

O uso de *Deep Learning* como Auxílio na Detecção de Doenças Visuais a partir de Exames OCT

Mateus Klein Lourenço¹, Gabriel Ramos de Souza²

¹Curso de ciências da computação – Universidade Presbiteriana Mackenzie (UPM)

10388729@mackenzista.com.br, 10419371@mackenzista.com.br

Abstract. *With the aging of the population, care for eye diseases may become increasingly urgent. A more common cause of irreversible loss of central vision is age-related macular degeneration (AMD), which occurs mainly in people over 50 years of age, and Several other illnesses are related to it, particularly macular edema and choroidal neovascularization, which when detected require medical referral and treatment urgently. Therefore, the use of neural networks has become a good strategy for automating medical diagnosis and also for quickly identifying these diseases.*

Resumo. *Com o envelhecimento da população os cuidados com doenças oculares podem se tornar cada vez mais urgente, uma causa mais comum de perda irreversível da visão central é a degeneração macular relacionada a idade (DMRI) que ocorre principalmente em pessoas com mais de 50 anos, e a ela estão relacionadas diversas outras enfermidades destacando principalmente o edema macular e a neovascularização coroidal que quando detectadas o encaminhamento médico e o tratamento precisa ser urgente. Através de um exame de OCT as doenças destacadas podem ser detectadas, e uma boa estratégia para a rápida identificação e encaminhamento médico pode ser através do treinamento de uma rede neural.*

1. Introdução

Devido ao avanço da medicina a expectativa de vida aumentou, com isso o envelhecimento da população se tornou uma realidade global. Com esse envelhecimento, os cuidados com a saúde oftalmológica vêm se tornando cada vez mais importantes, e uma doença que está relacionada a este cenário é a degeneração macular relacionada a idade (DMRI), que tem como causa muito comum quando não tratada a perda irreversível da visão central. Segundo Wong et al. (2014) em uma projeção sobre a prevalência da DMRI, estimava-se que 196 milhões de pessoas no mundo teriam a doença em 2020 e 288 milhões em 2040.

Conforme a BBC (2019) “Outro fator de risco importante é a idade. Para se ter uma ideia, mais de 82% de todas as pessoas cegas no mundo são maiores de 50 anos. A entidade pontua que, independentemente da classe social, a estimativa de perda total de visão é de 15 a 30 vezes maior em pessoas com mais de 80 anos do que nas com até 40 anos.”, além disso de acordo com dados da organização pan-americana da saúde globalmente estima-se que 1.3 bilhão de pessoas vivam com alguma forma de deficiência visual.

Há dois tipos de DMRI, o tipo seco (não exsudativa ou atrófica) e a úmida (neovascular ou exsudativa). Na DMRI não exsudativa a perda da visão central ocorre vagarosamente e sem dor ao longo do tempo, e com a progressão da mesma aparecem pontos cegos centrais que comprometem a visão, já na DMRI exsudativa a perda da visão central geralmente ocorre de maneira rápida, podendo haver a perda súbita caso um dos vasos se rompa. De acordo com Farah, Oshima, Costa e Sallum (2001), em seu estudo sobre "Degeneração Macular Relacionada à Idade: Modalidades Terapêuticas", quando a degeneração macular relacionada à idade apresenta uma complicação conhecida como membrana neovascular coroidal ou cicatriz disciforme, ela é classificada como "úmida" ou exsudativa. Isso ocorre em cerca de 15,3% dos casos. Apesar de ser menos comum do que a forma seca da doença, essa forma úmida é responsável por cerca de 80% dos casos de cegueira legal causada pela degeneração macular relacionada à idade.

Uma doença caracterizada pelo acúmulo anormal de fluído na mácula, e que também está relacionada a DMRI é o edema macular, mas a mesma também tem como causa comum o diabetes. Em pacientes diabéticos a cegueira resulta primariamente da progressão da retinopatia diabética e do edema macular clinicamente significativo (CSME, do inglês clinical significant macular edema) (POWERS, 2005).

