

Detecção Inteligente de Retinopatia Diabética com Redes Neurais Convolucionais Profundas

Diego Lucena de Medeiros, Elloá B. Guedes

Resumo— Este trabalho apresenta os resultados experimentais da utilização de redes neurais convolucionais aplicadas ao problema de detecção de retinopatia diabética em imagens médicas. Foram avaliadas arquiteturas canônicas e os melhores resultados obtidos possuem F_1 -Score igual a 75,83% com a arquitetura Inception, ressaltando o potencial do uso de tais modelos neste problema de Imagiologia Médica.

Palavras-Chave— Aprendizado Profundo; Redes Neurais Convolucionais; Visão Computacional.

Abstract— This work describes the experimental results of using Convolutional Neural Networks to detect diabetic retinopathy in medical images. Canonical architectures were considered and the best results obtained have F_1 -Score equal of 75.83% with Inception architecture, highlighting the potential of using such models in this Medical Imagiology problem.

Keywords— Deep Learning; Convolutional Neural Networks; Computer Vision.

I. INTRODUÇÃO

A Retinopatia Diabética (RD) é uma complicação crônica da perturbação metabólica chamada de Diabete Mellitus (DM), caracterizando-se por ser uma das principais causas de cegueira em certas populações portadoras de DM [1], [2]. Estima-se que, após 20 anos, quase todos pacientes com DM Tipo I, e cerca de 60% dos portadores de DM Tipo II, desenvolvem algum grau de RD [3].

Existem duas classificações principais para a RD: a forma não proliferativa (RDNP) e a proliferativa (RDP), sendo esta última uma evolução da primeira. RDNP pode manifestar-se respectivamente, em grau leve, moderado e avançado, em que pode então evoluir para a forma RDP [3]–[5]. Imagens de diferentes sujeitos sem retinopatia diabética e com RDNP em todos os seus graus e com RDP encontram-se ilustradas na Figura 1.

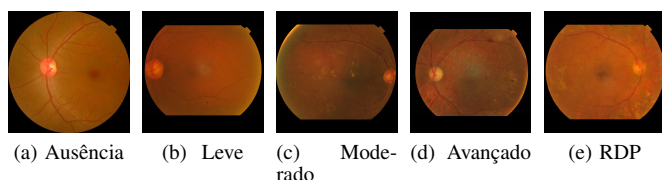


Fig. 1. Exemplos de imagens de retina saudáveis e com diversos graus de RD. Os graus leve, moderado e avançado correspondem à RDNP. Fonte: Aptos Challenge Dataset.

O diagnóstico de RD se dá em diversas formas e períodos após o diagnóstico de DM, mas quanto mais cedo for feito, menores são as chances de danos permanentes à visão [1], [4].

Grupo de Pesquisas em Sistemas Inteligentes, Laboratório de Sistemas Inteligentes, Escola Superior de Tecnologia, Universidade do Estado do Amazonas, Av. Darcy Vargas, 1200 – Manaus – Amazonas, {dldm.eng18, ebgcosta}@uea.edu.br

Considerando os resultados promissores do uso de técnicas de *Deep Learning* na Imagiologia Médica [6], [7], o presente trabalho apresenta resultados do uso de Redes Neurais Convolucionais na tarefa de Visão Computacional de detecção do grau de RD a partir de imagens de pacientes coletadas em um contexto real, na Índia [8].

Para apresentar os resultados obtidos, o restante deste artigo está organizado como segue: a metodologia adotada é apresentada na Seção II, os resultados obtidos são descritos e discutidos na Seção III e, por fim, a Seção IV contempla as considerações finais.

II. MATERIAIS E MÉTODOS

O problema de Visão Computacional considerado no escopo deste trabalho foi modelado como uma tarefa de classificação multiclasse mediante Aprendizado Supervisionado. Os dados experimentais consistem em 3.662 imagens coloridas oriundas da base de dados Aptos Challenge 2019, rotuladas previamente por especialistas [8]. O grupo de controle (ausência) e os quatro graus de manifestações de RD compõem as classes do problema, os quais estão distribuídos de maneira desbalanceada nos exemplos disponíveis, conforme Fig. 2.



