

Deep Learning em Imagens: aplicando CNNs com Keras e TensorFlow

Luis Vogado¹, Maíla Claro¹, Justino Santos^{1,2}, Rodrigo Veras¹

Programa de Pós-Graduação em Ciência da Computação - PPGCC

¹UFPI, ²IFPI

14 de novembro de 2019.



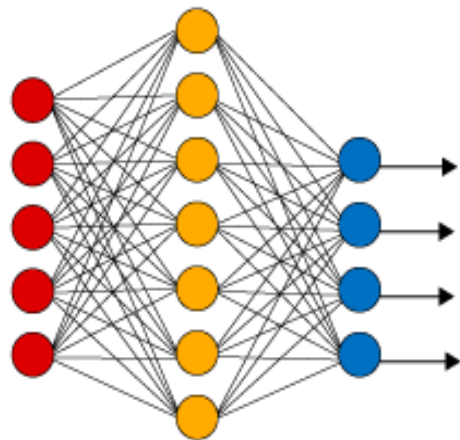
- A Inteligência Artificial (IA) é um ramo de pesquisa da ciência da computação;
- A IA se preocupa em desenvolver sistemas computacionais inteligentes, isto é, sistemas que exibem características, as quais nós associamos com a inteligência no comportamento humano.
 - Exemplo: Compreensão da linguagem, aprendizado, raciocínio, resolução de problemas, entre outros.
- Em outras palavras, são tarefas executadas de modo fáceis por pessoas, porém difíceis de serem descritas formalmente, tais como o reconhecimento da fala, padrões e imagens.

- O principal objetivo dos sistemas de inteligência artificial é executar tarefas como se um humano fosse executá-las.
 - **Capacidade de raciocínio:** aplicar regras lógicas a um conjunto de dados disponíveis para encontrar uma resposta.
 - **Aprendizagem:** aprender com um conjunto de dados disponíveis, atualizando o sistemas de erros e acertos, para no futuro agir de forma eficaz.
 - **Reconhecer padrões:** tanto padrões visuais, sensoriais como também de comportamento

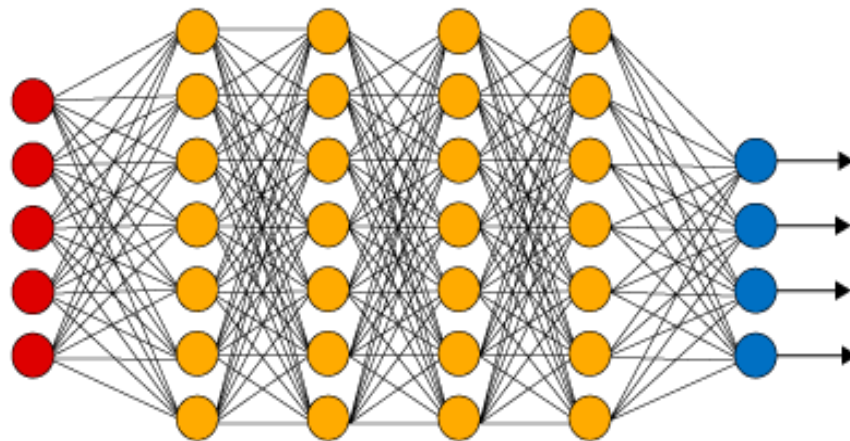
- Uma subárea da IA, que utiliza métodos computacionais para emular esse processo de intuição humana é o aprendizado de máquina (em inglês *Machine Learning* - ML);
- Em ML, a aprendizagem é feita por meio de treinamentos em banco de dados, que representam eventos e experiências passadas, possibilitando a construção de sistemas capazes de aprender de forma automática;

- As RNAs são modelos matemáticos que tentam simular algumas das estruturas neurais biológicas, possuindo capacidade computacional adquirida através do aprendizado e generalização.
- Pode-se dizer então que as RNAs são capazes de reconhecer e classificar padrões e posteriormente generalizar o conhecimento adquirido.
- Deep Learning ou aprendizagem profunda é uma técnica de aprendizado de máquina desenvolvida a partir das Redes Neurais Artificiais (RNAs).

