

Redes Neurais Artificiais

*Redes Convolucionais
- Algumas Arquiteturas*



INFORMAÇÃO,
TECNOLOGIA
& INOVAÇÃO

Diferentes Arquiteturas de Redes Neurais Convolucionais

Os desempenhos das CNNs podem variar muito dependendo da arquitetura e hiperparâmetros. As melhores CNNs profundas já criadas foram fruto de intuição, alguma intuição matemática, e muita tentativa e erro

- AlexNet
- VGG
- Network in Network (NiN)
- GoogLeNet
- ResNet



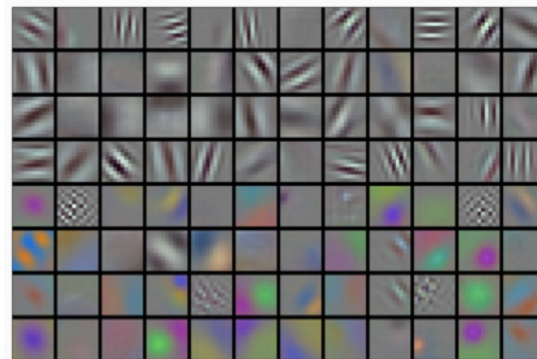
Deep Convolutional Neural Network (AlexNet)

AlexNet vem de Alex Krizhevsky, o primeiro autor do artigo propondo a nova variante de CNN

- Apresentou um aprendizado hierarquizado de características das imagens
- Primeiras camadas extraem características que lembram filtros tradicionais, e as próximas detectam estruturas mais complexas, sendo que a última aprende uma representação compacta resumando seu conteúdo, podemos separar as imagens em categorias

A AlexNet é uma CNN de 8 camadas

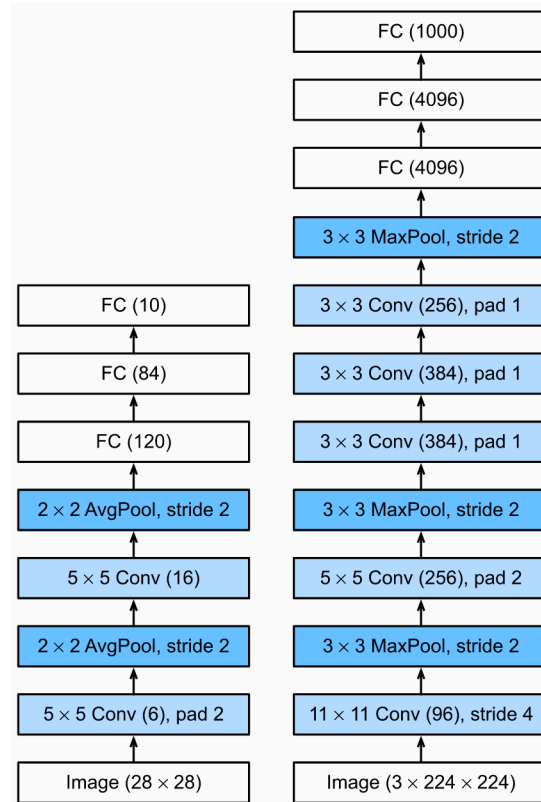
- É similar a LeNet, mas mais profunda



Deep Convolutional Neural Network (AlexNet)

AlexNet x LeNet

- Trocou a função de ativação sigmoidal por uma ReLU
- A complexidade do modelo nas camadas totalmente conectadas pode ser controlada por Dropout
- Usa data augmentation, mudando orientação e cor das imagens



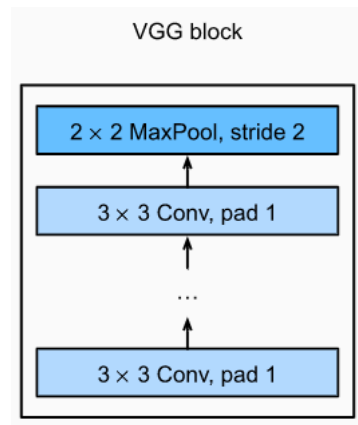
Networks Using Blocks (VGG)

Começa a ver as CNNs como conjuntos de blocos. Não pensamos mais individualmente em neurônios, mas sim em camadas, e agora blocos de camadas, repetindo padrões de camadas

- O bloco clássico de uma CNN é composto por (i) uma camada convolução com padding, (ii) uma não linearidade, como a ReLU, e (iii) e uma camada de pooling, como o max pooling.

