

Equações de Convolução para Estruturas Tridimensionais

A convolução é uma operação fundamental em redes neurais convulsionais, amplamente utilizada no processamento de imagens e vídeos. As equações a seguir descrevem como os valores dos mapas de características são calculados a partir das entradas e dos kernels em duas estruturas tridimensionais distintas.

1) Equações de Convolução para Dados Multicanais (Exemplo: Imagem RGB)

Para dados multicanais, como uma imagem RGB, a convolução é realizada somando as convoluções de cada canal individualmente. Suponha que temos uma entrada I com K canais, cada um de tamanho $n \times n$, e um kernel K de tamanho $k \times h \times h$. A equação de convolução para a posição (i, j) no mapa de características O é dada por:

$$O(i, j) = \sum_{c=1}^k \sum_{u=1}^h \sum_{v=1}^h I_c(i + u, j + v) \cdot K_c(u, v)$$

Nesta equação, a saída em um ponto específico do mapa de características é calculada a partir da soma ponderada dos valores dos canais de entrada, multiplicados pelos valores correspondentes do kernel.

2) Equações de Convolução para Dados Tridimensionais (Exemplo: Vídeos ou Imagens de Ressonância Magnética)

Para dados tridimensionais, como vídeos ou imagens de ressonância magnética (MRI), a convolução é realizada em três dimensões. Suponha que temos uma entrada I com dimensões $d \times n \times n$ e um kernel K de tamanho $h \times m \times m$. A equação de convolução para a posição (i, j, k) no mapa de características O é expressa como:

$$O(i, j, k) = \sum_{u=1}^h \sum_{v=1}^m \sum_{w=1}^m I(i + u, j + v, k + w) \cdot K(u, v, w)$$

Aqui, a saída em um ponto tridimensional do mapa de características é obtida somando os produtos dos valores da entrada em torno da posição especificada e os valores do kernel, refletindo o processamento de dados em várias dimensões.

Essas equações são cruciais para entender como as características relevantes são extraídas dos dados por meio da convolução, possibilitando avanços significativos em diversas aplicações, como visão computacional e reconhecimento de padrões.