Simple Neural Network



Deep Learning Neural Network

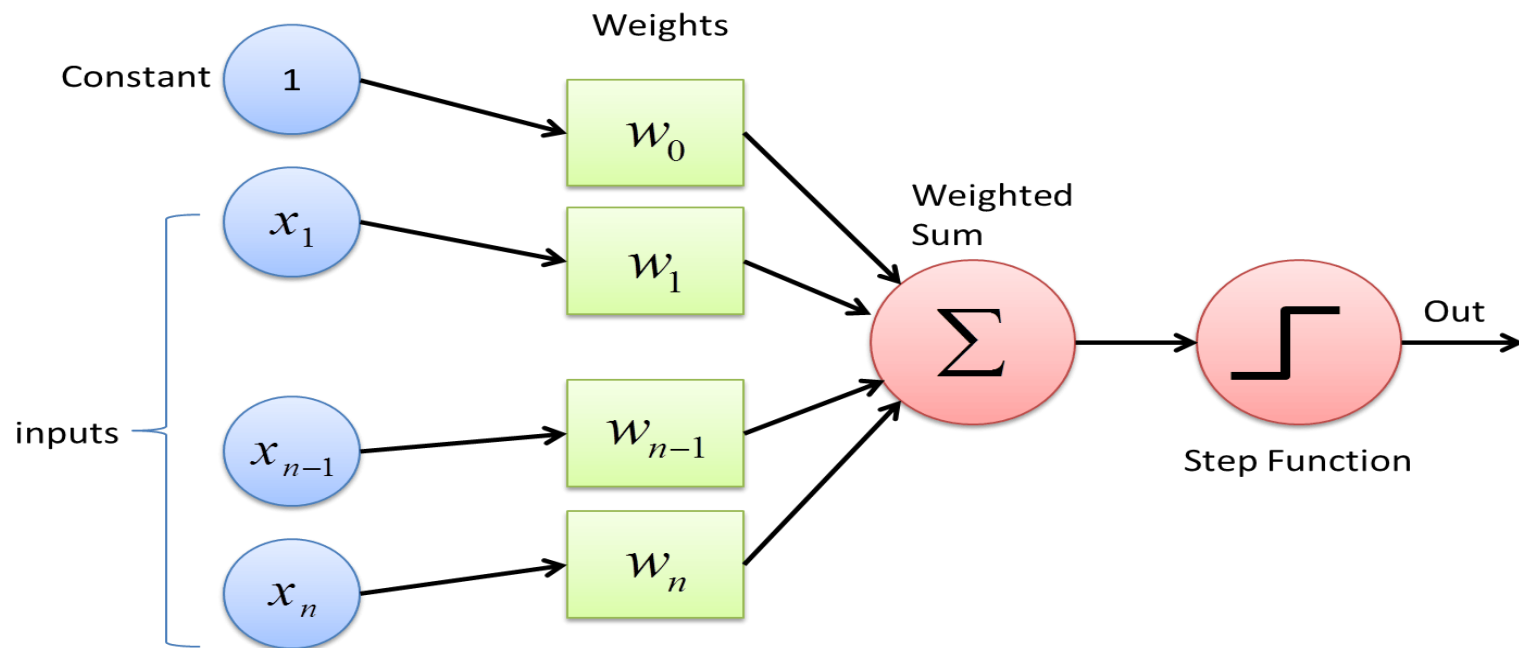


● Input Layer

● Hidden Layer

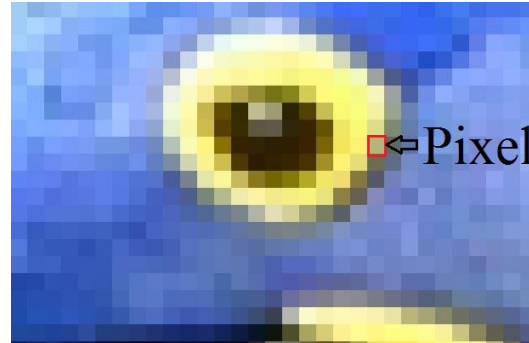
● Output Layer

Figura 1. Rede Neural Simples e Rede Neural Profunda (Deep Learning)



- As funções de ativação servem para trazer a não-linearidades ao sistema, para que a rede consiga aprender qualquer tipo de funcionalidade.
- Há muitas funções, como **Linear**, **Sigmoid**, **Tanh**, **Relu** e **Softmax**.
- A mais indicada para redes convolucionais é a Relu por ser mais eficiente computacionalmente sem grandes diferenças de acurácia quando comparada a outras funções.
- Essa função zera todos os valores negativos da saída da camada anterior.

- Uma imagem pode ser definida como sendo a representação visual de um objeto.
- Do ponto de vista matemático, uma imagem é considerada uma função bidimensional $f(x;y)$ onde x e y são coordenadas planas, e a amplitude de f em qualquer par de coordenadas $(x;y)$ é chamada de intensidade ou nível de cinza da imagem no referido ponto.



- Pixel: abreviatura de *Picture Element*.
- Menor elemento de uma imagem.
- Permite a representação pontual de uma cor
 - Combinação de três cores (**Red**, **Green**, **Blue**)
 - Em representações padrões, cada cor tem 256 tonalidades

