

14 - Prática: Redes Neurais Convolucionais 1 (Deep Learning) (II)

Thassiana C. A. Muller

Introdução

Redes neurais convolucionais (CNN) são amplamente utilizadas para trabalhar com reconhecimento de imagens. A seguir, serão expostos alguns dos principais conceitos para se trabalhar com essa técnica.

Pixel

O pixel, no padrão de cores RGB, possui 3 valores característicos que variam de 0 a 255 e definem a combinação total de cores possíveis.

Redes Neurais Convolucionais (CNN)

De forma parecida com Redes Neurais Artificiais Densas(ANNs), as CNN's recebem atributos previsores que, através de uma calibração de pesos, permitem que seja feita a classificação de novas entradas. No entanto, ao contrário das ANNs, as CNNs utilizam operações convolucionais e outras técnicas e escolhem apenas os atributos que diferenciam as imagens a serem classificadas de forma a evitar que todos os pixels da imagem se tornem atributos previsores, identificando padrões relevantes e aumentando acurácia

Convolução

A convolução é uma operação em que uma matriz menor, chamada de kernel, filtro ou máscara é aplicada a uma matriz maior de pixels. Ela envolve soma da multiplicação elemento a elemento entre o kernel e uma região da matriz da imagem de entrada. Esse processo é repetido por toda a imagem, gerando uma nova matriz chamada mapa de características que contém os pixels mais relevantes para diferenciar as imagens.

É definida por:

$$g(x, y) = \omega * f(x, y) = \sum_{i=-a}^a \sum_{j=-b}^b \omega(i, j) f(x - i, y - j),$$

Onde $g(x,y)$ são as coordenadas para os pixels da imagem de filtrada, ω é o Kernel, (i, j) são as coordenadas da matriz do Kernel e $f(x,y)$ são as coordenadas para os pixels da imagem de entrada

Kernel

O kernel é a matriz que multiplicará a imagem original e resalta algumas características da imagem original.

Operation	Kernel ω	Image result $g(x,y)$
Identity	$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$	
Ridge or edge detection	$\begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix}$	
	$\begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix}$	
Sharpen	$\begin{bmatrix} 0 & -1 & 0 \\ -1 & 5 & -1 \\ 0 & -1 & 0 \end{bmatrix}$	
Box blur (normalized)	$\frac{1}{9} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$	
Gaussian blur 3 × 3 (approximation)	$\frac{1}{16} \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}$	
Gaussian blur 5 × 5 (approximation)	$\frac{1}{256} \begin{bmatrix} 1 & 4 & 6 & 4 & 1 \\ 4 & 16 & 24 & 16 & 4 \\ 6 & 24 & 36 & 24 & 6 \\ 4 & 16 & 24 & 16 & 4 \\ 1 & 4 & 6 & 4 & 1 \end{bmatrix}$	
Unsharp masking 5 × 5 Based on Gaussian blur with amount as 1 and threshold as 0 (with no image mask)	$\frac{-1}{256} \begin{bmatrix} 1 & 4 & 6 & 4 & 1 \\ 4 & 16 & 24 & 16 & 4 \\ 6 & 24 & -476 & 24 & 6 \\ 4 & 16 & 24 & 16 & 4 \\ 1 & 4 & 6 & 4 & 1 \end{bmatrix}$	

(1)

Durante o treinamento da CNN são gerados filtros aleatórios para melhor detectar as características

Mapa de características

É o resultado da convolução da matriz de entrada (por exemplo uma imagem) com a matriz filtro ou kernel, ele contém os principais atributos que caracterizam a entrada.

Durante o treinamento, a rede irá gerar várias convoluções a fim de obter a mais relevante possível. Assim, a camada de convolução é um conjunto de mapas de características

Após a obtenção do mapa de características, é aplicado ponto a ponto a função ReLU, que espelha os valores positivos.

