

Relatório 20 - Prática Visão Computacional (III)

Lucas Scheffer Hundsdorfer

Descrição da atividade

2. Deep learning fundamentals:

Aula 3 - What is Deep Learning:

Deep learning é um subconjunto do Machine Learning, deep learning utiliza de redes neurais artificiais, é utilizada em áreas como visão computacional, processamento da linguagem natural e reconhecimento de fala.

Aula 4 - Introduction to PyTorch:

Pytorch é uma biblioteca de python dedicada ao machine learning de código aberto, desenvolvida pelo laboratório de IA do facebook, com um intuito de ser fácil de aprender tornando se popular entre desenvolvedores e pesquisadores

Aula 5 - Tensor:

Tensores são a estrutura de dados utilizada para treinar as redes neurais. Basicamente é a essência do deep learning, é como ele é construído e treinado, ele pode representar inúmeras coisas, como uma imagem, uma sequência de palavras e entre outros.

Aulas 6,7,8 foram práticas.

Aula 9 - Advantages of tensor:

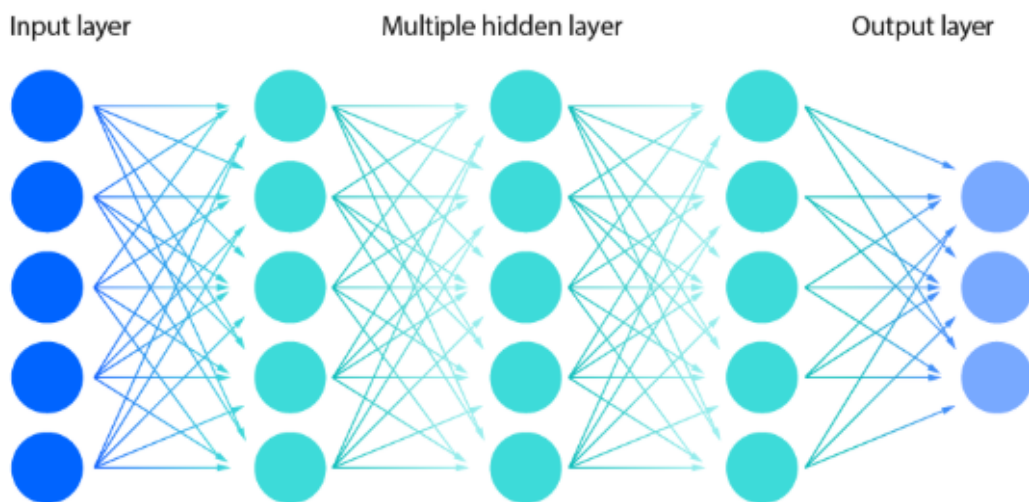
A vantagem de se utilizar os tensores do pytorch do que utilizar os arrays do numpy por exemplo é que os tensores já são otimizados para cálculos dentro do Deep Learning fazendo assim com que o treinamento seja mais eficaz e rápido, facilitando a computação de gradientes como backpropagation e tem uma API mais intuitiva.

3. Building Neural Networks with PyTorch:

Aula 10 - What is a Neural Network:

Redes neurais é um tipo de modelo de aprendizado de máquina onde é baseado no cérebro humano, é utilizado onde algoritmos normais têm maiores dificuldades, como assistentes de voz, como Alexa, Siri, Cortana entre outros. Funciona em camadas de neurônios ou nós que realizam cálculos dos dados de entrada, e cada neurônio produz uma saída baseada em pesos e vieses atribuídos a eles, esses vieses e pesos são ajustados na etapa de treinamento para ver qual é mais adequado, no momento de treinamento o objetivo é

reduzir o erro.



Aula 11 - Neural Network Training Workflow:

A melhor forma de treinar uma rede neural é separar as etapas, a primeira etapa é definir o objetivo da rede, o segundo é unir os dados e prepará-los para nossa rede, que envolve tanto normalizar, dimensionar e também dividir em dados de treino e dados de teste, após esses passos feitos pode se avançar para a próxima etapa que é projetar a arquitetura da rede, que envolve selecionar as camadas e as funções de ativação mais adequadas para seu conjunto de dados e o que quer resolver.

Aula 12 - Neural Network Architecture:

A arquitetura de uma rede neural se refere a layout dos nós e das camadas individuais dentro da rede. Um tipo comum de rede neural é a Rede neural convolucional ou CNN, essa arquitetura é utilizada para processamento de imagens. Outro tipo conhecido é o RNN ou rede neural recorrente usada para processamento de texto ou fala. A arquitetura é importantíssima para se chegar no resultado desejado, uma arquitetura mal desenvolvida provavelmente não vai ser eficiente e precisa.

