

DESENVOLVIMENTO DE MODELO DE INTELIGÊNCIA ARTIFICIAL PARA CLASSIFICAÇÃO DE REGIÕES EM CÉLULAS DE SANGUE

Diego Oliveira Bomfim

Marlon Rodrigues Garcia

Escola de Engenharia de São Carlos – Universidade de São Paulo

oliveiradiego63@usp.br

marlongarcia@usp.br

Objetivos

O presente projeto de pesquisa tem por objetivo o desenvolvimento de modelos de classificação de regiões em células de glóbulos brancos, especialmente eosinófilos, linfócitos, monócitos e neutrófilos. Pretende-se desenvolver algoritmos capazes de pavimentar o caminho para sistemas automáticos de detecção de condições e doenças em lâminas histológicas, para o caso de esfregaço de sangue. Para tanto, será abordada a biblioteca de redes neurais, *PyTorch*, de código aberto construída na linguagem de programação *Python*. Serão utilizadas redes neurais convolucionais, como a *U-Net*, desenvolvida para segmentar regiões biológicas.

dataset utilizado consiste de 1854 imagens, sendo 927 imagens de lâminas histológicas das células fotografadas e outras 927 imagens as *labels* segmentadas de treinamento e validação. Nesta divisão, 690 foram utilizadas para teste e 237 para validação. O estudo das ferramentas e funcionalidades oferecidas pela biblioteca *PyTorch* foi de grande importância a fim de desenvolver uma rede neural com os parâmetros corretos para realizar a segmentação das células. O ambiente de execução foi o *Google Colaboratory (Google Colab)*, de caráter *online* e gratuito para execução de códigos *Python* na nuvem, escolhido por sua versatilidade, diversos recursos de trabalho e o oferecimento de GPU, importante para executar os algoritmos de classificação.

Métodos e Procedimentos

O desenvolvimento do método empregado se deu a partir do embasamento teórico obtido na revisão bibliográfica, destacam-se Bain¹ (2005) pela proposta de automação no diagnóstico de esfregaço de sangue, Ronneberger⁴ (2015) pelo uso da arquitetura *U-Net* na segmentação de imagens biomédicas, Niazi² (2019) e Patil³ (2013) com análises sobre as vantagens e limitações do uso de inteligência artificial no diagnóstico de doenças e patologias.

A construção do *dataset* de treinamento e validação da rede neural desenvolvida se deu empregando o conjunto de imagens de células de glóbulos brancos de domínio público do portal de *datasets* de saúde *Raabin*. Ao todo, o

Resultados

Na figura 1 pode ser visto, como exemplo, a segmentação de uma célula de monócito dada pela rede neural desenvolvida. Nela, as imagens **a)**, **b)** e **c)** são respectivamente a imagem da célula de monócito na lâmina histológica, sua máscara para treinamento e validação e a saída da rede neural segmentada. Na imagem de saída **c)**, podem ser identificadas as regiões segmentadas, a de cor preta como a área externa à célula, cor roxa como as paredes celulares e citoplasma e o branco como o núcleo celular. Com um total de 10 épocas de treinamento, a rede atingiu 98,91% de acurácia para classificação com 3,33% de perda. Para acompanhar a acurácia,

também foi adotado o coeficiente de *Sørensen-Dice* (conhecido popularmente como *Dice Score*) que além de verificar os pixels corretos que a rede segmenta, também penaliza pelos “falsos” corretos. Nesta execução, o *Dice Score* final foi de 95,37%. A Figura 2 representa o gráfico da performance da rede neural convolucional para este treinamento.

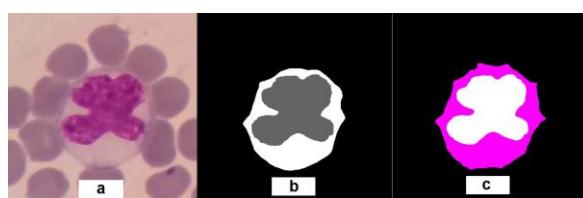


Figura 1: Segmentação de uma célula de monócito pela rede neural desenvolvida.

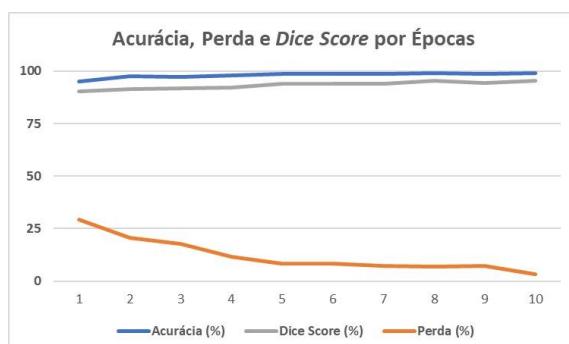


Figura 2: Gráfico da Acurácia, Perda e *Dice Score* por Épocas da rede neural desenvolvida.

Conclusões

O desenvolvimento proposto por este projeto de redes neurais convolucionais para a classificação de regiões de células de sangue concluiu em resultados satisfatórios que reforçam a eficácia e utilidade da arquitetura *U-Net* para a segmentação de imagens biológicas e biomédicas. Consequentemente, a acurácia e *Dice Score* respectivos de 98,91% e 95,37% obtidos com um *dataset* relativamente menor que a maioria das aplicações intensificam as expectativas de desempenho em situações de *datasets* maiores e com a análise ideal de épocas para treinamento, representando uma área promissora no âmbito de inspeção médica e pesquisa acadêmica.

A variedade de aplicações do material desenvolvido por esta pesquisa é muito vasta, a exemplo da análise do estado de saúde das células por meio de sua morfologia segmentada, a detecção da presença de parasitas e doenças relacionadas, e da contagem de células de um determinado tipo através da detecção e segmentação aqui apresentada.

Referências Bibliográficas

- [1] BAIN, Barbara J. Diagnosis from the blood smear. *New England Journal of Medicine*, v. 353, n. 5, p. 498-507, 2005.
- [2] NIAZI, M. K. K.; PARWANI, A. V.; GURCAN, M. N. Digital pathology and artificial intelligence. *The lancet oncology*, Elsevier, v. 20, n. 5, p. e253–e261, 2019.
- [3] PATIL, Dinesh D.; DEORE, Sonal G. Medical image segmentation: a review. *International Journal of Computer Science and Mobile Computing*, v. 2, n. 1, p. 22-27, 2013.
- [4] RONNEBERGER, O.; FISCHER, P.; BROX, T. U-net: Convolutional networks for biomedical image segmentation. In: SPRINGER. *International Conference on Medical image computing and computer-assisted intervention*. [S.I.], 2015. p. 234–241.