Tudo isso explica o aumento do uso dos exames de OCT para avaliar o quadro do paciente, segundo o Instituto de Oftalmologia de Curitiba (2022) “A Tomografia de Coerência Óptica é capaz de detectar e diferenciar as camadas presentes na Retina, bem como mapear sua conformação e espessura. Essas medidas são imprescindíveis na detecção e acompanhamento de diversas doenças que podem afetar a Retina. As principais aplicações da OCT estão na avaliação de Degeneração Macular Relacionada à Idade e à Retinopatia Diabética. Outras Retinopatias (como a causada pela Hipertensão), bem como o Edema Macular e Descolamentos da Retina ou do Vítreo também podem ser avaliados pelo exame.”

A avaliação manual de um exame de Tomografia de Coerência Óptica (OCT) pode ser um processo demorado, já que requer análise minuciosa das imagens. Quando levamos em conta a urgência de um diagnóstico a partir desse exame, fica claro que a agilidade e eficiência nesse processo são de extrema importância. Uma vez que diagnósticos rápidos podem ser cruciais para o tratamento eficaz de condições oftalmológicas, a implementação de métodos que garantam a análise precisa das imagens de OCT de maneira mais rápida e eficiente torna-se uma prioridade para melhorar os resultados clínicos e a experiência do paciente.

Dado o que foi discutido anteriormente, o treinamento de uma rede neural para reconhecimento da patologia em exames de OCT surge como uma técnica promissora e inovadora. Ao aproveitar a capacidade das redes neurais convolucionais em aprender padrões complexos diretamente das imagens, essa abordagem pode oferecer uma análise precisa e automatizada das imagens oftalmológicas.

A proposta deste trabalho é utilizar um conjunto de imagens públicas de exames de OCT para treinar uma Rede Neural Convolucional (CNN). Essas imagens incluem casos normais e condições patológicas como DRUSAS, edema macular diabético e neovascularização coroidal. O objetivo é desenvolver um modelo preditivo altamente preciso, empregando as melhores técnicas disponíveis.

O emprego de técnicas e métodos de aumento de dados (*data augmentation*) é crucial para evitar o *overfitting* e promover a generalização eficaz do modelo. Essa abordagem envolve a aplicação de diversas transformações às imagens, como rotações, translações, *zoom*, espelhamento, entre outras. Ao diversificar o conjunto de dados de treinamento dessa maneira, a rede neural convolucional torna-se mais capaz de reconhecer e extrapolar padrões relevantes, resultando em um modelo mais robusto e preciso na detecção de imagens.

Por fim será realizada uma discussão sobre as métricas que serão empregadas para avaliar o desempenho da CNN. Estas incluem a acurácia, precisão, *F1-score*, matriz de confusão, curva ROC, entre outras. Essas métricas serão importantes para determinar a eficácia do modelo na resolução do problema em questão, proporcionando uma avaliação abrangente e detalhada do seu desempenho.

2. Descrição do problema

Este estudo aborda a identificação de doenças oftalmológicas específicas usando imagens de tomografia de coerência óptica (OCT) e redes neurais convolucionais (CNN). As condições examinadas incluem drusas, edema macular diabético (EMD) e neovascularização de coroide (NVC).

2.1 Drusas

Drusas são depósitos de proteínas e lipídios sob a retina, associadas ao envelhecimento e à degeneração macular relacionada à idade (DMRI). Existem dois tipos principais: drusas duras, pequenas e bem definidas, geralmente inofensivas, e drusas moles, maiores e de margens indefinidas, associadas a um maior risco de DMRI. O acúmulo de lipofusina nas células da retina, devido à luz e oxigênio, pode levar à formação de drusas, inflamação e potencial agravamento da DMRI.

2.2 Edema Macular Diabético (EMD)

O EMD ocorre quando a retinopatia diabética (RD) afeta a mácula devido à hiperglicemia persistente, sendo uma das principais causas de cegueira. Estudos mostram que a RD é a segunda maior causa de cegueira em países ocidentais (Motta, Coblenz, & Melo, 2008). A detecção precoce e o tratamento adequado podem prevenir ou minimizar a perda visual, destacando a importância de avançar no entendimento da fisiopatologia da RD e desenvolver novos métodos diagnósticos e terapêuticos.

2.3 Neovascularização de Coróide (NVC)

A NVC é caracterizada pelo surgimento de novos vasos sanguíneos anormais na coróide, frequentemente associada à forma úmida da DMRI. Esses vasos podem causar complicações visuais significativas, como sangramento e formação de cicatrizes, levando à perda permanente da visão. O diagnóstico precoce e o tratamento agressivo são essenciais para evitar complicações graves e preservar a função visual.

