

Universidade de Pernambuco

Programa de Pós-Graduação em Engenharia da Computação (PPGEC)

Proposta de Projeto de Doutorado

Área: Computação Inteligente

Título: Utilização de novas arquiteturas profundas para procurar melhorar o desempenho no reconhecimento de padrões.

Orientador – Mêuser J S Valença (meuservalenca@gmail.com)

Descrição

As Redes Neurais Convolucionais (ConvNets ou CNNs) são redes neurais artificiais profundas que podem ser usadas para classificar imagens, agrupá-las por similaridade (busca de fotos) e realizar reconhecimento de objetos dentro de cenas. São algoritmos que podem identificar rostos, indivíduos, sinais de rua, cenouras, ornitorrincos e muitos outros aspectos dos dados visuais. A eficácia das redes convolucionais no reconhecimento de imagem é uma das principais razões pelas quais o mundo testemunhou a eficácia do aprendizado profundo. Este tipo de rede está impulsionando grandes avanços em [Visão Computacional](#). Entretanto como diz o Prof. Hinton o grande pulo do “gato” que é a operação de maxpooling é também seu grande gargalo: *Hinton: “A operação de pooling usada em redes neurais convolucionais é um grande erro e o fato de funcionar tão bem é um desastre.”*

A representação interna de dados de uma rede neural convolucional não leva em conta importantes hierarquias espaciais entre objetos simples e complexos.

De acordo com o Prof. Hinton para aprender a diferenciar os dedos, o cérebro humano precisa ver apenas algumas dúzias de exemplos, centenas no máximo. As CNNs, por outro lado, precisam de dezenas de milhares de exemplos para alcançar um desempenho muito bom, o que parece ser uma abordagem de força bruta que é claramente inferior ao que fazemos com nossos cérebros.

O aprendizado de máquina com eficiência de dados é a palavra da moda no momento. A idéia é projetar algoritmos de ML que tenham bom desempenho, mas não exijam centenas de milhares de pontos de dados anotados. Um algoritmo ideal de ML exigiria apenas um ponto de dados para aprender um conceito inteiro. Diz-se que esse sistema de aprendizado de conceitos é um sistema de aprendizado único. O desafio de fazer o One-shot Learning se deve à dificuldade de ter Representações dinâmicas. O objetivo é utilizar Comparadores recorrentes atentos, que propõe um modelo simples, porém poderoso, para o aprendizado eficiente de dados.

Referências Bibliográficas

- [1] P. Sermanet, K. Kavukcuoglu, S. Chintala, and Y. Lecun, “Pedestrian detection with unsupervised multi-stage feature learning,” in The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2013.
- [2] M. D. Zeiler and R. Fergus, Computer Vision – ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part I. Springer International Publishing, 2014, ch. Visualizing and Understanding Convolutional Networks, pp. 818–833.
- [3] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” Nature, vol. 521, pp. 436–444, 2015.
- [4] W. Ouyang and X. Wang, “Joint deep learning for pedestrian detection,” in The IEEE International Conference on Computer Vision (ICCV), 2013.
- [5] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, “Gradient-based learning applied to document recognition,” in Proceedings of the IEEE, 1998, pp. 2278–2324.
- [6] P. Dollár, C. Wojek, B. Schiele, and P. Perona, “Pedestrian detection: An evaluation of the state of the art,” PAMI, vol. 34, 2012.
- [7] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “Imagenet classification with deep convolutional neural networks,” in Advances in Neural Information Processing Systems 25. Curran Associates, Inc., 2012, pp. 1097–1105.
- [8] Martín Abadi, Ashish Agarwal, Paul Barham, Eugene Brevdo, Zhifeng Chen, Craig Citro, Greg S Corrado, Andy Davis, Jeffrey Dean, Matthieu Devin, et al. Tensorflow: Large-scale machine learning on heterogeneous distributed systems. arXiv preprint arXiv:1603.04467, 2016.

[9] Attentive Recurrent Comparators. Pranav Shyam, Shubham Gupta, Ambedkar Dukkipati ([arXiv:1703.00767](https://arxiv.org/abs/1703.00767))