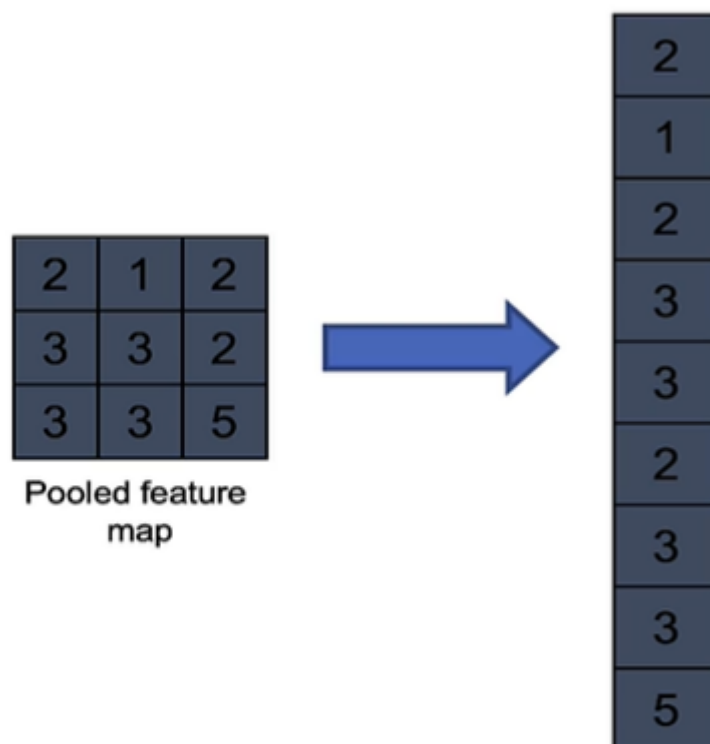
Pooling

Também é uma técnica de redução de dimensionalidade e possibilita que a rede neural se adapte às mudanças de contexto da imagem original.

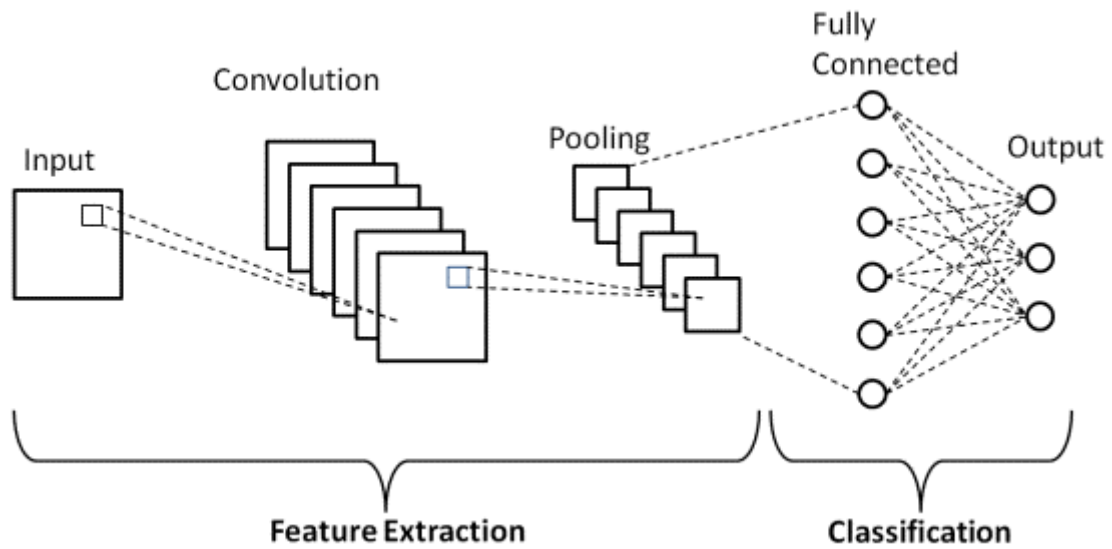
A técnica age sobre o mapa de características, por exemplo, selecionando os maiores valores regionais. Dessa forma, obtêm-se as principais características da matriz de entrada que irão diferenciar uma classe da outra.

Flattening

Refere-se a vetorizar a matriz resultante do Pooling, de forma a reduzir os dados a somente uma dimensão. Assim, os dados já poderão ser atributos previsores de uma Rede Neural Densa.



Assim, todo o processo fica



(2)

Referências

- (1) WIKIPEDIA CONTRIBUTORS. Kernel (image processing). Disponível em: <[https://en.wikipedia.org/wiki/Kernel_\(image_processing\)](https://en.wikipedia.org/wiki/Kernel_(image_processing))>.
- (2) Basic Introduction to Convolutional Neural Network in Deep Learning. Disponível em: <<https://www.analyticsvidhya.com/blog/2022/03/basic-introduction-to-convolutional-neural-network-in-deep-learning/>>.