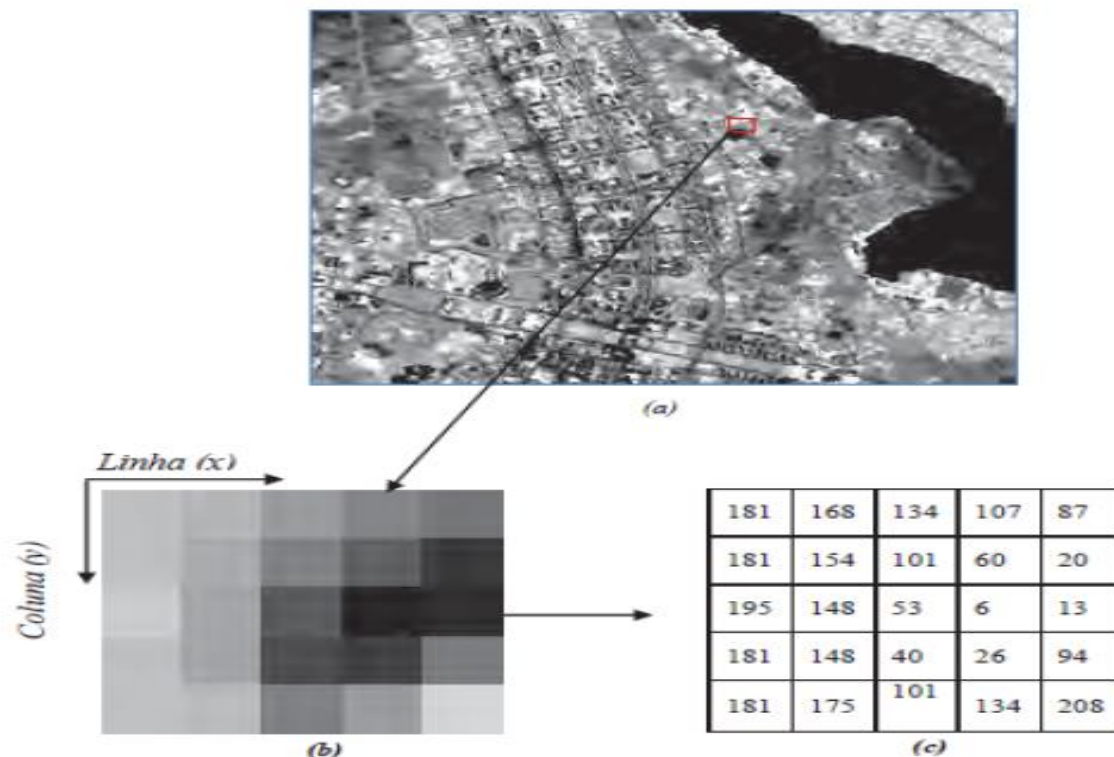


Figura 2. Imagem digital (a) com área de zoom de grupo de pixels em valores de cinza (b) e correspondentes valores digitais em (c)

- As CNNs fazem parte do conjunto de técnicas de Deep Learning. Essas redes convolucionais são uma classe de RNAs que modelam abstrações em alto nível através de imagens e camadas convolucionais dispostas de forma sequencial ou não.
- O conceito de CNN foi apresentado por Yann LeCun [Lecun et al. 1998] e Fukushima [Fukushima 1988] na década de 90. No entanto, apenas no século XXI essa tecnologia foi desenvolvida com eficácia.

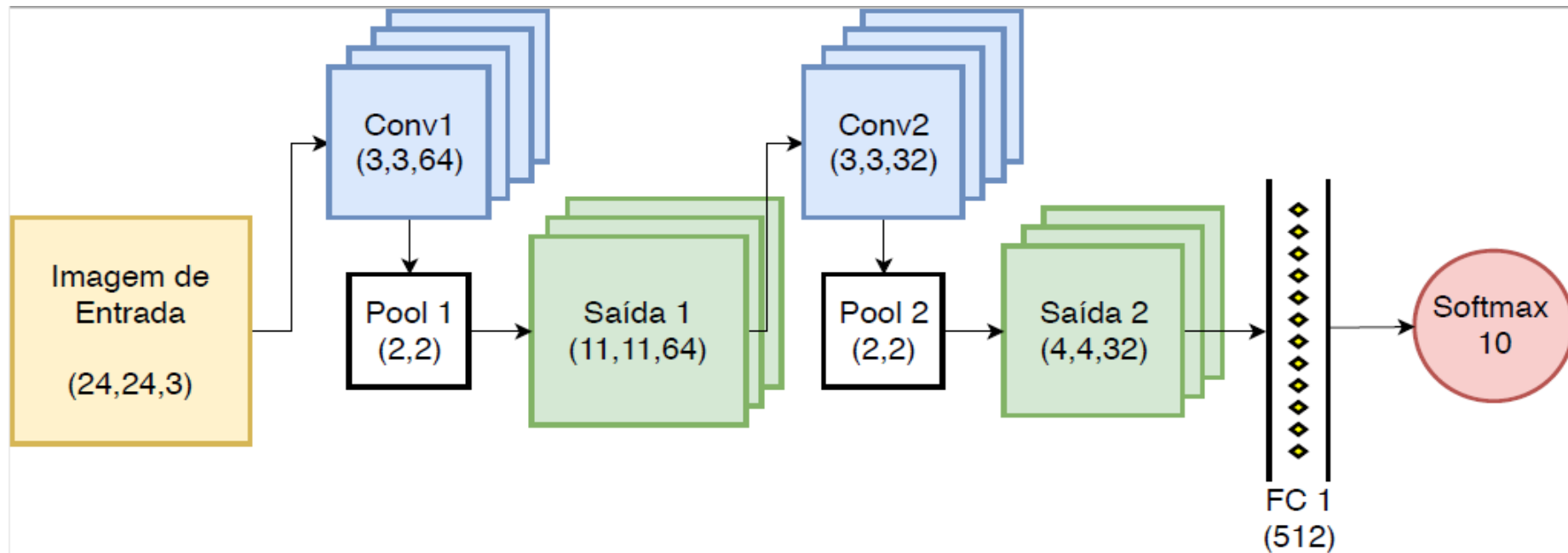


Figura 3. As CNN modernas possuem milhares ou até milhões de parâmetros para serem otimizados

1	1	0	1	0
0	1	0	0	1
1	1	1	1	0
0	1	1	0	0
1	1	1	0	1

Entrada

1	1	0
0	1	1
1	0	0

Filtro

1x1	1x1	0x0	1	0
0x0	1x1	0x1	0	1
1x1	1x0	1x0	1	0
0	1	1	0	0
1	1	1	0	1

a

1	1x1	0x1	1x0	0
0	1x0	0x1	0x1	1
1	1x1	1x0	1x0	0
0	1	1	0	0
1	1	1	0	1

b

...

4	2	3
3	4	2
5	4	3

Saída

Figura 4: Operação realizada nas camadas convolucionais.

