

RECONHECIMENTO DE PADRÕES EM IMAGEM EMPREGANDO APRENDIZAGEM PROFUNDA

Helena Cristina dos Reis Braga¹
Dr. Clarimar José Coelho²

Acadêmica do x Período do Curso de Engenharia Civil¹ UniEVANGÉLICA
Programa de Pós-Graduação em Sociedade Tecnologia e Meio Ambiente² UniEVANGÉLICA

INTRODUÇÃO

Reconhecimento de padrões é a classificação de dados de entrada a partir da extração de características importantes provenientes de muitos dados ruidosos ou erros aleatórios (*outlier*) (GONZALEZ, THOMAS, 1978). O termo aprendizagem profunda (*Deep Learning*, DL) foi introduzido em aprendizagem de máquina (*Machine Learning*, ML) e depois passou a ser usado em redes neurais artificiais (*Artificial Neural Network*, ANN). Os métodos de aprendizagem DL são compostos por várias camadas de aprendizagem para extrair conhecimento dos dados com múltiplos níveis de abstração (SCHMIDHUBER, 2015).

A rede neural convolucional profunda (*Convolutional Neural Network*, CNN) é uma arquitetura de aprendizagem profunda usada para a classificação de imagem para a solução de problemas em medicina, biomedicina e sensoriamento remoto. A classificação de imagem consiste em categorizar imagens em uma das várias classes predefinidas. Trata-se de um problema básico de visão computacional como a localização, detecção e segmentação (DU, CAI, WANG, 2016). Por exemplo, numa aplicação em análise ambiental usando imagem de satélite, as classes podem ser área desmatada, área com vegetação rasteira. Dada uma imagem e um sistema de reconhecimento de objetos é possível definir quais serão os objetos de interesse para o reconhecimento automático na imagem (KAMILARIS, PRENAFETA-BOLDÚ, 2017).

Nos últimos anos, a aprendizagem profunda vem sendo usada em diferente áreas do conhecimento. Duarte e colaboradores desenvolveram uma estrutura CNN usando conexões residuais e convoluções dilatadas considerando tanto as amostras de imagens aéreas tripuladas quanto as não-tripuladas para realizar a classificação da imagem de satélite de danos de construção civil (DUARTE, *et al.*, 2018).

Neste trabalho é desenvolvida uma arquitetura CNN para o reconhecimento de face de pessoas a partir de fotos de um banco de imagens público. Esse banco de imagem é usado para o desenvolvimento didático inicial da CNN. No futuro, esse banco de imagem pode ser composto por imagens de satélite.

METODOLOGIA

Uma CNN é um *pipeline* (tubulação) com vários estágios. As imagens vão para um lado e a probabilidade da imagem ser a face de interesse sai do outro lado. Existem três tipos de camadas: convolucionais, agrupamento (*pooling*), totalmente conectadas. A camada convolucional é composta por um conjunto de filtros (*kernels*) onde a CNN realiza a aprendizagem a partir do conjunto de dados de treinamento. A camada de agrupamento reduz o tamanho das matrizes resultantes do processo de convolução (operador linear ou integral de superposição) que diminui a quantidade de parâmetros de aprendizagem da rede. Na camada totalmente conectada, cada elemento das matrizes de saída da convolução é conectada em um neurônio de entrada. A camada totalmente conectada é similar a última camada de uma ANN tradicional. A organização das camadas em uma CNN é conhecida como a arquitetura CNN (VENKATESAN, LI, 2018).

A implementação de uma CNN requer a construção de várias funções auxiliares como a função de ativação, treinamento e teste da CNN. Para o treinamento e teste da CNN é utilizado um conjunto de fotos do banco de imagem público FEI *Database* disponível em <https://fei.edu.br/~cet/facedatabase.html>. Um subconjunto com onze imagens é organizado a partir de diferentes conjuntos de dados do banco de dados FEI para o treinamento e teste da CNN proposta.

RESULTADOS

A função (programa) para ativação é construída baseada na função tangente hiperbólica e sigmóide. A função é executada e gera um gráfico do comportamento das duas funções para ativação da CNN. Uma função para o enquadramento das fotos é implementada para que sejam razoavelmente consistentes em tamanho e minimizem o plano de fundo. Cada foto é convertida para tons de cinza no final do procedimento.



Figura 1. Fotos de entrada e suas equivalentes convertidas para tons de cinza.

A função de treinamento foi implementada e usada para treinar a CNN a partir das fotos e resultou na probabilidade da última foto do conjunto de dados ser uma pessoa em 15%. A princípio, essa probabilidade parece baixo. Porém, levando em conta que uma CNN deve ser treinada com conjuntos de dados grandes, na casa de milhares. Um conjunto de treinamento com onze fotos mostra que a CNN é promissora para o reconhecimento de objetos em imagens. A função de teste da CNN também é desenvolvida e usada para o teste de entrada. A função teste retornou uma

mensagem que seria necessário um computador com maior capacidade de memória para ser executada. Em resumo, um protótipo com os métodos propostos foi implementado e testado. A CNN é uma tecnologia recente e complexa e requer maiores investimentos de tempo e adaptações para se tornar operacional. Pode-se considerar que o plano de trabalho foi desenvolvido com sucesso.

REFERÊNCIAS

DU, X. CAI, Y. WANG, S. Overview of deep learning, 31st Youth Academic Annual Conference of Chinese Association of Automation, Wusan, China, p.11-13, 2016.

DUARTE, D., NEX, F., KERLE, N., VOSSSELMANA, G. Satellite image classification of building Damages using Airborne and Satellite image samples in a deep learning approach, ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, v. IV-2, 2018, ISPRS TC II Mid-term Symposium "Towards Photogrammetry 2020", 4–7 June 2018, Riva del Garda, Italy.

GONZALEZ, R. C. THOMAZ, M. G. Syntatic Pattern Recognition: an Introduction, Addison Wesley, Reading, MA, 1978.

KAMILARIS, A., PRENAFETA-BOLDÚ, F. X., A review of the use of convolutional neural networks in agriculture, The Journal of Agricultural Science 1–11. <https://doi.org/10.1017/S0021859618000436>, 2017.

NGUYEN, T. JIHO, H., DONG-CHUL, J., Satellite Image Classification Using Convolutional Learning. 1558. 2237-2240. [10.1063/1.4825984](https://doi.org/10.1063/1.4825984), 2013.

SCHMIDHUBER, J. Deep Learning. Scholarpedia, v. 10, n. 11:32832, URL http://www.scholarpedia.org/article/Deep_Learning. revision #152272, 2015. doi: 10.4249/scholarpedia.32832.

VENKATESAN, R. LI, B. Convolutional neural network in visual computing: A concise guide, CRC Press, 2018.