2.4 Tomografia de Coerência Óptica (OCT)

O OCT utiliza luz infravermelha para escanear a retina, produzindo imagens tridimensionais detalhadas sem usar radiação. É uma ferramenta essencial para diagnosticar e monitorar várias patologias oculares, permitindo diagnósticos precisos e tratamentos eficazes.

2.5 Redes Neurais Convolucionais (CNN)

As CNNs são especialmente projetadas para tarefas de processamento de imagens, utilizando camadas convolucionais para extrair características das imagens e camadas de *pooling* para reduzir a dimensionalidade. Estudos indicam que CNNs treinadas adequadamente podem identificar padrões complexos e características relevantes nas imagens de OCT, proporcionando precisão e eficiência significativas na detecção e diagnóstico precoce de condições médicas (Costa, 2007).

2.6 Trabalhos Relacionados

Diversos estudos demonstram o sucesso das CNNs em detectar patologias oftalmológicas em imagens de OCT. Kermany et al. (2018) utilizaram aprendizado por transferência para diagnosticar imagens de OCT, obtendo alta precisão, sensibilidade e especificidade. Schlegl et al. (2018) desenvolveram uma abordagem para detectar fluidos retinianos, mostrando alta correlação com leituras manuais. Lee et al. (2017) usaram a rede VGG16 para diferenciar entre imagens de OCT normais e associadas à DMRI, com métricas impressionantes de precisão, sensibilidade e especificidade. Srinivasan et al. (2014) utilizaram descritores de gradiente orientados (HOG) e SVM para classificar imagens de OCT, demonstrando alta eficiência, especialmente para DMRI e EMD.

2.7 Métricas de Avaliação

Para avaliar a eficácia das CNNs, serão usadas métricas como acurácia, precisão e matriz de confusão. A acurácia mede a proximidade das previsões com os valores reais. A precisão avalia a capacidade do modelo de identificar corretamente exemplos positivos. A matriz de confusão apresenta a frequência das classificações, permitindo uma análise detalhada do desempenho do modelo.

Em resumo, o uso de CNNs treinadas em imagens de OCT oferece um método promissor para a identificação precisa e eficiente de doenças oftalmológicas, potencialmente melhorando a saúde ocular e a qualidade de vida dos pacientes.

3. Metodologia

Para a realização deste trabalho, foi selecionado o *dataset* "Retinal OCT Images" disponibilizado por Paul Mooney obtido por meio de Exames de Tomografia de Coerência Óptica (OCT). Esse *dataset* contém quatro classes de doenças visuais (normal, drusas, DME e CNV) e é composto por 83.484 imagens para treino, 968 para teste e 32 para validação.

No conjunto de treino, o *dataset* possui 26.315 imagens normais, 8.616 de drusas, 11.348 de DME e 37.205 de CNV. Para o conjunto de teste, há 242 imagens de cada classe. Já no conjunto de validação, temos 8 imagens para cada classe. Este balanceamento é fundamental para garantir que o modelo não seja tendencioso em relação a nenhuma classe específica.

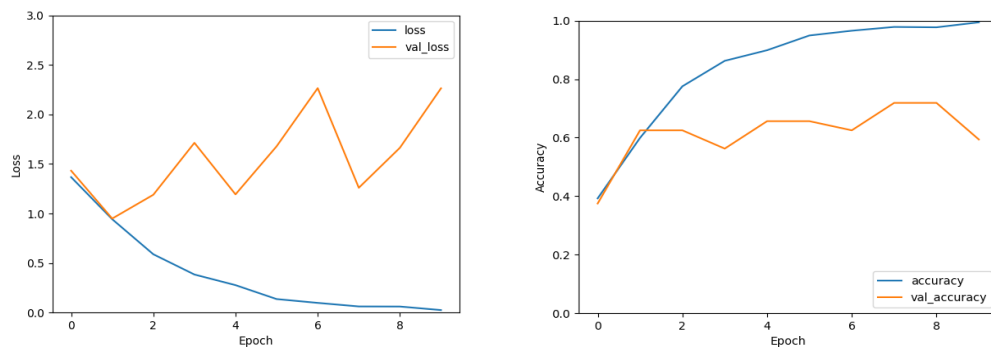
Todas as imagens presentes no *dataset* são do tipo .jpeg e estão em escala de cinza (1 canal), onde cada pixel é representado por um único valor que reflete a intensidade da luz, variando do preto ao branco. As dimensões das imagens variam em termos de largura e altura, o que torna essencial o pré-processamento. Este processo, que será abordado no próximo tópico, foi realizado para padronizar as dimensões de todas as imagens, garantindo consistência e compatibilidade com o modelo de treinamento.