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				...
				...
				...
				...
...

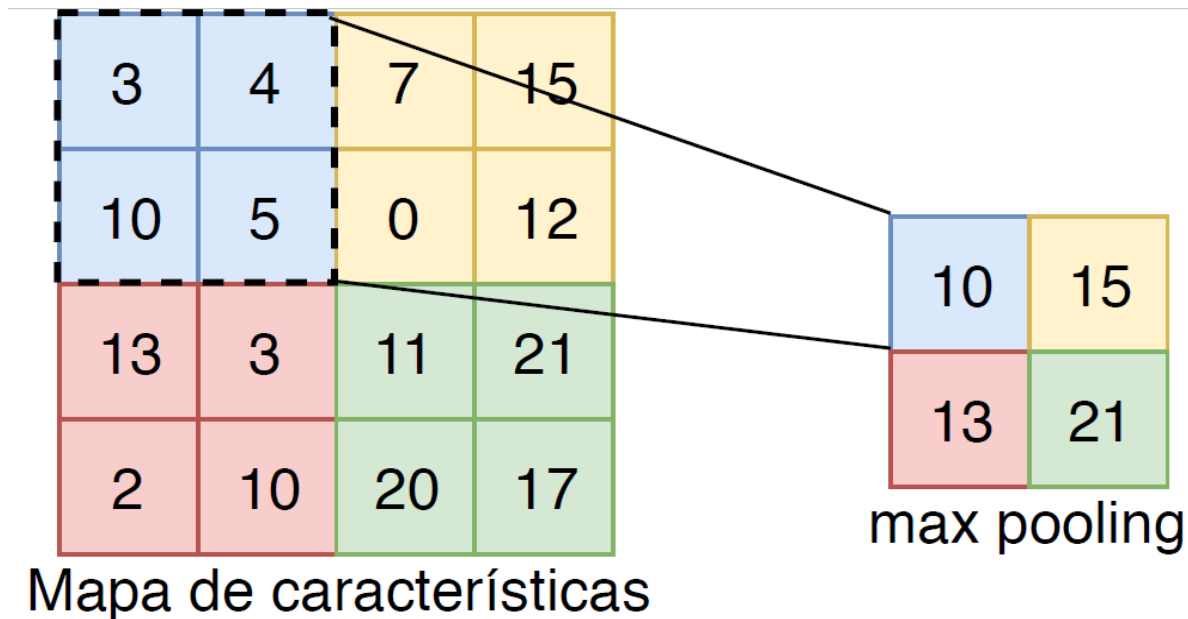
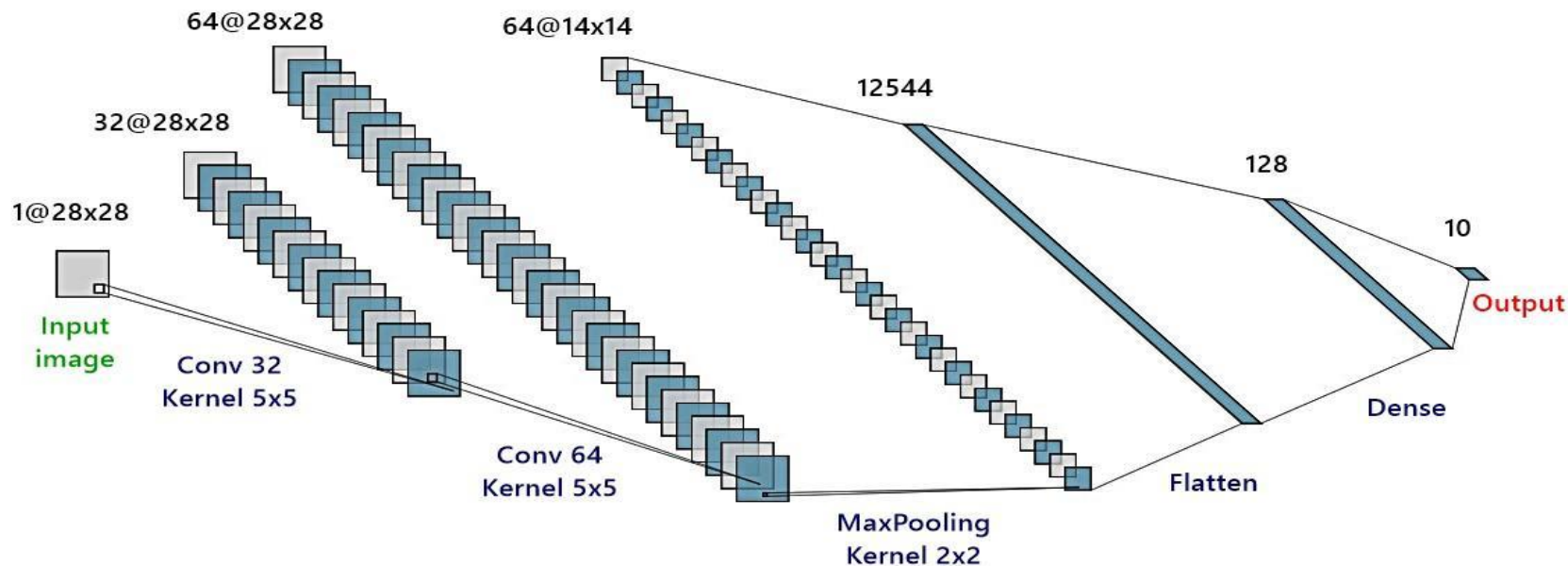


Figura 5: Operação de *maxpooling* realizada em um mapa de características aleatório.

