



Inteligência Artificial Trabalho Prático 02

CCENS UFES – Departamento de Computação
Prof. Dr. Jacson Rodrigues Correia da Silva

Questões

1. O que é uma convolução? Explique com um exemplo.

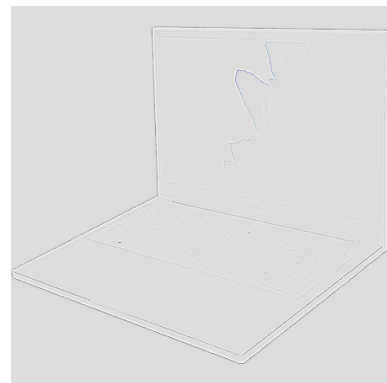
Na matemática, é uma operação que envolve o somatório do produto de duas funções (f e g) ao longo da região em que f e g se sobrepõem.

Ela é utilizada em campos como processamento de imagem aplicando filtros para por exemplo realçar algumas características da imagem ou então suavizar.

2. Escolha 3 imagens aleatórias. Após isso, utilizando o GIMP, aplique uma convolução em cada imagem escolhida, sendo as convoluções necessárias: aguçar, detectar bordas, destacar relevo.



Imagem original



Resultado da convolução detectar bordas



Imagem original



Resultado da convolução aguçar



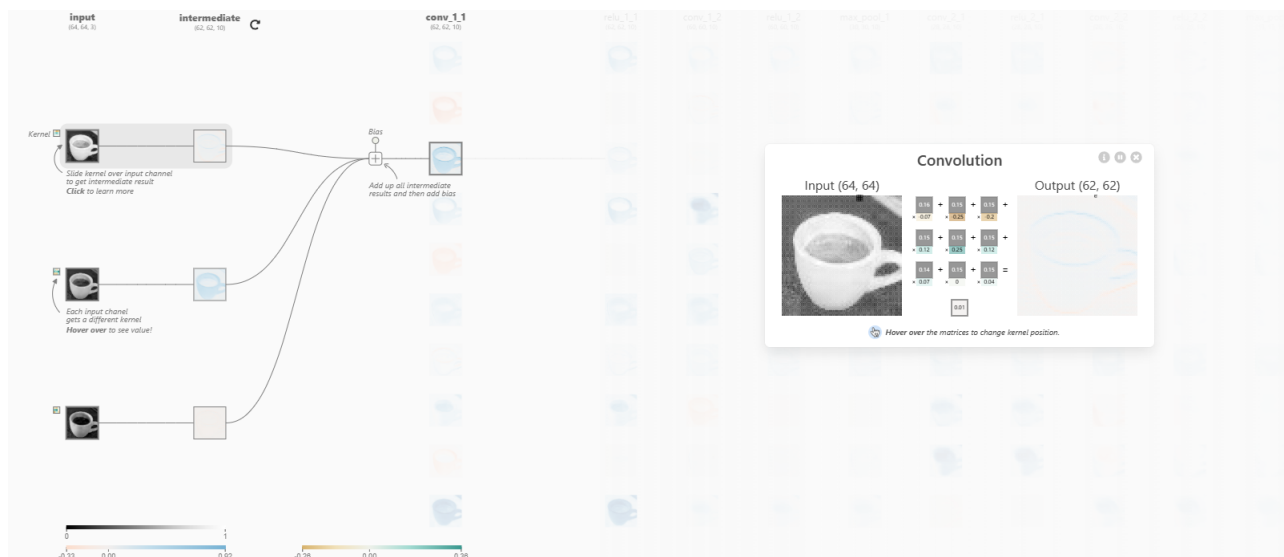
Imagem original



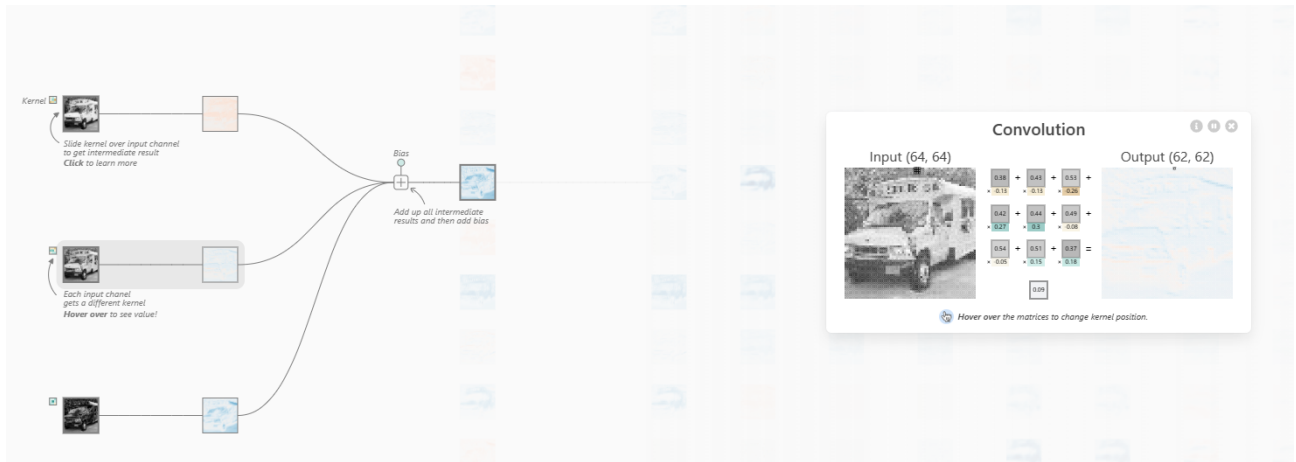
Resultado da convolução destacar relevo

3. Entre nesse site: <https://www.jeiks.net/copycat-cnn-explainer/>
 Interaja com o sistema apresentado e encontre um local onde é feita a convolução. Após isso, escolha uma imagem diferente da xícara, tire um print da convolução e apresente-a aqui. Quantas convoluções são realizadas ao todo nesse sistema?

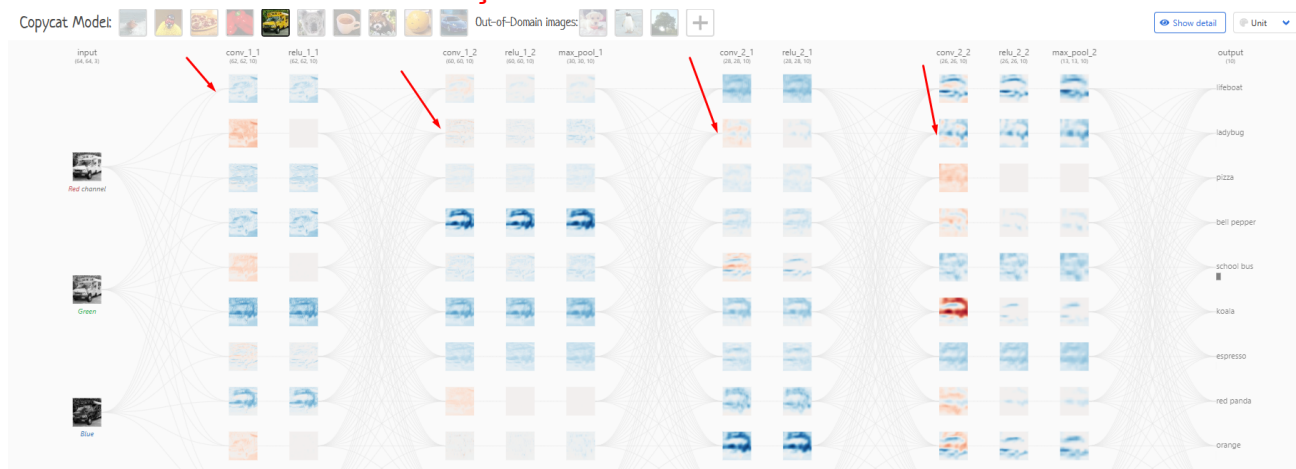
Primeiro local onde ocorre uma convolução:



Convolução diferente da xícara:



Ao total são realizadas 4 convoluções:



4. O que é uma função de ativação? Dê um exemplo.

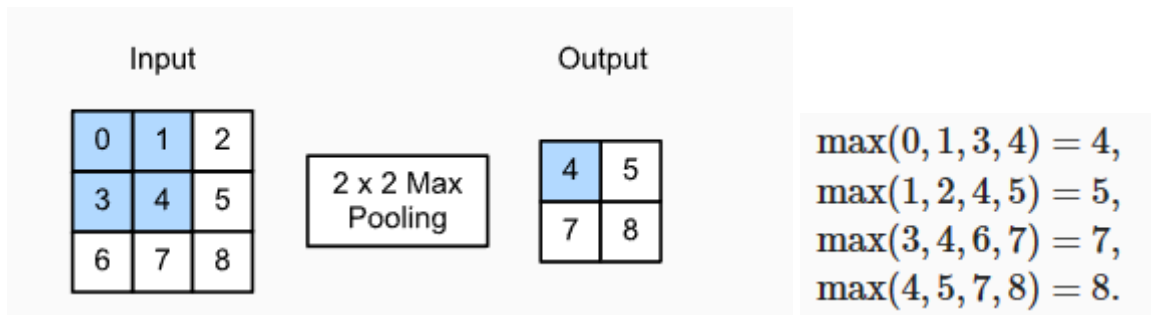
São funções que decidem se um neurônio irá ou não ativar, calculando a soma ponderada e adicionando o bias à ela. A grande maioria delas adiciona um fator de não-linearidade à rede neural. Um exemplo de função de ativação é a sigmóide, ela é geralmente utilizada quando a saída da rede precisa possuir apenas números positivos

Sua fórmula é:

$$\phi(x) = \frac{1}{1 + e^{-x}}$$

5. O que é uma operação de Pooling? Dê um exemplo.

O pooling é usado para reduzir a dimensionalidades das características que foram geradas nas camadas convolucionais mas mantendo as características mais importantes. Um exemplo de pooling é o Max pooling que verifica uma janela de valores e pega o maior deles como resultado. Na imagem abaixo possui um exemplo de 2x2 Max pooling. A primeira “janela” de verificação são os valores $\max([0, 1, 3, 4]) = 4$, o próximo seria o $\max([1, 2, 4, 5]) = 5$, depois seria $\max([3, 4, 6, 7]) = 7$ e por fim o $\max([4, 5, 7, 8]) = 8$ gerando assim $[4, 5, 7, 8]$.



6. Vamos construir uma rede neural convolucional simples:

Efetue o download do seguinte arquivo:

<https://s3.amazonaws.com/fast-ai-imageclas/cifar10.tgz> Ele possui uma pasta “train” e uma pasta “test”. Dentro delas, há 10 pastas com diferentes tipos de imagens. Como os nomes sugerem, a pasta “train” será utilizada no treinamento da rede e a de “test” nos testes.

Extraia o conteúdo do arquivo em uma pasta e, aleatoriamente, escolha só 3 classes (não valem ser as 3 primeiras) e apague as demais (ou seja, mantenha só três nomes de subdiretórios em “train” e os mesmos nomes de subdiretórios em “test”).

Agora, instale o Python+PyTorch+Matplotlib

(há vários tutoriais. Tente seguir um e, se não conseguir, me procure).

Execute então o seguinte código disponível em:

<https://drive.google.com/file/d/1Fnsyyfj9a-Tay3CM7gHNmXGEn-CLQHNI/view?usp=driv>

[e_link](#)

Apresente o texto (que apareceu na tela) referente à execução dele.



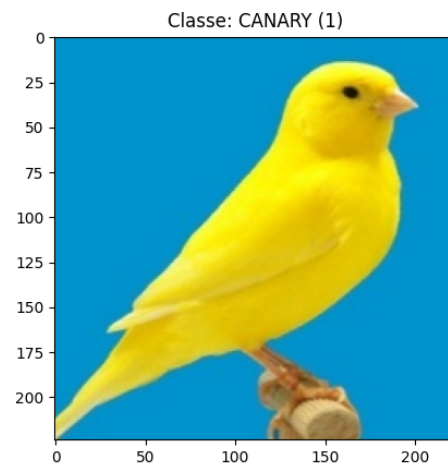
Qual foi sua taxa de acerto?

```
O treino acabou.
Salvando o modelo em "model.pth"
Testando:
Accuracia da rede: 82.57 %
```


7. Escolha um novo dataset de imagens em: <https://www.kaggle.com/datasets?search=image> Modifique o código apresentado no exercício 6 e também organize corretamente o dataset que escolheu para fornecer 3 classes para o treinamento e testes.

Após isso, efetue o treinamento e testes 3 vezes e faça um gráfico boxplot para apresentar a acurácia dos treinamentos.

```
C:\repo\codigos\py\convolutional
(venv) λ python ./main.py
Classes do train_data: ['ALBATROSS', 'CANARY', 'VERDIN']
Tamanho da figura: torch.Size([3, 224, 224])
```



```
Testando:
Accuracia da rede 1: 93.33 %
Accuracia da rede 2: 93.33 %
Accuracia da rede 3: 93.33 %
```

