

Relatório LAMIA 15

Prática: Redes Neurais Convolucionais 2 (Deep Learning) (II)

Kaique Medeiros Lima

1 Introdução

Neste card, é apresentado o conteúdo de Redes Neurais Convolucionais, comumente utilizada em tarefas de processamento de imagens e também, de textos. Foi-se utilizado o conteúdo do curso *Deep Learning: Convolutional Neural Networks in Python* da Udemy.

2 Descrição da atividade

2.1 Convolução

A convolução é uma operação matemática fundamental em redes neurais, utilizada nas *Convolutional Neural Networks* (CNNs). É uma forma de combinar duas funções, resultando em uma terceira, sendo particularmente aplicada no processamento de imagens. A convolução permite extrair características importantes, como bordas, texturas e objetos, funcionando como um “modificador” que destaca padrões relevantes nas imagens.

2.1.1 Kernel ou Filtro

Durante a convolução, a matriz de entrada (imagem) é multiplicada por uma matriz menor chamada *kernel* ou filtro. Esse kernel tem tamanhos pré-definidos, como 3×3 , 5×5 ou 7×7 . A operação de convolução é realizada movendo o filtro sobre a matriz de entrada e multiplicando seus elementos, gerando uma matriz de saída menor que contém as informações extraídas.

2.1.2 Stride

O *stride* define o número de pixels que o filtro “anda” a cada iteração. Um *stride* maior significa que mais pixels são desconsiderados, resultando em uma matriz de saída ainda menor. Strides maiores podem ser úteis para reduzir o tamanho da imagem processada sem perder muita informação.

2.1.3 Padding

O *padding* consiste em adicionar pixels com valor zero nas bordas da imagem de entrada, garantindo que o tamanho da imagem original seja mantido após a aplicação do filtro. Ele também ajuda a evitar a perda de informações nas extremidades da imagem durante a convolução.

2.1.4 Mapa de Características

O mapa de características é o resultado gerado após a aplicação dos filtros. Cada filtro extrai uma característica diferente da imagem (como bordas ou texturas), gerando múltiplos mapas de características, que representam as informações obtidas de acordo com o filtro utilizado.

2.2 Redes Neurais Convolucionais (CNNs)

As Redes Neurais Convolucionais (CNNs) são compostas por várias camadas, cada uma com a função de manipular e extrair informações das imagens de entrada. Essas redes são projetadas especificamente para trabalhar com dados de imagens e são amplamente utilizadas em tarefas como classificação de imagens, detecção de objetos e reconhecimento facial. Nas CNNs, os filtros aplicados nas camadas convolucionais extraem padrões visuais, como bordas, formas e texturas, em diferentes níveis de abstração. Essas características são, então, passadas para as camadas densas, onde as informações são combinadas e processadas para realizar a tarefa de classificação.

2.3 Estrutura de uma CNN

2.3.1 Camada de Convolução

A camada de convolução é responsável pela extração de características das imagens. Ela aplica filtros sobre a imagem de entrada, gerando o mapa de características. Além disso, técnicas de normalização, como *Batch Normalization*, são frequentemente aplicadas nesta etapa para acelerar o treinamento e melhorar a estabilidade do modelo.

2.3.2 Camada de Pooling

A camada de *pooling* é usada para reduzir a dimensionalidade do mapa de características, o que diminui a complexidade do modelo e melhora a eficiência computacional, preservando as informações mais importantes. As técnicas mais comuns de pooling são o *Max Pooling*, que seleciona o valor máximo de cada região do mapa, e o *Average Pooling*, que calcula a média dos valores.

2.3.3 Flattening

O *flattening* transforma os dados multidimensionais gerados pelas camadas anteriores em um vetor unidimensional, necessário para a entrada nas camadas densas. Essa operação é crucial, pois as camadas densas requerem uma representação unidimensional dos dados para funcionar corretamente. Alternativamente, o *Global Max Pooling* (GMP) pode ser usado, que reduz cada mapa de características a um único valor representando a característica mais forte. A principal diferença entre o *flattening* e o GMP é que o primeiro preserva todas as características aprendidas, enquanto o segundo sintetiza as informações em valores mais compactos e representativos.

2.3.4 Camadas Densas

As camadas densas são compostas por neurônios completamente conectados aos neurônios da próxima camada. É nelas que ocorre o processamento final das características extraídas

anteriormente, e onde a classificação ou predição é realizada. Essas camadas combinam as informações extraídas das imagens para fornecer o resultado final do modelo.

3 Conclusão

Neste card, foi apresentado o conceito de convolução e Redes Neurais Convolucionais (CNNs), destacando a importância dessas técnicas no processamento de imagens e reconhecimento de padrões. As CNNs são amplamente utilizadas em tarefas de visão computacional, como classificação de imagens, devido à sua capacidade de extrair características relevantes das imagens e realizar tarefas complexas de forma eficiente. O conhecimento desses conceitos é fundamental para entender o funcionamento das CNNs e aplicar esses em projetos de *deep learning*.