Um bloco VGG é composto por uma sequência de camadas convolucionais seguida por uma camada de max pooling

- O artigo original propunha: convoluções 3x3 com padding=1, e uma max pooling 2x2 com stride igual a 2



AlexNet x VGG

-
- The diagram illustrates the architecture of three convolutional neural networks: AlexNet, VGG, and VGG blocks.
- AlexNet:** The architecture consists of the following layers (from bottom to top):
- 11×11 Conv (96), stride 4
 - 3×3 MaxPool, stride 2
 - 5×5 Conv (256), pad 2
 - 3×3 MaxPool, stride 2
 - 3×3 Conv (384), pad 1
 - 3×3 Conv (384), pad 1
 - 3×3 Conv (384), pad 1
 - 3×3 MaxPool, stride 2
 - 3×3 Conv (128), pad 1
 - 3×3 Conv (128), pad 1
 - 3×3 Conv (128), pad 1
 - FC (4096)
 - FC (4096)
 - FC (1000)
- VGG:** The architecture consists of the following layers (from bottom to top):
- 3×3 Conv (64), pad 1
 - 3×3 Conv (64), pad 1
 - 3×3 Conv (128), pad 1
 - 3×3 Conv (128), pad 1
 - 3×3 Conv (256), pad 1
 - 3×3 Conv (256), pad 1
 - 3×3 Conv (512), pad 1
 - 3×3 Conv (512), pad 1
 - 3×3 MaxPool, stride 2
 - 3×3 MaxPool, stride 2
 - FC (4096)
 - FC (4096)
 - FC (1000)
- VGG block:** The VGG block consists of the following layers (from bottom to top):
- 3×3 Conv (64), pad 1
 - 3×3 Conv (64), pad 1
 - 2×2 MaxPool, stride 2

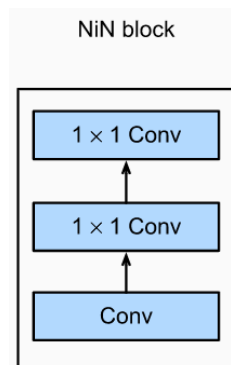
Network in Network (NiN)

LeNet, AlexNet e VGG todas extraem características via uma sequência de camadas de convolução e pooling, e depois um pós-processamento via camadas totalmente conectadas. E se utilizarmos camadas totalmente conectadas no meio do processo?

- Temos que tomar cuidado para não destruir a estrutura espacial da representação. O uso de camadas totalmente conectadas pode ocasionar isso
- Network in Network (NiN) são uma alternativa: usar uma MLP em cada pixel separadamente

Podemos pensar na NiN como uma camada de convolução 1×1

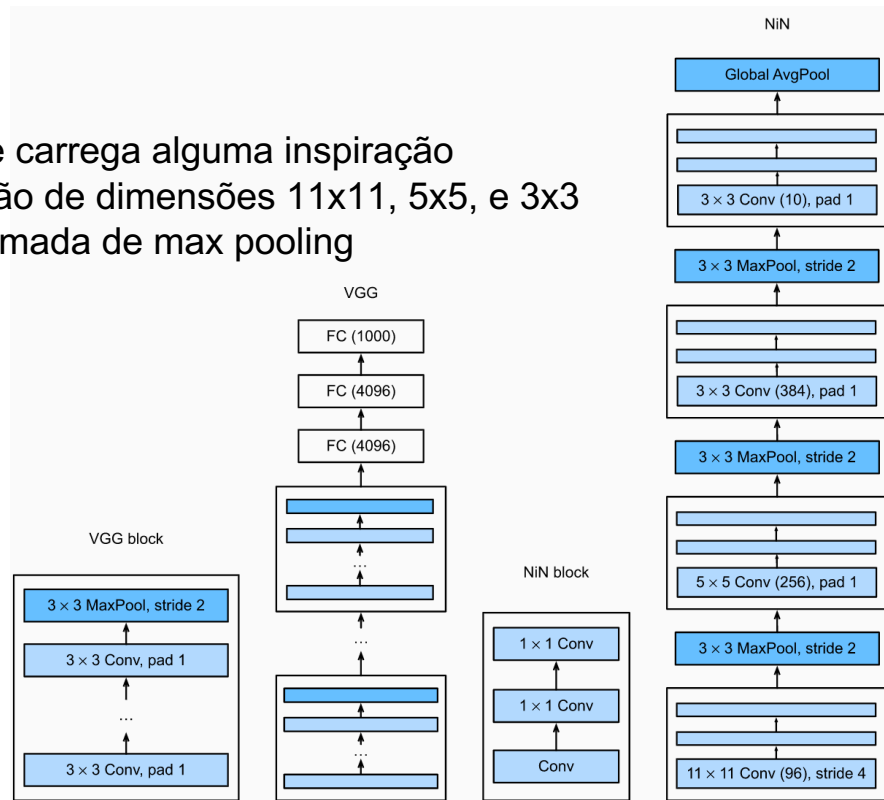
- Uma camada totalmente conectada atuando de forma independente em cada pixel
- O bloco NiN consiste de uma camada de convolução seguida de duas camadas de convolução 1×1 , agindo como camadas totalmente conectadas por pixel e função de ativação ReLU