- As camadas densas ou totalmente conectadas (*Fully-Connected layers* - FCs) foram apresentadas inicialmente nas RNAs.
- São constituídas de neurônios que representam pesos e guardam o aprendizado da rede neural.
- Nas CNN's elas apresentam a mesma função e geralmente aparecem ao final da arquitetura, após inúmeras camadas convolucionais.
- A camada de classificação presente nas CNNs é do tipo densa, no entanto, apresenta uma função de ativação do tipo softmax, enquanto as FCs contam com ativação por meio da ReLu.



- A popularidade das CNNs cresceu durante o ImageNet *Large Scale Visual Recognition Challenge* (ILSVRC) [Russakovsky et al. 2015].
- Essa competição ocorreu a partir de 2010 por 7 anos ininterruptos e teve como principais desafios a classificação de imagens e detecção de objetos em larga escala.
- Durante essa competição foram propostas várias arquiteturas, dentre elas temos a AlexNet [Krizhevsky et al. 2012], Inception ou GoogLeNet [Szegedy et al. 2015], VGGNet [Simonyan and Zisserman 2014] e ResNet [He et al. 2016].

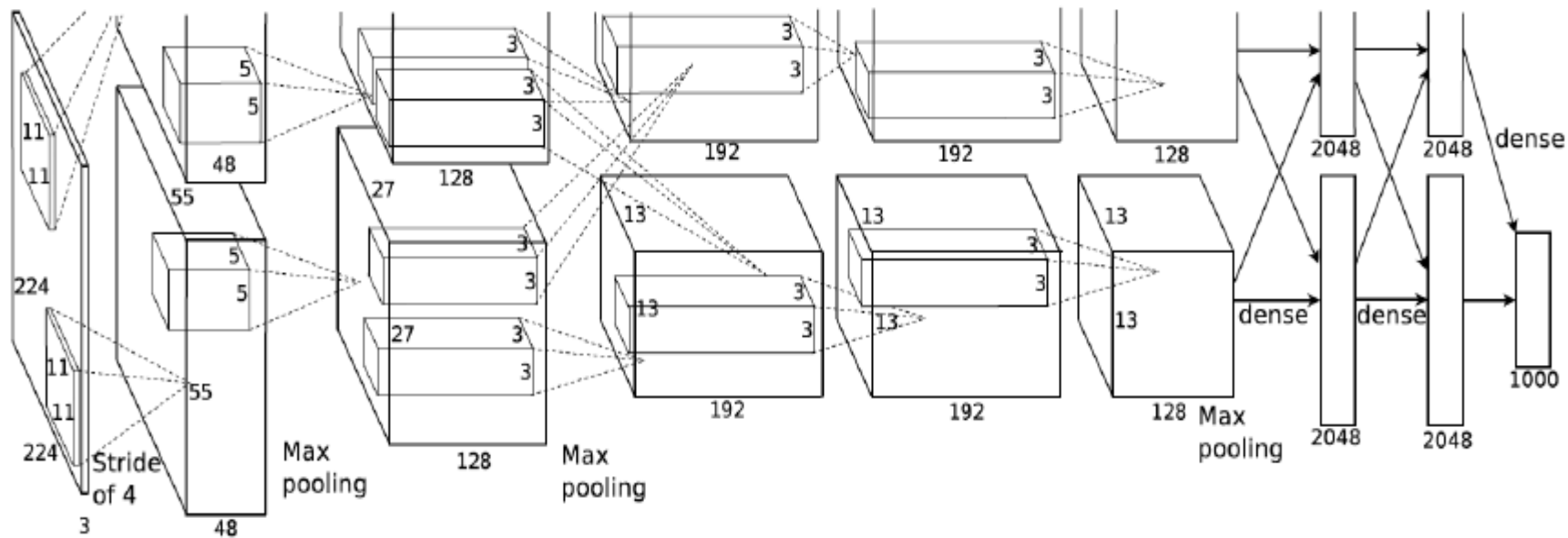


Figura 6: Ilustração da AlexNet. Fonte: Krizhevsky et al. 2012

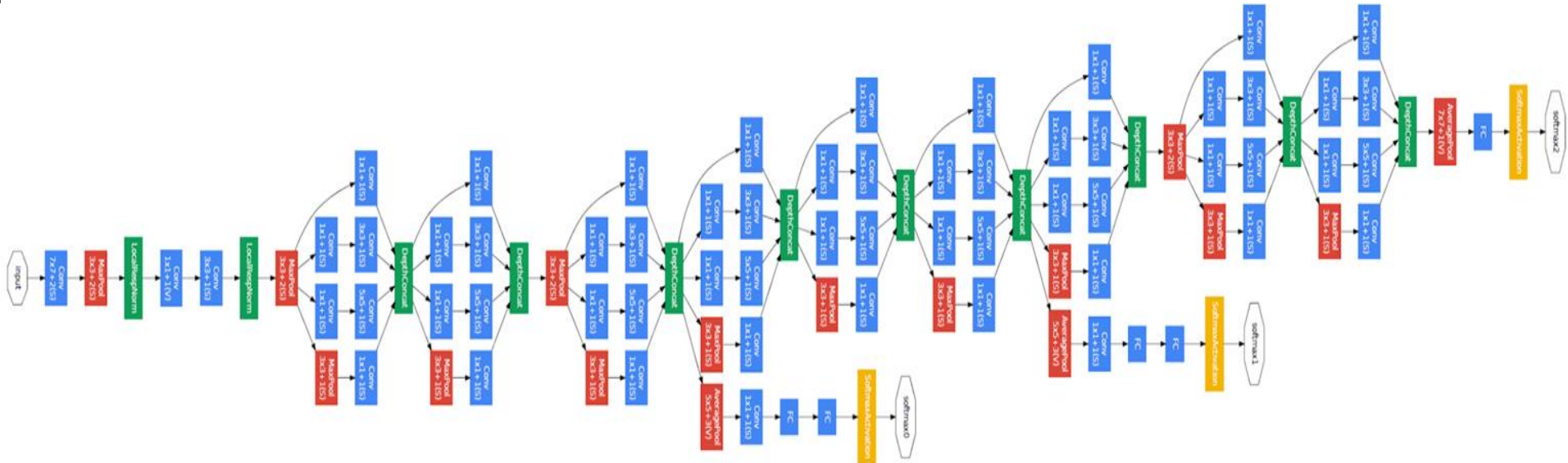


Figura 7: Ilustração da GoogLeNet. Fonte: Szegedy et al. 2015

34-layer residual

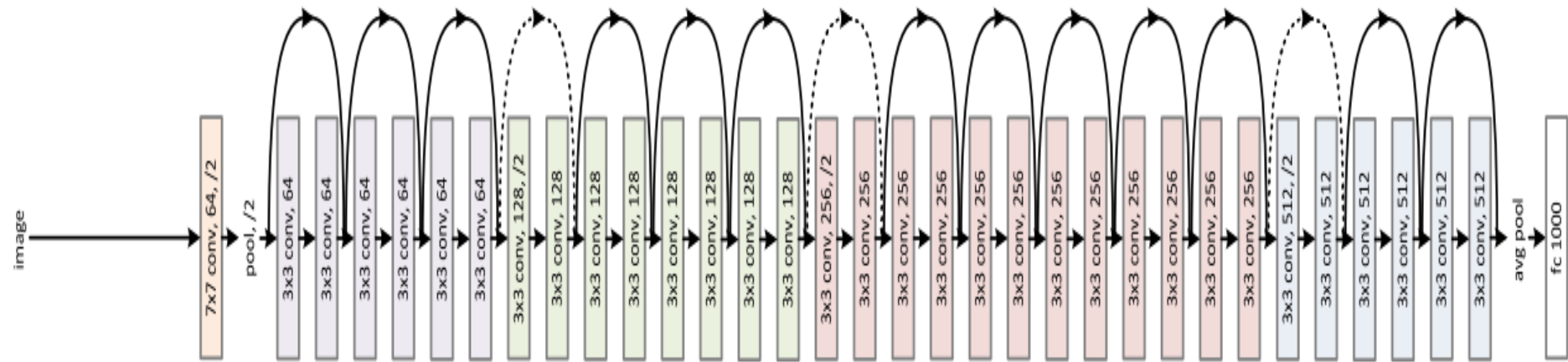


Figura 8: Residual Network com 34 camadas treináveis. Fonte: He et al. 2016

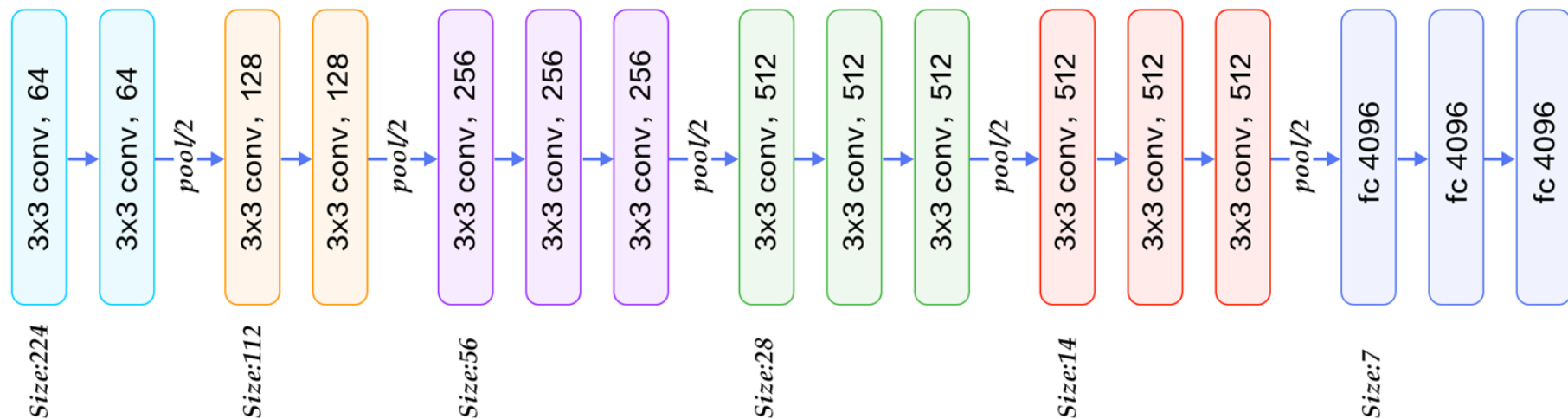


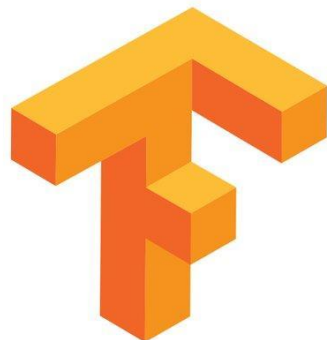
Figura 9: Ilustração da arquitetura VGGNet. Fonte: Simonyan and Zisserman 2014

- Visualização:
 - Sumário
 - Filtros
 - Mapas de características



Keras.io

Keras é uma API de alto nível para Redes Neurais Artificiais, escrita em python. capaz de executar sobre o **TensorFlow**, CNTK ou Theano.



tensorflow.org

O TensorFlow é uma plataforma de código aberto para aprendizado de máquina. Possui um ecossistema flexível e abrangente que permite que os desenvolvedores construam e implantem facilmente aplicativos com ML.