O pré-processamento é fundamental para o desenvolvimento de um modelo de *deep learning*, garantindo seu desempenho adequado. Neste estudo, as imagens do *dataset*, apresentando dimensões variadas, foram padronizadas para 512x496 pixels para assegurar consistência nos dados de entrada. Utilizando um script em *Python* com as bibliotecas "os" e "PIL", as imagens foram redimensionadas e cortadas a partir da região central, que contém informações relevantes para a análise de OCT, como a mácula. Além disso, para uma futura melhora do modelo pode ser aplicadas técnicas de *data augmentation*, como rotação e translação, para enriquecer o conjunto de dados de treinamento, aumentando a robustez e a capacidade de generalização do modelo na detecção de doenças oculares.

Com o auxílio da biblioteca *TensorFlow*, foram desenvolvidos três modelos preliminares de CNNs em *Python*. O primeiro modelo foi treinado utilizando 2.400 imagens, sendo 600 para cada classe. O segundo modelo, por sua vez, foi treinado com um conjunto maior, contendo 7.600 imagens no total, com 1.900 imagens para cada classe. E por fim o terceiro modelo foi treinado com 11.200 imagens, com 2.800 imagens para cada classe.

4. Resultados

O primeiro modelo alcançou uma acurácia de teste de 0,5335. Durante as 10 épocas de treinamento, foram calculadas tanto a acurácia quanto a perda (*categorical crossentropy*). A *categorical crossentropy* avalia as diferenças entre a distribuição prevista pelo modelo e a distribuição verdadeira (real). Além disso, para os dados de validação, foi gerada a matriz de confusão, bem como a acurácia e a precisão para cada classe.



Fig

Figura 1. Gráficos de perda e acurácia ao longo das 10 épocas para o modelo 1

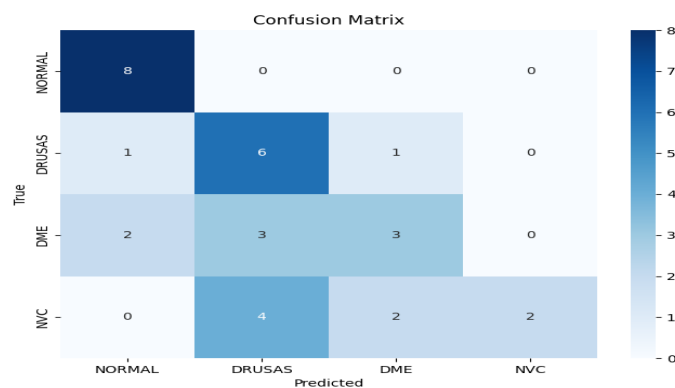


Figura 2. Matriz de confusão para os dados de validação para o modelo 1

O segundo modelo alcançou uma acurácia de teste de 0.61. De maneira similar ao primeiro modelo, foram calculadas as métricas de acurácia e perda ao longo de 10 épocas de treinamento. Adicionalmente, foi gerada a matriz de confusão para os dados de validação, juntamente com as métricas de acurácia e precisão para cada classe específica.

