	158	...
0	153	154	157	159	159	...
0	149	151	155	158	159	...
0	146	146	149	153	158	...
0	145	143	143	148	158	...
...	...	...	...	...	...	...

Input Channel #1 (Red)

0	0	0	0	0	0	...
0	167	166	167	169	169	...
0	164	165	168	170	170	...
0	160	162	166	169	170	...
0	156	156	159	163	168	...
0	155	153	153	158	168	...
...	...	...	...	...	...	...

Input Channel #2 (Green)

0	0	0	0	0	0	...
0	163	162	163	165	165	...
0	160	161	164	166	166	...
0	156	158	162	165	166	...
0	155	155	158	162	167	...
0	154	152	152	157	167	...
...	...	...	...	...	...	...

Input Channel #3 (Blue)

-1	-1	1
0	1	-1
0	1	1

Kernel Channel #1



308

+

1	0	0
1	-1	-1
1	0	-1

Kernel Channel #2



-498

+

0	1	1
0	1	0
1	-1	1

Kernel Channel #3



164

+ 1 = -25



Bias = 1

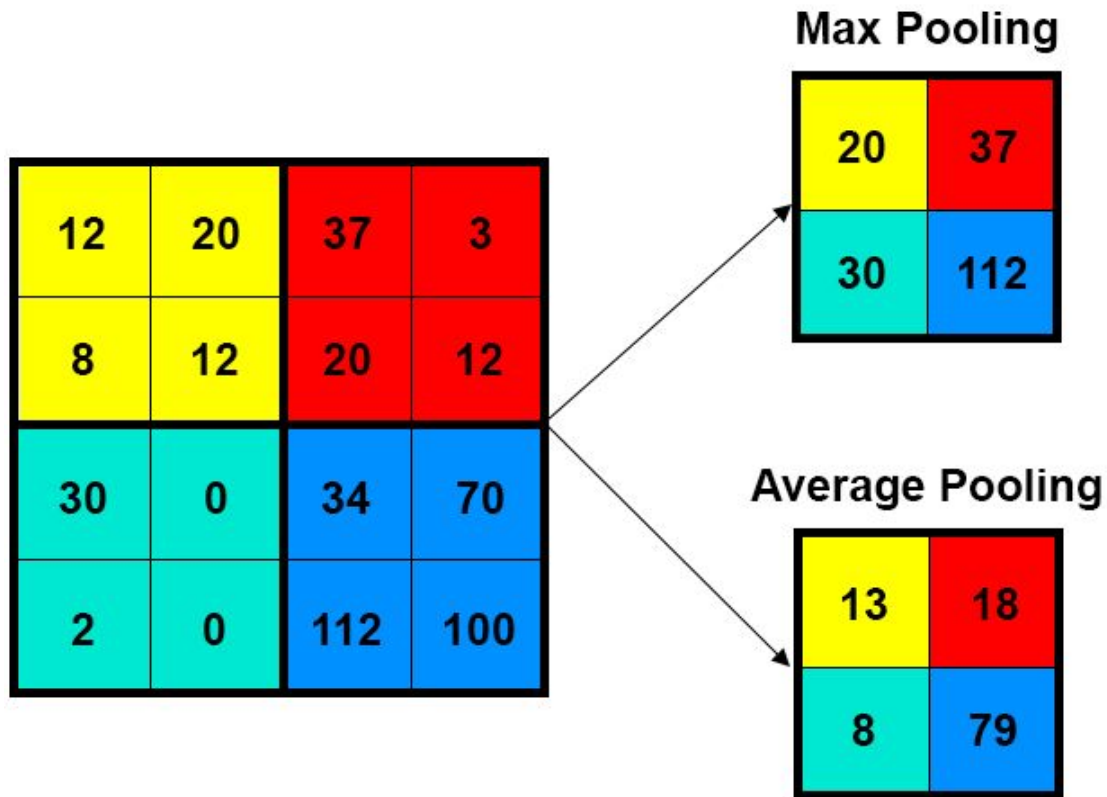
Output

-25				...
				...
				...
				...
...	...	...	...	...

## ***Pooling***

- extrair características importantes das imagens, como bordas, texturas, formas e, em camadas mais profundas, padrões mais complexos.