<https://colab.research.google.com/>



- Criando novo Python notebook.
- Configurar runtime
- Explorar o ambiente

- Prática I
 - Criar CNN genérica
 - Visualização da rede (*summary*)
 - Carregar base de imagens (dataset interno)
 - Treinar o modelo
 - Avaliação dos resultados

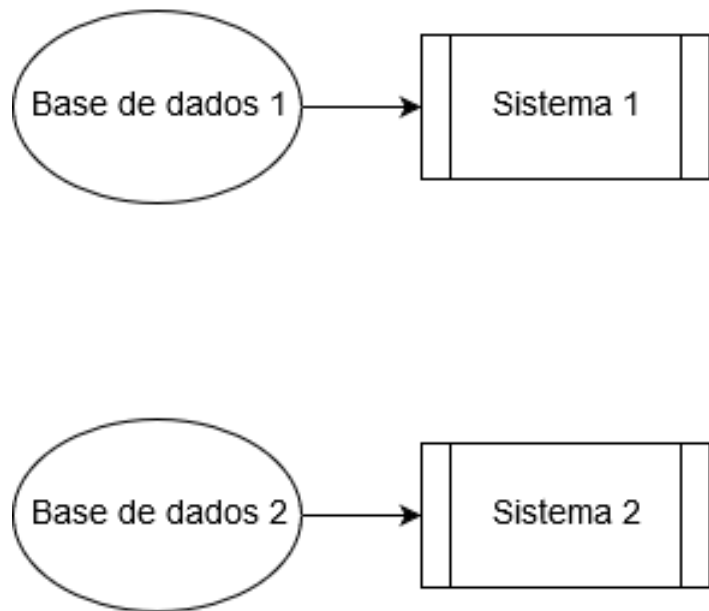
- O que foi aprendido em um cenário é explorado para facilitar a aprendizagem em outro.



- Vantagens
 - Volume de dados de treinamento
 - Tempo
 - Hardware

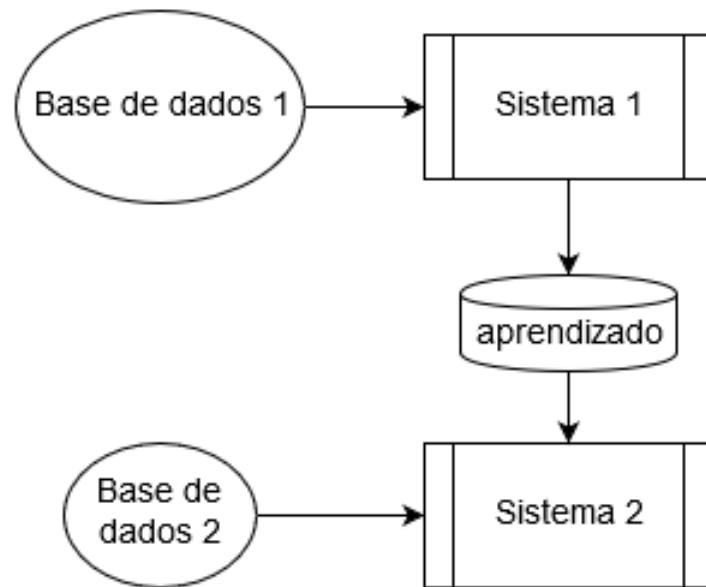
Aprendizagem de máquina tradicional:

Ocorre isoladamente. O conhecimento/aprendizado adquirido em uma tarefa não é reaproveitado em outra.