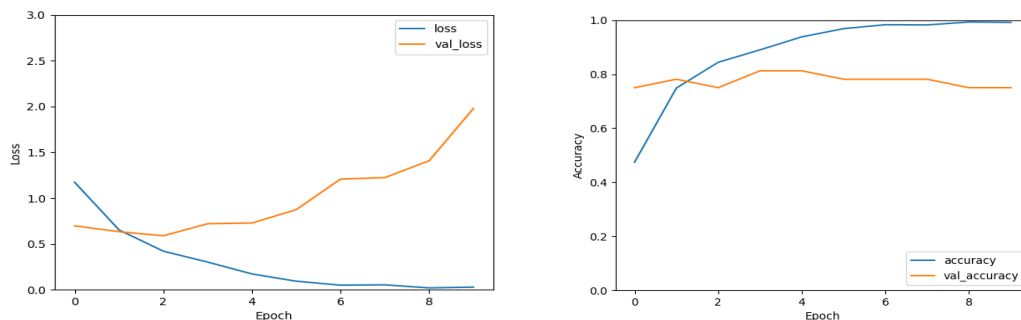


Figura 3. Gráficos de perda e acurácia ao longo das 10 épocas para o modelo 2

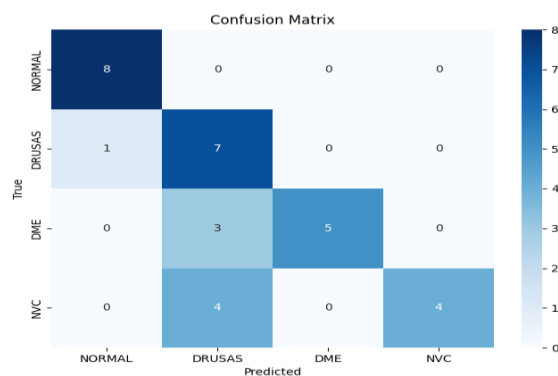


Figura 4. Matriz de confusão para os dados de validação para o modelo 2

O terceiro modelo obteve uma acurácia de teste de 0.78. Da mesma maneira dos outros dois modelos anteriores foram calculadas as métricas de acurácia e perda ao longo de 10 épocas de treinamento. E também foi gerada a matriz de confusão para os dados de validação, juntamente com as métricas de acurácia e precisão para cada classe específica.

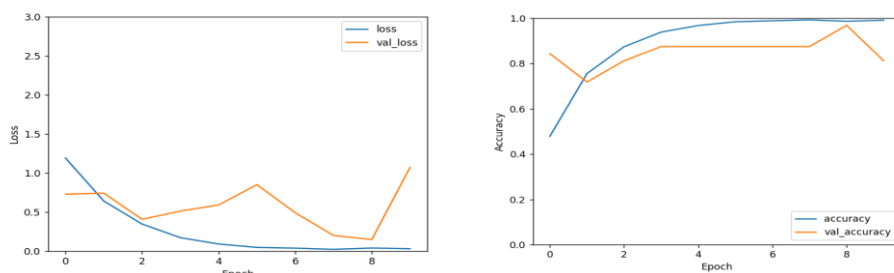


Figura 5. Gráficos de perda e acurácia ao longo das 10 épocas para o modelo 3

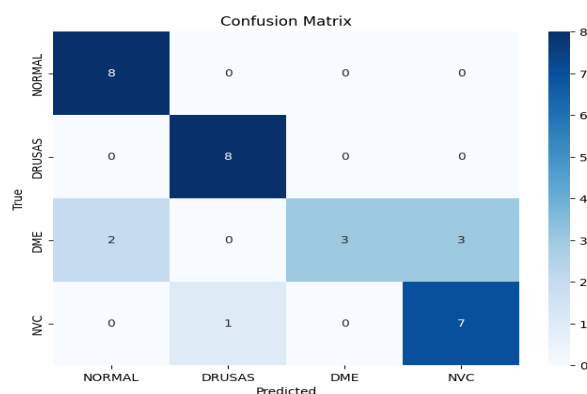


Figura 6. Matriz de confusão para os dados de validação para o modelo 3

5. Conclusão

Ao analisar o gráfico de perda dos modelos, notamos que a curva azul (perda de treinamento) diminui consistentemente ao longo das épocas, indicando que o modelo está aprendendo e ajustando-se melhor aos dados de treinamento. Este é um sinal positivo de que o modelo está convergindo adequadamente. No entanto, a curva laranja,

que representa a perda de validação, inicialmente diminui, mas começa a aumentar após algumas épocas, sugerindo que o modelo pode estar superajustando (*overfitting*) aos dados de treinamento.