# Pooling

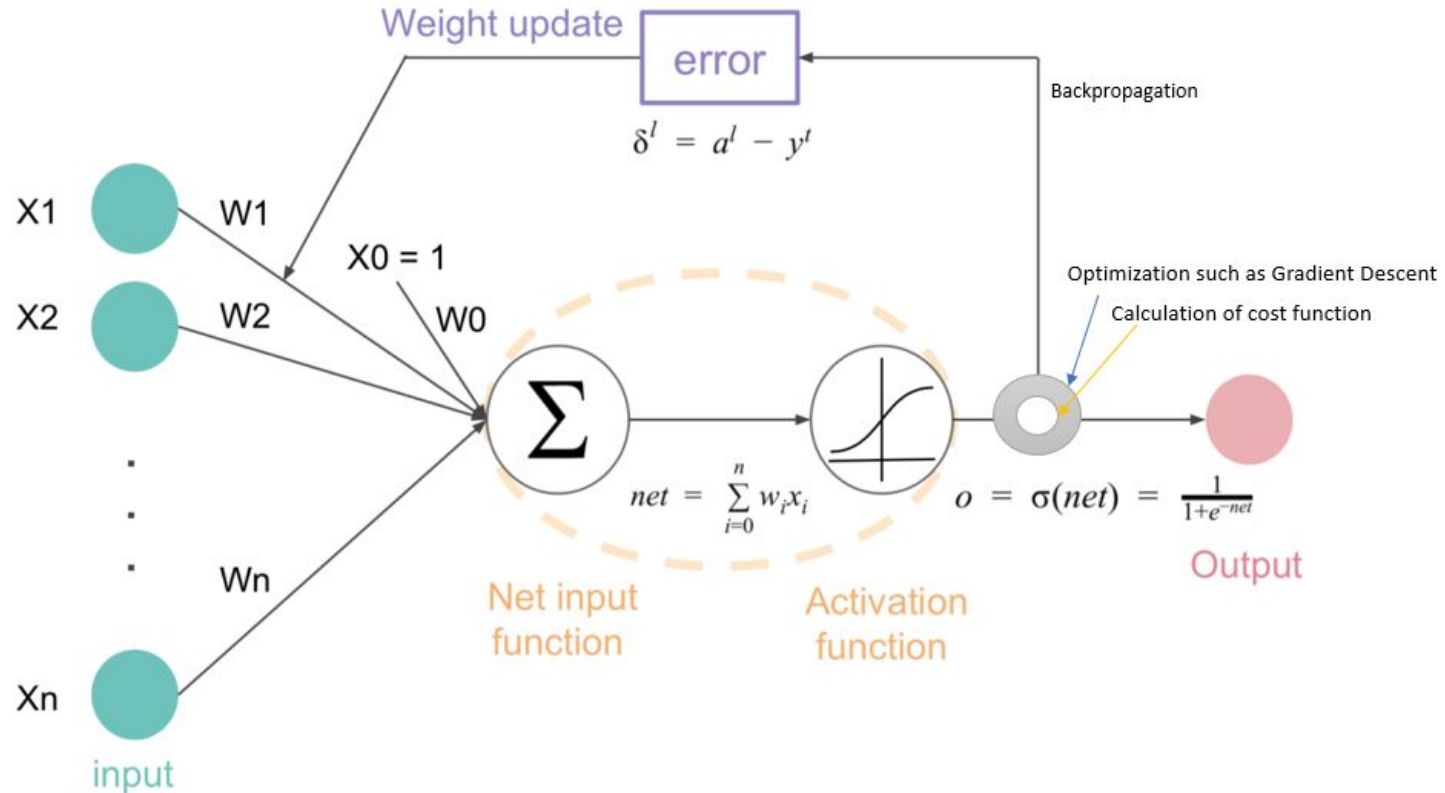


# ***Pooling***

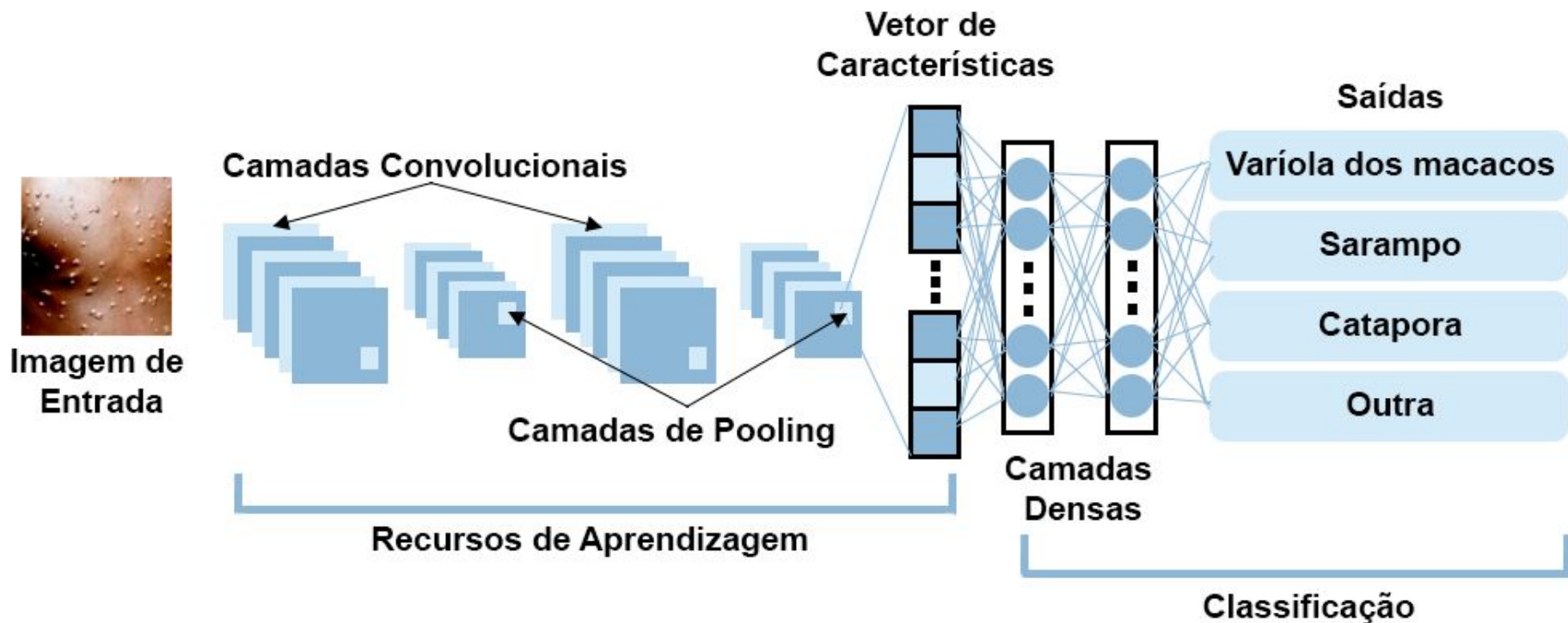
- A *Pooling*, que é responsável por reduzir a dimensionalidade da imagem (diminuindo a largura e a altura dos mapas de características).
- Benefícios:
  - diminuir o custo computacional
  - extrair características dominantes das imagens.



# Treinamento e Backpropagation



# Estrutura de uma CNN



# Transferência de Aprendizado

ooo

“Técnica de transferência de aprendizado é utilizada para extrair as características aprendidas em um problema e reutilizar esse conhecimento em um novo problema semelhante”

— Orenstein e Beijbom 2017

# Transferência de Aprendizado

- A transferência de aprendizado é normalmente utilizada quando há um conjunto de dados com poucas imagens para treinar um modelo classificação real com CNNs.

# Transferência de Aprendizado e Ajuste Fino

- É adicionado um novo classificador (camadas totalmente conectadas), que é treinado sob o modelo pré-treinado para que possa redirecionar os mapas de características aprendidos anteriormente para o novo conjunto de dados.

# Transferência de Aprendizado e Ajuste Fino

Segundo [Fchollet 2020], no contexto de DL, o guia geral para realizar a aprendizagem por transferência é o seguinte:

1. Reutilizar todas as camadas de um modelo pré-treinado (uma CNN pré-treinada, por exemplo);

# Transferência de Aprendizado e Ajuste Fino

2. Congelar todas essas camadas para evitar perder as informações que foram aprendidas durante as próximas iterações de treinamento. As camadas congeladas evitam que os pesos dos neurônios sejam atualizados;

# Transferência de Aprendizado e Ajuste Fino

3. Adicionar novas camadas totalmente conectadas sob as camadas congeladas. Com isso essas novas camadas irão transformar o aprendizado das camadas congeladas em previsões para o novo conjunto imagens;



# Transferência de Aprendizado e Ajuste Fino

4. Treinar essas novas camadas recém adicionadas no novo conjunto de imagens.

# Agora vamos ao estudo de caso!