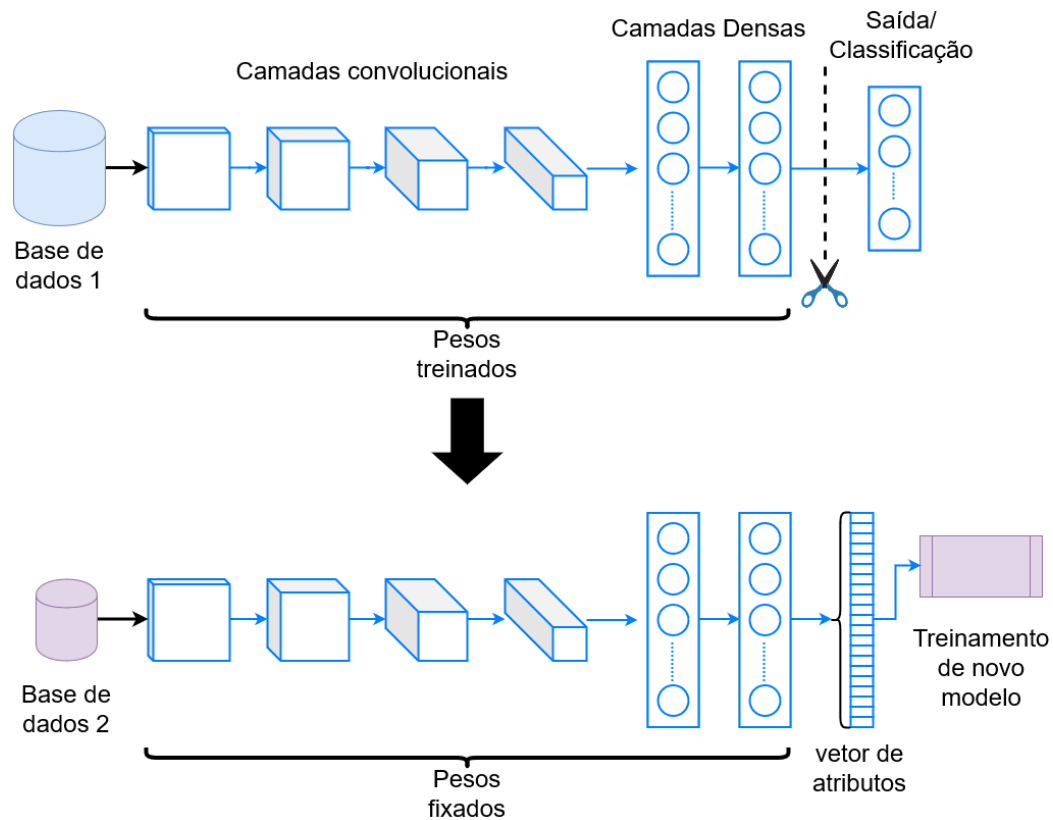


Transferência de aprendizado:

A aprendizagem em novas tarefas utiliza o conhecimento adquirido em sistema anteriormente definidos.

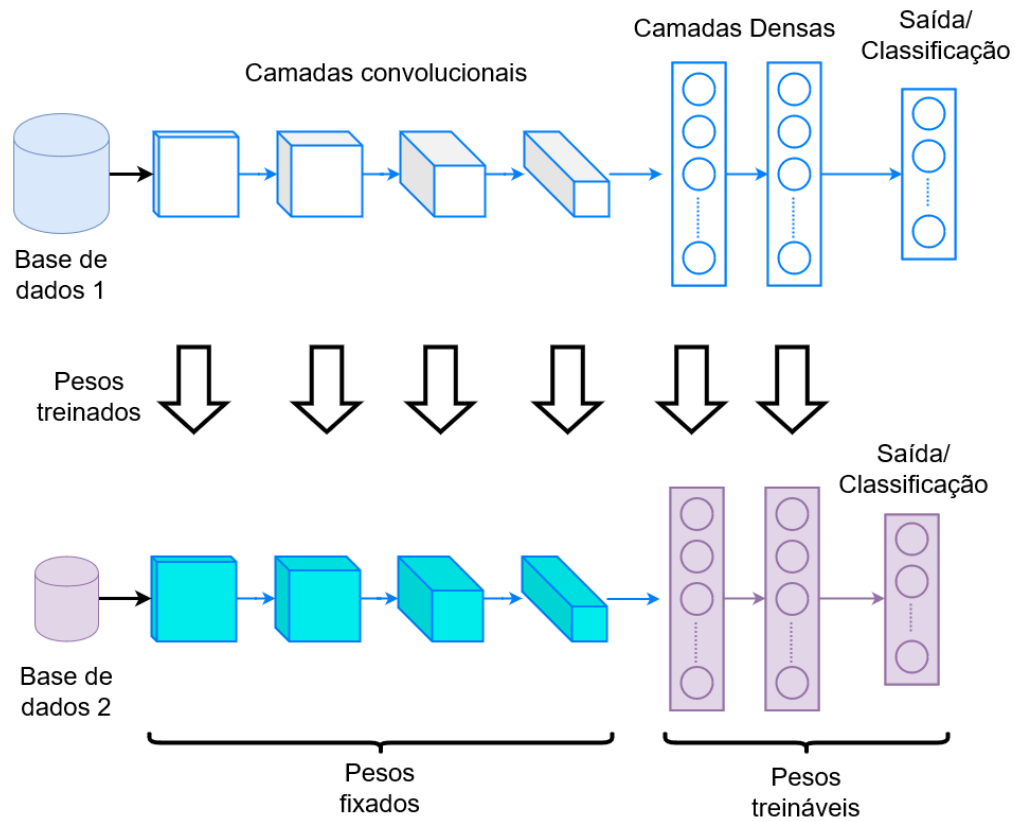


Extração de características



- Prática II (Transferência de Aprendizado)
 - Importar modelo pré treinado
 - Carregar base de imagens externa (google drive)
 - Extração de características
 - Visualização da rede
 - Análise do vetor de saída
 - Treino e teste de classificador

Ajuste Fino



- Prática III (Transferência de Aprendizado)
 - Ajuste fino
 - Visualização da rede
 - *Freeze* de camadas específicas
 - Re-treino e teste

Título: *Leukemia diagnosis in blood slides using transfer learning in CNNs and SVM for classification.*

Deep Learning em Imagens: aplicando CNNs com Keras e TensorFlow

Luis Vogado¹, Maíla Claro¹, Justino Santos^{1,2}, Rodrigo Veras¹

Programa de Pós-Graduação em Ciência da Computação - PPGCC

¹UFPI, ²IFPI

14 de novembro de 2019.

