

Relatório 19 - Como Estamos Ensinando os Computadores a Compreender Fotos

Guilherme Loan Schneider

Descrição da atividade

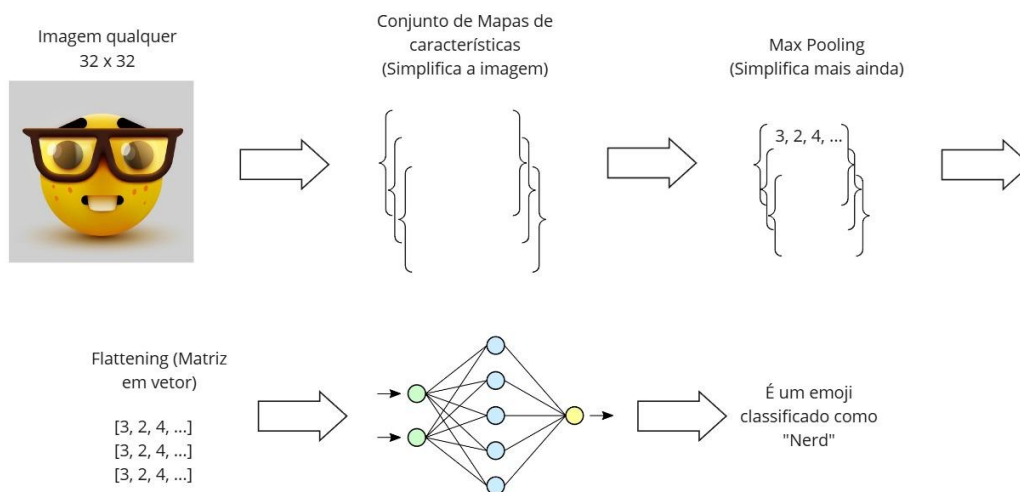
O vídeo inicia comentando sobre a forma que os humanos conseguem compreender cenários em sua vida, ou seja, aprendendo o que são objetos, animais, veículos, aviões, dentre inúmeros outros. Correlacionado a isso, a palestrante retrata a dificuldade que é fazer um computador compreender imagens, pois é necessária uma quantidade grande de dados para obter um resultado aceitável.

A percepção dos humanos é feita no cérebro, a partir das sinapses nervosas, que irão compreender a partir dos trilhões de neurônios, as informações naquilo que ele viu. A ideia para fazer os computadores compreenderem imagens é a mesma, onde o algoritmo tentará se assimilar ao máximo a um cérebro humano.

Nos humanos, quando observamos imagens em que existe um animal tampado por algum objeto, ou até mesmo um cavalo com a sombra de uma cerca, fazendo-o parecer uma zebra, ele é capaz de identificar o contexto da imagem e perceber o que está ali, já para um computador, essa tarefa é muito mais árdua, já que ele buscará informações na imagem que o diga com o que aquilo se assemelha ao que ele já viu.

A rede neural artificial utilizada pela palestrante é do tipo convolucional, amplamente utilizada para reconhecimento de imagens.

Rede Neural Convolucional



Ainda no âmbito da rede neural, retrata-se a dificuldade de obter imagens de qualidade para treinar a CNN (Convolutional Neural Network), onde naquela época mobilizou-se um montante de quase 50 mil pessoas ao redor do mundo para reunir imagens de diferentes objetos, aviões, casas, animais, dentre outros. A base de dados obtida totalizou 15 milhões de imagens com 22 mil features (classes) diferentes.

Um fato curioso é a adição de imagens que não existem de verdade, como a mostrada abaixo, um animal que não existe. É possível perceber que o algoritmo

conseguiu extrair características da imagem, indicando que ali possivelmente existe um animal, mas o algoritmo não sabe qual é visto que é uma combinação de um jacaré e um pato.



Os resultados obtidos foram muito satisfatórios de acordo com a palestrante, criando um marco na compreensão de imagens por computadores. Além disso, comenta que a maior dificuldade será fazer o computador compreender o que está além das imagens.



Referencias

 <https://www.youtube.com/watch?v=40riCqvRoMs>