Network in Network (NiN)

Arquitetura da Rede NiN

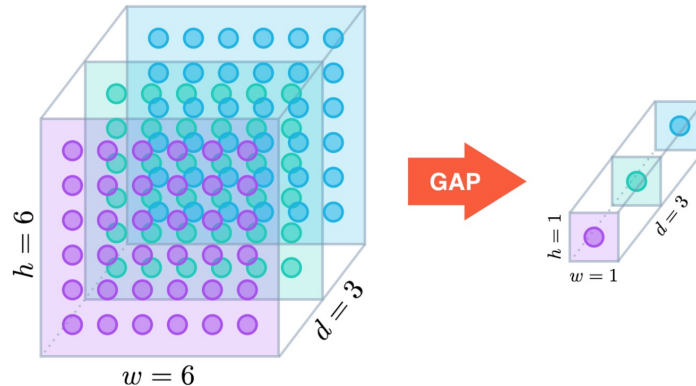
- A rede NiN veio logo após a AlexNet e carrega alguma inspiração
- A rede NiN usa camadas de convolução de dimensões 11x11, 5x5, e 3x3
- Cada bloco NiN é seguido por uma camada de max pooling
- Troca as camadas totalmente conectadas da saída por um global average pooling
- Essa remoção reduz muito o número de parâmetros



Network in Network (NiN)

Global Average Pooling

- Como se tivéssemos um feature map para cada categoria
- Ao invés de adicionar camadas totalmente conectadas em cima dos feature maps, é feita feita a média de cada feature map
- O vetor resultante é dado como entrada para uma camada softmax



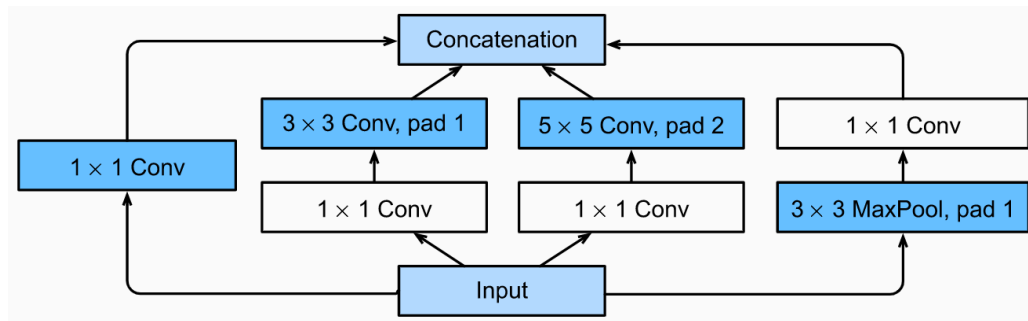
Networks with Parallel Concatenations (GoogLeNet)

Combina os benefícios da NiN com o paradigma de blocos repetidos. No artigo original, os autores mostraram que pode ser vantajoso combinar kernels de diferentes tamanhos

- O bloco básico da GoogLeNet é chamada Inception

A combinação de filtros explora a imagem de várias maneiras.

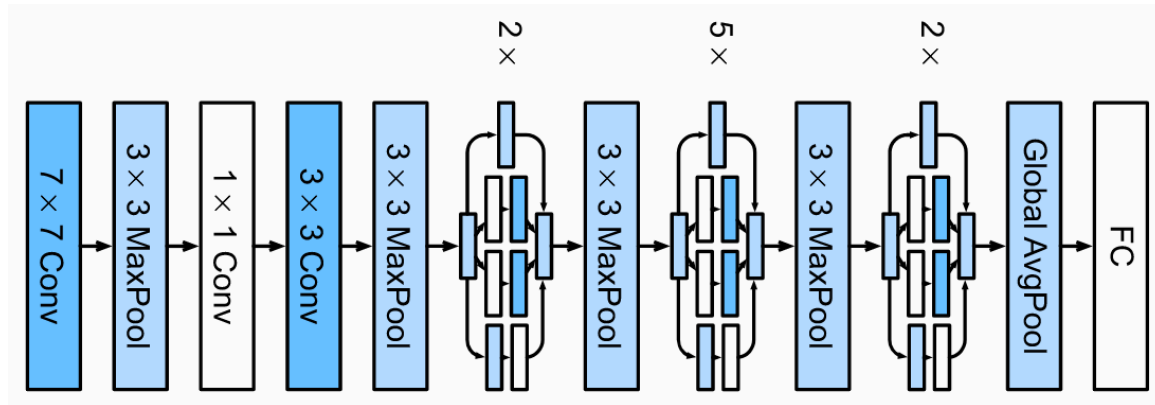
- Diferentes detalhes podem ser capturados por filtros de diferentes tamanhos
- Cada filtro tem um número diferente de parâmetros



