

Uso de aprendizado profundo na classificação de imagens histopatológicas

Ana Kely Lopes. Ferreira¹, Críston Pereira de Souza¹,
Francisco Camilo de Lima Filho¹, Francisco Ivanilso Soares Araujo¹

¹Universidade Federal do Ceará (UFC) – Campus Quixadá
Caixa Postal 63902.580 – Quixadá – CE – Brasil

{kely}@alu.ufc.br, {cristonsouza}@ufc.br

{caamilolima, ivanilson.soares.is}@gmail.com

Abstract. *This article deals with the classification of histopathological images of breast cancer, using a convolutional neural network and deep learning techniques. The database used in this study comes from the Data Science competition platform, Kaggle and is part of the Breast Cancer Histopathological Image Classification database (BreakHis), built in collaboration with the PD Pathological Anatomy and Cytology Laboratory of Paraná.*

Resumo. *Este artigo trata-se da classificação de imagens histopatológicas de câncer de mama, com a utilização de uma rede neural convolucional e técnicas de aprendizado profundo, bem como a análise dos resultados obtidos a partir da aplicação de tais técnicas. A base de dados utilizada neste trabalho é oriunda da plataforma de competições em Ciência de Dados, Kaggle e faz parte do banco de dados Classificação de Imagens Histopatológicas do Câncer de Mama (BreakHis), que foi construído em colaboração com o Laboratório PD Anatomia Patológica e Citologia do Paraná, para trabalhar com imagens de tumores de câncer de mama a fim de contribuir na prevenção deste tipo de câncer.*

1. Introdução

O câncer de mama é uma doença caracterizada pela multiplicação das células da mama de maneira desordenada, causando anomalias que se multiplicam, formando um tumor. Segundo o Instituto Nacional de Câncer (2020), o câncer de mama é um dos mais incidentes em mulheres no mundo, representando cerca de 24,2% do total de casos no ano de 2018, sendo considerado como a quinta causa de morte por câncer em geral com 626.679 óbitos. Além disso, segundo o Atlas de mortalidade por câncer no Brasil, no ano de 1979 e 1999, esse tipo de câncer foi considerado como a primeira causa de óbito por câncer em mulheres, destacando como uma das principais características, a idade.

Além disso, o diagnóstico do câncer de mama pode ser feito por exames clínicos das mamas, exames de imagem, como mamografia, ultrassonografia ou ressonância magnética e pelo exame simples de toque. Assim, quando detecta-se alguma alteração suspeita na mamografia, é feita uma biópsia que, [Santos and Oguissso 2007] destaca como o único exame que pode estabelecer com certeza o diagnóstico de câncer de mama.

O aprendizado profundo é um tipo de aprendizado de máquina, que tem o objetivo de reconhecer padrões que [Chollet 2018] define como um campo da inteligência artificial que tem se tornado um dos mais importantes nos últimos anos e que trouxe diversas

contribuições para todos os setores da sociedade. Nesse contexto, uma Rede Neural Convolutacional, do inglês *Convolutional Neural Network* (CNN), é um algoritmo de aprendizado profundo utilizado para captar padrões em imagens, composta de neurônios que se otimizam por meio de aprendizado [Albawi et al. 2017].

Dessa forma, objetiva-se com este trabalho classificar imagens histopatológicas do conjunto de dados BreaKHis, oriundo da plataforma Kaggle, retiradas de exames de biópsia para diagnóstico de câncer de mama, utilizando técnicas de aprendizado profundo, bem como o uso de uma rede neural convolutacional. Na Seção 2, é apresentada a metodologia utilizada. Na Seção 3 são apresentados os resultados e discussões. Por fim, na Seção 4 são apresentadas as considerações finais e perspectivas de trabalhos futuros.

2. Metodologia

O presente estudo foi realizado seguindo as etapas: obtenção e análise dos dados, compreensão da referente problemática, definição da rede neural e das camadas de convolução, aumento de dados, escolha de hiperparâmetros e com a utilização de técnicas de aprendizado profundo. A realização deste trabalho e das pesquisas, foram feitas com o intermédio de artigos científicos e monografias acadêmicas oriundas da plataforma Google Scholar. Buscando assim, compreender a aplicação das técnicas e do algoritmo que foram utilizados.

2.1. O conjunto de dados BreaKHis

O conjunto de dados utilizado é oriundo da plataforma de competições em ciência de dados, Kaggle e faz parte do banco de dados Classificação de Imagens Histopatológicas do Câncer de Mama (BreakHis), construído em colaboração com o Laboratório de PD - Anatomia Patológica e Citopatologia, Paraná, Brasil. Para realizar deste trabalho, foram utilizadas cerca de 1693 imagens, contendo 547 amostras benignas e 1146 amostras malignas, com fatores de aumento de 400X, com 3 canais RGB com profundidade de 8 bits em cada canal, no formato PNG. Além disso, o conjunto de dados foi dividido em dados de treinamento e dados de teste em pastas diferentes.