Um aspecto importante é que, à medida que aumentamos a quantidade de imagens, a perda de validação ao final das épocas diminui consideravelmente, apesar de ainda aumentar após certo ponto. Isso indica que o aumento no volume de dados contribui para uma melhor generalização do modelo. Além disso, a implementação de técnicas de *data augmentation*, que introduzem maior diversidade ao conjunto de treinamento, pode reduzir ainda mais a perda de validação. Com mais imagens e *data augmentation*, podemos esperar um modelo mais robusto, preciso e confiável na detecção de doenças oculares em imagens de OCT.

Na validação do Modelo 1, obtivemos uma acurácia de 0.59, com precisões de 0.72, 0.46, 0.5 e 1 para as classes 0, 1, 2 e 3 (normal, drusas, DME e CNV), respectivamente. No Modelo 2, a acurácia aumentou para 0.75, com precisões de 0.88, 0.5, 1 e 1 para as mesmas classes. Finalmente, o Modelo 3 alcançou uma acurácia de 0.8125, com precisões de 0.8, 0.88, 1 e 0.7 para as respectivas classes. Esses resultados demonstram que o aumento do número de imagens contribuiu significativamente para a melhoria das métricas da CNN, especialmente na sensibilidade de cada classe, reduzindo o número de falsos positivos.

6. Referências

COSTA, M. G. F. Redes neurais convolucionais na saúde. *Journal of Health Informatics*, v. 9, n. 4, 2017. Disponível em: <https://jhi.sbis.org.br/index.php/jhi-sbis/article/view/615>. Acesso em: 11 mai. 2024.

FARAH, M. E.; OSHIMA, A.; COSTA, R. A.; SALLUM, J. F. Degeneração macular relacionada à idade: modalidades terapêuticas. *SciELO*, 2001. Disponível em: <https://www.scielo.br/j/abo/a/xX4P3LbmbkdtjMNq467f59j/?lang=pt>. Acesso em: 05 mai. 2024.

INSTITUTO DE OFTALMOLOGIA DE CURITIBA. Exame OCT: Tomografia de Coerência Óptica. 2022. Disponível em: <https://www.ioc.med.br/blog/exame-oct-tomografia-de-coerencia-optica>. Acesso em: 10 mai. 2024.

KERNY, D. S. et al. Identifying Medical Diagnoses and Treatable Diseases by Image-Based Deep Learning. *Cell*, v. 172, n. 5, p. 1122-1131, 2018.

LEE, C. S.; BAUGHMAN, D. M.; LEE, A. Y. Deep learning is effective for the classification of OCT images of normal versus age-related macular degeneration. *Ophthalmology Retina*, v. 1, n. 4, p. 322-327, jul./ago. 2017. doi: 10.1016/j.oret.2016.12.009. PMID: 30693348; PMCID: PMC6347658.

MOONEY, P. Retinal OCT Images (optical coherence tomography) [Data set]. Kaggle, 2018. Disponível em: <https://www.kaggle.com/datasets/paultimothymooney/kermany2018>. Acesso em: 17 mai. 2024.

MOTTA, M. M. dos S.; COBLENTZ, J.; MELO, L. G. N. de. Aspectos atuais na fisiopatologia do edema macular diabético. *Revista Brasileira de Oftalmologia*, v. 67, n. 1, p. 45–9, jan. 2008. Disponível em: <https://doi.org/10.1590/S0034-72802008000100009>. Acesso em: 09 mai. 2024.

POWERS, A. Diabetes Mellitus. In: KASPER, et al. *Harrison's Principles of Internal Medicine*. 16. ed. New York: McGraw-Hill, 2005. cap. 323, p. 2152-2180.

SCHLEGL, Thomas, et al. Fully automated detection and quantification of macular fluid in OCT using deep learning. *American Journal of Ophthalmology*, v. 125, n. 4, p. 549-558, 2018. Acesso em: 17 mai. 2024.

SRINIVASAN, Pratul P., et al. Detecção totalmente automatizada de edema macular diabético e degeneração macular seca relacionada à idade de imagens de tomografia de coerência óptica. *Biomedical Optics Express*, v. 5, p. 3568-3577, 2014.

WONG, W. L., et al. Global prevalence of age-related macular degeneration and disease burden projection for 2020 and 2040: a systematic review and meta-analysis. *The Lancet Global Health*, v. 2, n. 2, p. e106-116, 2014. doi: 10.1016/S2214-109X(13)70145-1. PMID: 25104651.