Fig. 2. Distribuição dos exemplos por classe.

Os exemplos foram particionados em três conjuntos: treinamento, validação e teste, com proporções 70%, 10% e 20%, respectivamente, seguindo a abordagem *holdout* de validação cruzada. Os dados de treino passaram por um aumento artificial de dados para favorecer a captura de padrões durante o aprendizado dos modelos, sendo este aumento caracterizado por transformações do tipo espelhamento vertical e horizontal, aumento e cisalhamento.

O modelo de *Deep Learning* escolhido para este problema foram as Redes Neurais Convolucionais (CNNs, do inglês *Convolutional Neural Networks*), as quais têm se mostrado uma ferramenta eficaz para a tarefa de classificação de imagens, por possuírem boa capacidade de reconhecimento de padrões em dados de alta dimensionalidade [9], o que justifica a adoção deste modelo no contexto deste trabalho. As CNNs possuem organização inspirada no córtex visual de mamíferos e topologia hierárquica composta de camadas convolucionais, de amostragem e completamente conectadas. Para o ajuste de pesos em CNNs mediante o Aprendizado Supervisionado, aplica-se o algoritmo *backpropagation*, no qual três fatores colaboram para um bom aprendizado: interação esparsa, compartilhamento de parâmetros e representação equivariante [10].

TABELA I
RESULTADOS EXPERIMENTAIS.

	Função de Ativação	Acurácia	Precisão	Revocação	F ₁ -Score	Épocas	Early Stopping	Tempo (s)
LeNet	ReLU	73,35%	66,87%	73,35%	68,57%	150	30	1742
LeNet	Leaky ReLU	73,50%	70,31%	73,50%	70,51%	150	25	1776
AlexNet	ReLU	69,16%	57,99%	69,16%	62,13%	100	10	560
AlexNet	Leaky ReLU	62,13%	70,70%	62,13%	63,71%	100	25	2155
Inception	ReLU	78,74%	78,49%	78,74%	75,84%	150	20	1666
ResNet	ReLU	73,20%	68,98%	73,20%	67,85%	300	20	3795

A proposição de uma arquitetura de CNN é uma tarefa complexa, pois diz respeito à escolha de tipos e quantidades de camadas, sua disposição e definição de uma grande quantidade de parâmetros e hiperparâmetros [11]. Para contornar esta dificuldade, no escopo deste trabalho foram consideradas arquiteturas canônicas da literatura, a citar: LeNet [12], AlexNet [13], Inception (v3) [14] e ResNet-152 [15]. Nas duas primeiras arquiteturas foram consideradas as funções de ativação do tipo ReLU (*REctified Linear Unit*) e Leaky ReLU, enquanto as duas últimas foram consideradas tal como concebidas pelos seus proponentes.

A avaliação de desempenho das CNNs será realizada por meio da aferição da correteza das previsões realizadas no conjunto de testes perante as métricas de acurácia, precisão, revocação e F_1 -Score, calculadas considerando o desbalanceamento do conjunto de dados.

III. RESULTADOS E DISCUSSÃO

A metodologia proposta foi implementada com a linguagem de programação Python com o auxílio dos *frameworks* Scikit-learn, Keras, Tensorflow e Matplotlib. Os *scripts* foram executados em um servidor com processador Intel Core i7, 32 GB de memória principal, 960 GB de memória secundária e 2 placas de vídeo NVIDIA GTX 1080 Ti com 11 GB para aceleração em *hardware* do treinamento. Os resultados obtidos encontram-se dispostos na Tabela I, juntamente com as épocas de treinamento e o parâmetro de regularização para o *early stopping*.

Considerando o F_1 -Score como métrica de referência, observa-se que a melhor estratégia de classificação foi obtida com a arquitetura Inception. Ressalta-se o aprendizado automático de características como uma das vantagens dos modelos utilizados, sem a necessidade de etapas de pré-processamento. Observa-se ainda que apesar de possuir uma grande quantidade de parâmetros, o desempenho da ResNet no cenário experimental em questão não superou o de arquiteturas mais simples, como a LeNet, o que ressalta a não-trivialidade da tarefa considerada.