2.2. Rede neural convolutacional

Uma Rede Neural Convolutacional é um algoritmo de aprendizado profundo utilizado no campo de visão computacional para reconhecimento de imagens. Segundo [Li et al. 2016], a utilização de redes neurais convolucionais trouxe uma série de resultados de pesquisas revolucionárias em diversas áreas. Tais como: classificação de imagens, detecção de objetos e segmentação semântica.

2.3. Arquitetura da CNN

2.3.1. Camadas de convolução

Em problemas de classificação de imagens, as entradas para uma rede neural convolutacional são, geralmente, matrizes tridimensionais que seguem as dimensões de cada imagem. Usualmente, as imagens contêm três canais RGB, contendo valores de cada pixel. Um Rede Neural Convolutacional é formada por camadas de convoluções, que funcionam

como filtros ponderadas por kernels, que têm o objetivo diminuir o tamanho de uma imagem e melhorar o processamento, buscando destacar as principais características da imagem. Neste trabalho, foram utilizadas duas camadas de convolução, com 64 filtros cada uma, utilizando um kernel (3,3) para que a rede não fosse considerada profunda e, ainda assim, pudesse extrair características sem que a quantidade de convoluções fosse um empecilho.

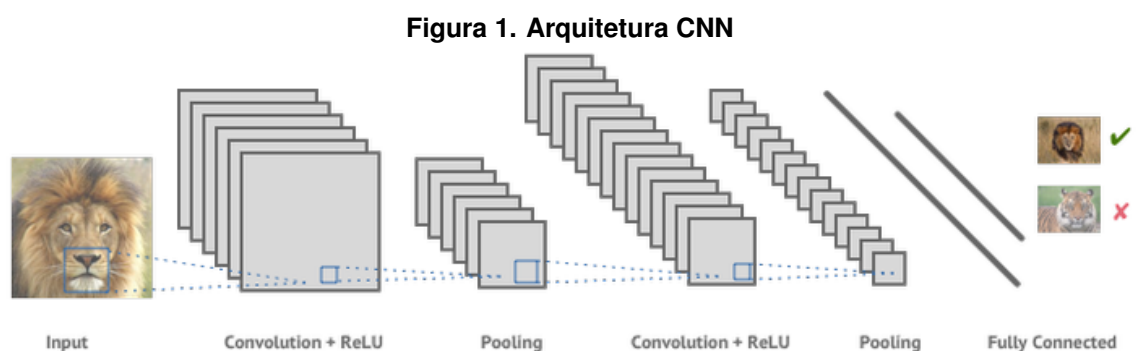
2.3.2. Função ReLU

Para oferecer não-linearidades ao algoritmo, são utilizadas as funções de ativação que dão suporte à rede no processo de aprendizado. Dessa forma, a função utilizada neste trabalho foi a função ReLU que, em Redes Neurais Convolucionais ou para aprendizado profundo é considerada a mais eficiente [Rizzo and Canato 2020], que retorna 1 ou 0 em problemas que envolvem classificação binária.

2.3.3. Pooling

A etapa de *pooling* tem o objetivo de enfatizar ambientes distintos de uma imagem, criando um mapa de características contribuindo para diminuir o *overfitting* e ruídos. Nesta etapa, foram utilizadas duas camadas de *pooling* com o método de *MaxPooling*, que obtém os valores máximos do mapa de características com uma unidade de área de (2,2).

Assim como na convolução, é escolhida uma unidade de área, por exemplo 2x2, para transitar por toda a saída da camada anterior. A unidade é responsável por resumir a informação daquela área em um único valor. Se a saída da camada anterior for 24x24, a saída do *pooling* será 12x12. Além disso, é preciso escolher como será feita a sumarização. O método mais utilizado é o *maxpooling*, no qual apenas o maior número da unidade é passado para a saída. Essa sumarização de dados serve para diminuir a quantidade de pesos a serem aprendidos e também para evitar *overfitting*. Por fim, foi utilizado uma camada de *Flatten* para modificar a matriz de uma imagem e alterar seu formato, transformando-a em um array.



Fonte: [e Library.com 2020]

2.4. Rede Neural Densa

Uma Rede Neural é um modelo computacional capaz de reconhecer padrões através de uma gama de dados, utilizando neurônios artificiais baseados no sistema nervoso humano que, MacCullon and Pitts (1943), define essa utilização como um ponto marcante na história das redes neurais.

Na implementação da rede neural artificial, foram utilizadas três camadas densas com 128 unidades cada. Ademais, para as duas primeiras camadas densas, a função de ativação definida foi a função ReLU. Além disso, como a presente problemática trata-se de uma classificação binária, a função de ativação sigmoid foi escolhida, por ser, uma função não-linear e apresenta rapidez no treinamento de redes neurais densas [Karlik and Olgac 2011].