Aula 14 - Activation and Loss Functions:

As funções de ativação e de erro são essenciais para mapear entradas para saídas e medir a eficiência da rede. A função de ativação é uma função matemática aplicada na entrada de um neurônio, introduzindo uma não linearidade, fazendo assim que a rede consiga aprender padrões complexos. Funções de erro são usadas para ver o desempenho da rede, elas comparam o resultado esperado com o resultado obtido retornando um valor que indica o quão perto foi essa comparação.

Aula 16 - Optimizers:

Otimizador é um algoritmo que atualiza os pesos e vieses para minimizar o erro durante todo o treinamento, o otimizador calcula os gradientes em da função de perda em relação aos parâmetros da rede e ajusta esses parâmetros para reduzir o erro. Existem vários tipos de otimizadores e a escolha dele pode alterar muito o resultado da sua rede.

Aula 18 - Dataset and Data Loader:

O dataset e o data loader são ferramentas poderosas para auxiliar no treinamento de redes neurais, ajudando a carregar e preparar a sua base de dados para o treinamento e na escolha do 'batch size' um hiperparâmetro que permite a divisão dos dados que impacta significativamente no processo de treinamento.

4. Neural Network for Images:

Aula 21 - Introduction to Image Classification:

Através das redes neurais é possível extrair dados e padrões das imagens. Normalmente para a classificação de imagens é utilizada as redes neurais convolucionais.

Aula 24 - HyperParameters:

Hiperparâmetros são os parâmetros da rede neural que são definidos previamente ao treinamento, esses parâmetros não podem ser aprendidos pelo modelo e afeta diretamente como o modelo aprende com os dados.

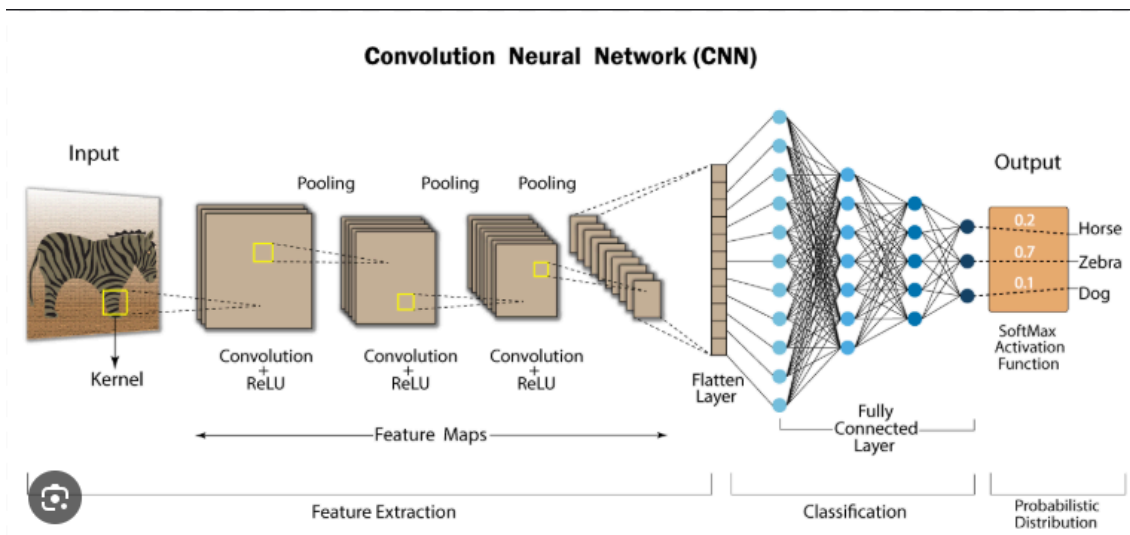
Aula 26 - Data Normalization:

A normalização dos dados é uma etapa antes do processamento dos dados. Normalizar significa ajustar os valores dos pixels das imagens para um intervalo conhecido, quase sempre entre 0 e 1 ou -1 e 1. Um dos motivos para ser necessário a normalização é que algumas imagens tem variação de brilho, a normalização retira isso.

5. Convolutional Neural Networks (CNN's):

Aula 27 - Introduction to CNN:

As redes neurais convolucionais é a melhor maneira que conhecemos hoje para o reconhecimento de imagens, tudo isso por causa das camadas convolucionais que existem nessas redes neurais, camadas que conseguem reconhecer padrões observando pequenas partes da imagem por vez.