IV. CONSIDERAÇÕES FINAIS

Este trabalho apresentou os resultados experimentais do uso de CNNs no problema de detecção de RD em imagens médicas. Os resultados obtidos destacam a arquitetura Inception como tendo melhor performance no cenário experimental considerado. Embora promissores, os resultados aqui apresentados precisam ser sujeitos à outras técnicas de avaliação e validação antes de utilização em contextos reais, não se propondo neste momento a substituir especialistas médicos.

Em trabalhos futuros almeja-se explorar outras arquiteturas de CNNs bem como técnicas de otimização e regularização, o que pode colaborar para um melhor desempenho. Além disso, considera-se analisar os resultados obtidos em contraste com outras técnicas da literatura.

AGRADECIMENTOS

Os autores agradecem o apoio financeiro da FAPEAM e do CNPq por meio do Edital PPP 004/2017 e do Programa PAIC/FAPEAM/UEA 2020-2021.

REFERÊNCIAS

- [1] C. J. Flaxel, R. A. Adelman, S. T. Bailey, A. Fawzi, J. I. Lim, G. A. Vemulakonda, and G.-s. Ying, "Diabetic retinopathy preferred practice pattern," *Ophthalmology*, vol. 127, no. 1, pp. P66–P145, 2020.
- [2] X. Zhang, J. B. Saaddine, C.-F. Chou, M. F. Cotch, Y. J. Cheng, L. S. Geiss, E. W. Gregg, A. L. Albright, B. E. Klein, and R. Klein, "Prevalence of diabetic retinopathy in the united states, 2005-2008," *Jama*, vol. 304, no. 6, pp. 649–656, 2010.
- [3] D. S. Fong, L. Aiello, T. W. Gardner, G. B. George L King, J. D. Cavallerano, F. L. F. 3rd, and R. Klein, "Retinopathy in diabetes," *Diabetes Care*, vol. 27, no. 1, pp. S84–S87, 2004.
- [4] J. A. Pereira, M. A. T. Bertolin, G. D. C. Pereira, L. C. Corgozinho, L. A. da Matta Faria, and M. S. M. Pereira, "Atualizações sobre retinopatia diabética: uma revisão narrativa," *Revista Eletrônica Acervo Saúde*, no. 49, pp. e3428–e3428, 2020.
- [5] J. M. Vlisel and T. A. Oetting, "Diabetic Retinopathy for Medical Students. EyeRounds.org - Ophthalmology - The University of Iowa," 2010. [Online]. Available: <https://webeye.ophth.uiowa.edu/eyeforum/tutorials/Diabetic-Retinopathy-Med-Students/index.htm>
- [6] H. Greenspan, B. van Ginneken, and R. M. Summers, "Guest editorial Deep Learning in medical imaging: Overview and future promise of an exciting new technique," *IEEE Transactions on Medical Imaging*, vol. 35, no. 5, pp. 1153–1159, 2016.
- [7] J. G. Lee, S. Jun, Y. W. Cho, H. Lee, G. B. Kim, J. B. Seo, and N. Kim, "Deep learning in medical imaging: General overview," *Korean J. Radiol.*, vol. 18, no. 4, pp. 570–584, 2017.
- [8] A. P. T.-O. S. APTOS, "APTOS 2019 blindness detection," 2019, acesso em 26 de julho de 2021. [Online]. Available: <https://www.kaggle.com/c/aptos2019-blindness-detection>
- [9] S. Khan, H. Rahmani, S. A. A. Shah, and M. Bennamoun, *A Guide to Convolutional Neural Networks for Computer Vision*, 1st ed., ser. Synthesis Lectures on Computer Vision. San Rafael, California, USA: Morgan & Claypool, 2018.
- [10] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge: The MIT Press, 2016, vol. 1.
- [11] F. Chollet, *Deep Learning with Python*, 1st ed. Shelter Island, New York: Manning Publications, 2017.
- [12] Y. LeCun, L. Bottou, Y. Bengio, P. Haffner *et al.*, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.
- [13] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," 2012.
- [14] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, vol. 1. Boston, Massachusetts, USA: IEEE, 2015, pp. 1–9.
- [15] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.