Networks with Parallel Concatenations (GoogLeNet)

O modelo empilha 9 blocos Inception com um global average pooling no final

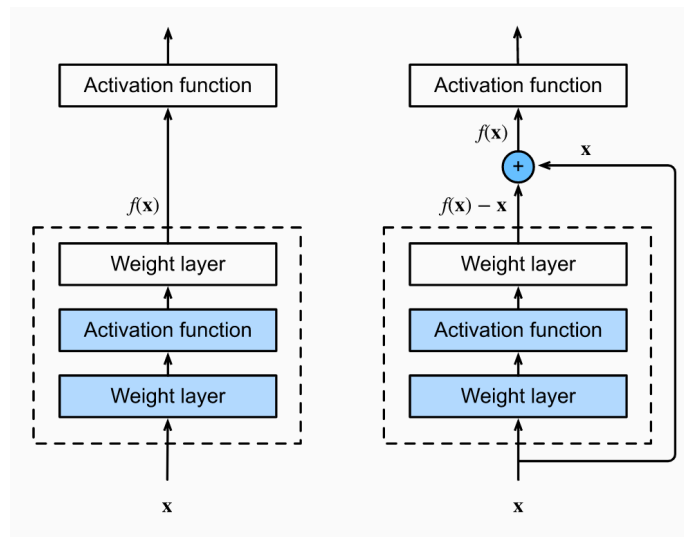
- O bloco Inception é equivalente a uma subrede com 4 caminhos
- Os blocos são conectados em série



Residual Networks (ResNet)

A ideia é que cada camada adicional deve conter a função identidade como um de seus elementos. Isso leva a um bloco residual

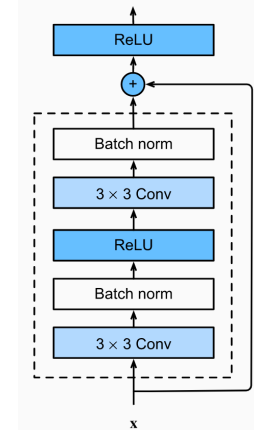
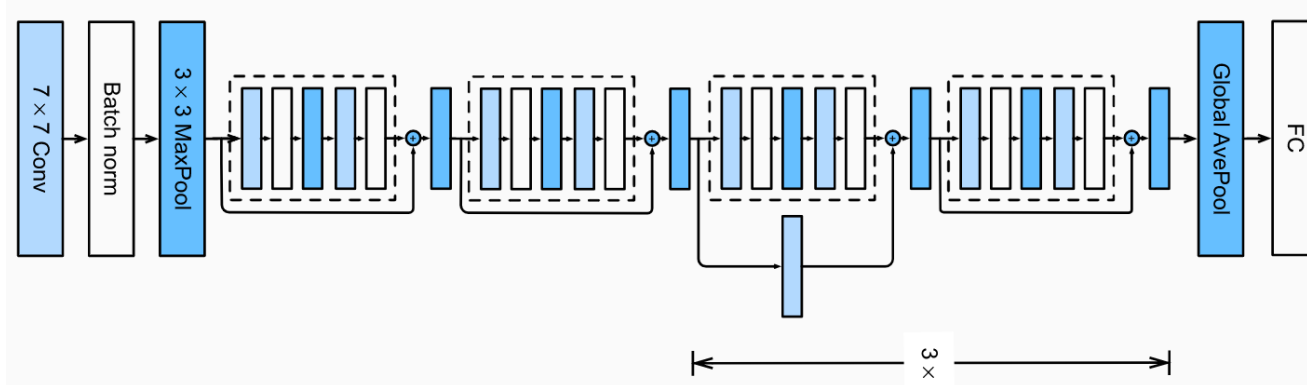
- Os blocos residuais permitem que informações (memória) fluam de uma camada inicial até as próximas camadas
- Ajudar na construção de arquiteturas mais profundas
- Conforme as camadas ficam mais profundas, a função identidade ajuda a reduzir o erro da rede
- São skip connections
- É como decompor uma função f em um termo linear simples e uma não linearidade mais complexa



Residual Networks (ResNet)

A ideia é que cada camada adicional deve conter a função identidade como um de seus elementos. Isso leva a um bloco residual

- Os blocos residuais permitem que informações (memória) fluam de uma camada inicial até as próximas camadas



Perguntas?

