

# Relatório CARD 18 - Prática: Visão Computacional (III)

Willian Augusto Soder de Souza

O objetivo deste relatório é apresentar os principais conhecimentos adquiridos durante o curso 'The Ultimate Computer Vision and Deep Learning Course'. Nesse curso, o apresentador aborda conceitos teóricos importantes no campo da visão computacional, além de ensinar a parte prática utilizando a biblioteca PyTorch. Abaixo, segue um resumo dos principais conceitos aprendidos ao longo do curso.

- **Tensor:** é uma estrutura de dados multidimensional que permite a representação e manipulação eficiente de informações complexas, sendo fundamental para o processamento de imagens e vídeos em visão computacional.
- **Operações com tensores:** operações como adição e soma com tensores são fundamentais para manipular e transformar dados em aprendizado de máquina, especialmente em redes neurais e visão computacional.
- **Rede Neural:** é um modelo computacional inspirado no funcionamento do cérebro humano, usado principalmente para aprender padrões em dados e fazer previsões ou classificações. Ela é composta por várias camadas de nós, chamadas de neurônios ou unidades, que processam informações de forma semelhante às células nervosas do cérebro.

Arquitetura de um rede neural:

- **Camada de Entrada (Input Layer):** a primeira camada, onde os dados brutos entram na rede. Cada neurônio nesta camada representa uma característica ou entrada dos dados (como pixels em uma imagem).
- **Camadas Ocultas (Hidden Layers):** localizadas entre a camada de entrada e a camada de saída, essas camadas processam os dados. Quanto mais camadas ocultas uma rede tiver, mais complexos os padrões que ela pode aprender.
- **Camada de Saída (Output Layer):** A última camada da rede, que gera a previsão final. O número de neurônios nesta camada depende do tipo de tarefa. Por exemplo, em uma tarefa de classificação com três classes, haveria três neurônios.
- **Conexões e Pesos:** neurônios de uma camada são conectados a neurônios da próxima camada por meio de conexões com pesos ajustáveis. Esses pesos determinam a influência de cada neurônio na saída da camada seguinte.
- **Função de Ativação:** aplicada em cada neurônio das camadas ocultas e da camada de saída para introduzir não-linearidade, permitindo que a rede aprenda padrões complexos.
- **Loss Function:** tem a função de medir o quão bem o modelo está performando, comparando as previsões feitas pelo modelo com os valores reais esperados.
- **Optimizer:** é um algoritmo que ajusta os pesos e vieses de uma rede neural durante o treinamento para minimizar a função de perda. O objetivo do otimizador é encontrar o conjunto de parâmetros que resulte na melhor performance do modelo, ou seja, a menor perda possível.

- **Dataset:** é uma coleção de dados usados para treinar, validar e testar modelos de aprendizado de máquina. Os dados podem ser imagens, textos, sons, ou qualquer outra forma de informação que o modelo precisa processar.
- **Data loader:** é uma ferramenta que facilita o carregamento de dados em mini-lotes (mini-batches) para o modelo durante o treinamento e validação. Em vez de carregar todo o dataset de uma vez, o que pode ser ineficiente em termos de memória e processamento, o data loader divide os dados em partes menores e os carrega de forma iterativa.
- **Sequential:** é uma forma simples de construir redes neurais que permite empilhar camadas de forma sequencial, onde a saída de uma camada é a entrada da próxima. É ideal para arquiteturas lineares e diretas, onde os dados fluem em uma única direção. O modelo Sequential é fácil de usar e adequado para construir redes neurais simples, como Multilayer Perceptrons (MLPs) e redes convolucionais básicas.
- **Rede Neural Profunda:** um tipo de rede neural que possui múltiplas camadas ocultas entre a camada de entrada e a camada de saída. O termo "profundo" refere-se ao grande número de camadas ocultas, que permite à rede aprender representações complexas e hierárquicas dos dados.
- **Data normalization:** é o processo de ajustar os valores dos dados para uma escala padrão, o que facilita o treinamento de modelos de aprendizado de máquina e melhora a performance e a convergência dos algoritmos.
- **Convolutional Neural Networks (CNNs):** são um tipo especializado de rede neural projetado para processar dados com uma estrutura de grid, como imagens. Elas são amplamente utilizadas em tarefas de visão computacional, como reconhecimento de objetos, detecção de faces e segmentação de imagens.
- **Auto-encoders:** são um tipo de rede neural utilizada para aprender uma representação compacta e eficiente dos dados, geralmente para redução de dimensionalidade, compressão de dados ou extração de características. Eles são compostos por duas partes principais: o encoder e o decoder.

## CONCLUSÃO:

Entender os conceitos discutidos neste relatório é crucial para avançar em aprendizado de máquina e visão computacional. O conhecimento sobre tensores, redes neurais profundas, e autoencoders é fundamental para manipulação de dados, extração de características e construção de modelos complexos. A normalização de dados e a aplicação de encoders garantem uma preparação eficaz dos dados para o treinamento de modelos.

Além disso, a biblioteca PyTorch oferece as ferramentas necessárias para implementar e otimizar esses conceitos de maneira prática. Dominar essas técnicas e ferramentas é essencial para resolver problemas complexos na área de visão computacional e aprendizado de máquina.