Em alguns modelos de aprendizado de máquina, ocorre o problema de *overfitting*, que indica alta adaptação aos dados de treinamento e baixa generalização para dados de validação. Para reduzir esse tipo de ocorrência, existe uma técnica chamada *Dropout* que consiste na eliminação aleatória temporária de alguns neurônios durante a atualização dos pesos da rede no processo treinamento de redes neurais. Dessa forma, a rede neural densa utilizou duas camadas de dropout com 0.2 cada. Na etapa de compilação da rede são feitas algumas configurações, como: função de perda (*loss function*), otimizador (*optimizer*) e métricas de avaliação do modelo. Assim, foi definida como função de perda a *binary crossentropy* que é indicada para problemas de classificação binária, como otimizador o *adam*, pois este é utilizado como otimizador padrão e como métrica de avaliação a acurácia para analisar o desempenho da rede com relação aos dados de validação.

3. Resultados e discussão

O treinamento de uma rede recebe alguns valores, como: dados de treinamento, número de amostras por época, número de épocas, dados de validação e número de amostras para a validação. Os resultados obtidos com cinco épocas são mostrados na Figura 2:

Figura 2. Resultados do treinamento da rede.

Época	Loss	Accuracy	Validation Loss	Validation Accuracy
Época 1	0.1913	0.9246	0.9677	0.8277
Época 2	0.1363	0.9461	0.6397	0.8018
Época 3	0.1184	0.9545	0.531	0.7999
Época 4	0.0935	0.9660	1.2702	0.8020
Época 5	0.0795	0.9724	1.6310	0.7781

As três primeiras épocas tiveram valores de perda relativamente baixos: 0.9677, 0.6397 e 0.5931. Além disso, os valores das acurácias das três primeiras épocas foram, respectivamente: 0.8277, 0.8018 e 0.7999. Em análise, a perda de informação foi diminuindo e a acurácia de cada época também. Nas duas últimas épocas, a perda de informação pode ser considerada relativamente alta e 1.2702 e 1.6310, enquanto a acurácia foi de 0.8020 e 0.7781, o que pode ser considerada alta.

4. Considerações finais

Diante dos resultados obtidos, por meio do treinamento da rede, os objetivos deste trabalho foram alcançados. Uma vez que, a métrica de avaliação acurácia permaneceu acima da média durante o treinamento e validação do modelo. Além disso, há um ponto relevante que torna-se necessário ser destacado é o valor da perda de informação que foi aumentando ao longo da validação. Possivelmente, o modelo pode ter se encontrado na problemática de *overfitting*, quando o modelo adapta-se aos dados de treinamento e não consegue classificar de forma satisfatória os dados de teste.

Para trabalhos futuros, objetiva-se realizar uma análise comparativa de hiperparâmetros, afim de diminuir o valor da perda de informação e aumentar a quantidade de épocas, uma vez que esta ajusta os pesos da rede para diminuir a perda de informação. Além disso, objetiva-se também, construir modelos com outras arquiteturas de redes neurais convolucionais com maior quantidade de dados para construção do modelo.

Referências

- Albawi, S., Mohammed, T. A., and Al-Zawi, S. (2017). Understanding of a convolutional neural network. In *2017 International Conference on Engineering and Technology (ICET)*, pages 1–6. IEEE.
- Chollet, F. (2018). *Deep Learning mit Python und Keras: Das Praxis-Handbuch vom Entwickler der Keras-Bibliothek*. MITP-Verlags GmbH & Co. KG.
- de Câncer, I. N. (2002). Atlas de mortalidade por câncer no brasil 1979-1999.
- e Library.com (2020). Exemplo estrutura cnn.
- Karlik, B. and Olgac, A. V. (2011). Performance analysis of various activation functions in generalized mlp architectures of neural networks. *International Journal of Artificial Intelligence and Expert Systems*, 1(4):111–122.
- Li, Y., Hao, Z., and Lei, H. (2016). Survey of convolutional neural network. *Journal of Computer Applications*, 36(9):2508–2515.
- McCulloch, W. S. and Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, 5(4):115–133.
- Mendonça, G. A., Silva, A. M. d., Caula, W. M., et al. (2004). Características tumorais e sobrevida de cinco anos em pacientes com câncer de mama admitidas no instituto nacional de câncer, rio de janeiro, brasil. *Cadernos de Saúde Pública*, 20:1232–1239.
- Rizzo, I. V. and Canato, R. L. C. (2020). Inteligência artificial: Funções de ativação. *Prospectus*, 2(2):51–65.
- Santos, A. E. d. and Oguisso, T. (2007). Hospital sírio-libanês: história de sua criação e contribuição para a assistência à saúde. *Rev. paul. enferm*, pages 4–9.