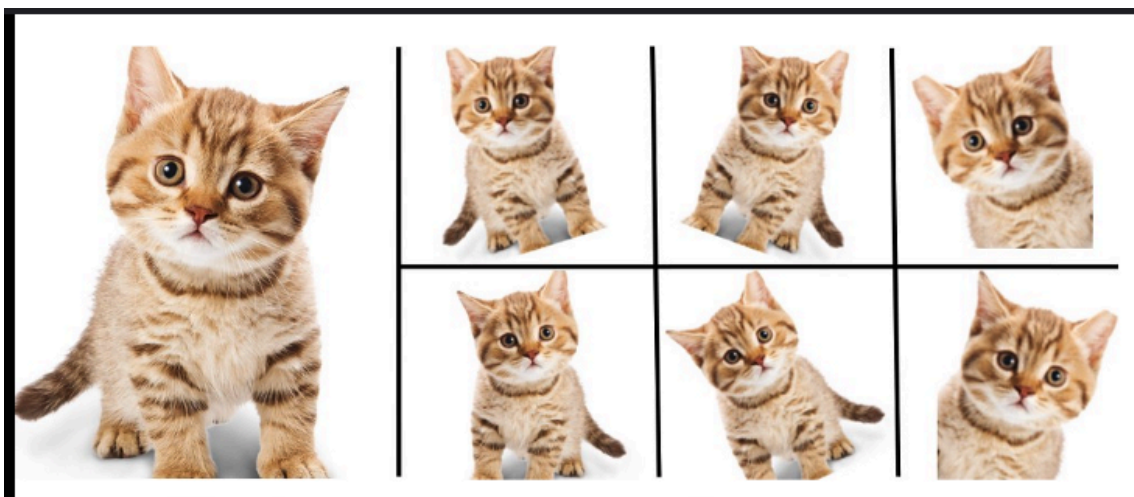


Aula 28 - Why CNN?:

As redes neurais convolucionais são as mais adequadas para tarefas no campo da visão computacional por causa que elas são mais eficientes em reconhecer padrões como bordas, cantos e texturas. E a eficácia dela nisso é dado a aplicação de uma série de filtros que cria 'mapas de recurso' onde esses mapas de recurso conseguem captar a existência ou não de certos padrões. Além das camadas de convolução existem outros recursos dentro da rede, como camadas de pooling e também camadas totalmente conectadas.

Aula 30 - Data Augmentation:

Data Augmentation é uma técnica de aumentar a sua base de dados aplicando transformação nas imagens já presentes no dataset. O exemplo abaixo demonstra isso muito bem, a mesma imagem do mesmo gato, porém com algumas transformações. Além de deixar a base de dados mais robusta, deixa o modelo melhor treinado com maiores precisões.



6. Auto Encoders:

Aula 33 - Introduction to Auto Encoders:

Auto encoders é um tipo de rede neural que consegue codificar e decodificar imagens em outras coisas. Consiste em 2 partes, um codificador e um decodificador. O codificador recebe uma imagem de entrada e mapeia para um espaço latente e o decodificador recebe esse espaço latente e reconstrói a imagem original. Um dos vários benefícios dos autoencoders é que eles podem ser treinados sem rótulos e podem ser usados para compressão de dados.

Aula 36 - Introduction to Variational Auto Encoders:

Auto encoders variacionais ou VAE é um tipo de auto encoder que incorpora inferência bayesiana no processo de aprendizado. O treinamento de um VAE envolve minimizar o erro de reconstrução e também maximizar um limite na probabilidade de log dos dados

7. Hands on Project:

Aula 38 - Section Overview:

Com o conhecimento adquirido durante as sessões do curso é possível criar efeitos visuais impressionantes como o Deep Fake, Neural Style Transfer e também Image Colorization. Deep fake é manipular vídeos e áudios através de algoritmos complexos. Neural Style Transfer é aplicar o estilo de uma imagem no conteúdo de outra. Image Colorization é adicionar cor ao preto e branco, a utilização seria dar vida a fotos antigas.

Exemplo de Deep Fake:

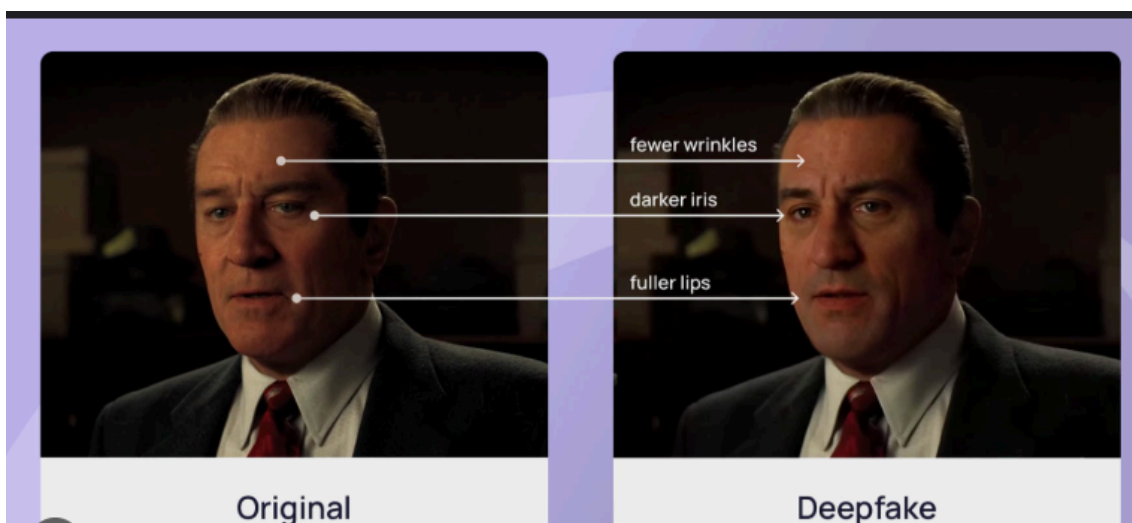
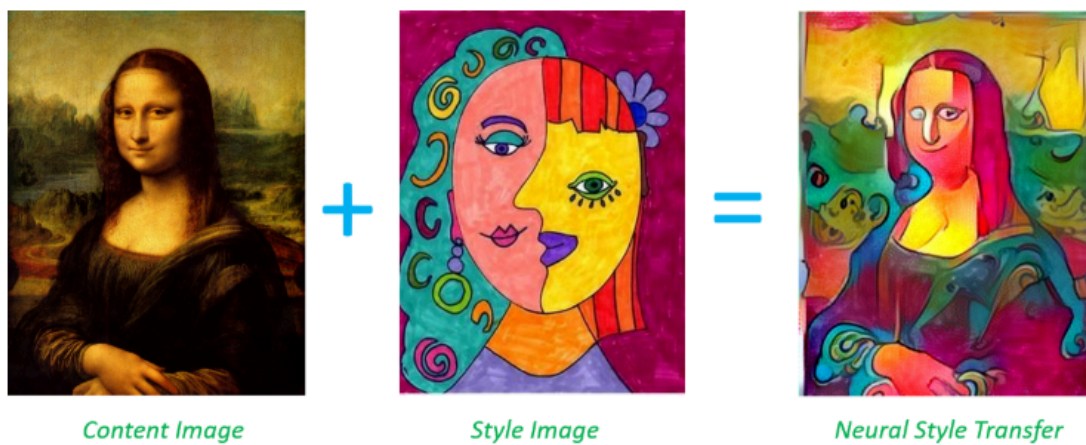


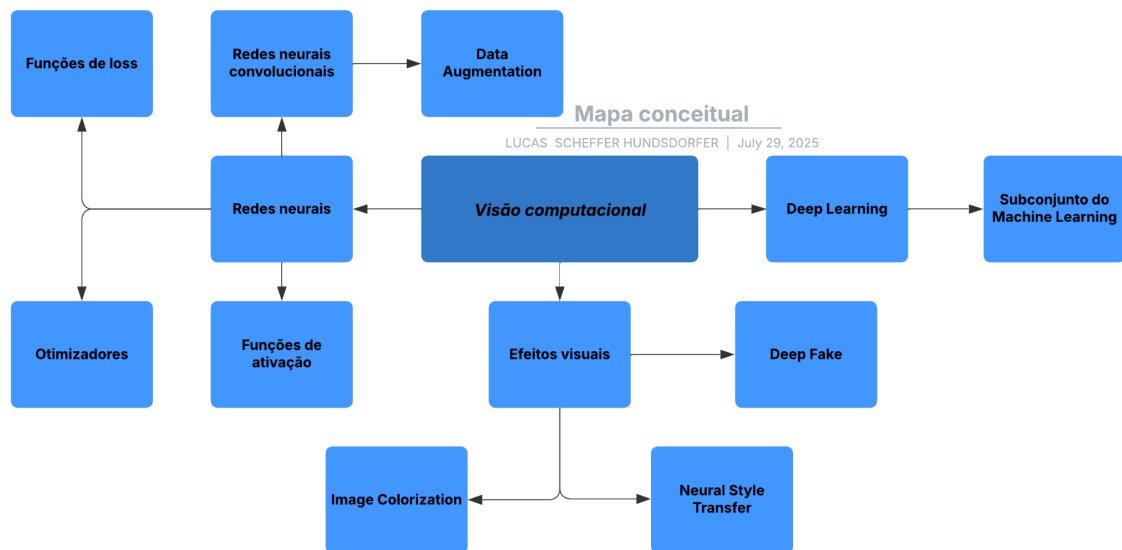
Image Colorization exempo:



Neural Style Transfer exempo:



Insight visual original:



Conclusões

Com base no conhecimento que foi passado ao longo do curso, fica claro que as redes neurais e principalmente as redes neurais convolucionais representam uma ferramenta poderosa ainda mais quando aplicada nos conceitos vistos nos vídeos como a visão computacional. É uma área complexa de ser compreendida, porém com os esforços necessários se torna um conhecimento útil tanto para trabalho e também para entender como algumas coisas do dia a dia funcionam.

Referências