CLASSIFICAÇÃO DE IMAGENS DE EDEMA MACULAR DIABÉTICO USANDO REDES NEURAIS CONVOLUCIONAIS COM TRANSFERÊNCIA DE APRENDIZADO

Gilberto Luis De Conto Junior[1], Ionildo José Sanches[2]

[1]*Universidade Federal Tecnológica do Paraná, Ponta Grossa, Brasil*
gilbertoluis@utfpr.edu.br
[2]*Universidade Federal Tecnológica do Paraná, Ponta Grossa, Brasil*
ijsanches@utfpr.edu.br

Resumo: Este trabalho tem por objetivo a utilização de um modelo pré-treinado de rede neural convolucional, com transferência de aprendizado, para classificar edema macular diabético em imagens de tomografia de coerência óptica. A rede foi treinada com o modelo ResNet-50 e adicionou camadas *Dropout*, *Flatten* e *Dense* com ativação sigmoidal. Foram utilizadas 22.696 imagens para treinamento, divididas igualmente entre exames normais e com edema macular. A acurácia obtida foi de 96,40%, com precisão de 96,77%, *recall* de 96,0%, *F-measure* de 96,39% e especificidade de 96,80%.

Palavras-chave: edema macular diabético; tomografia de coerência óptica; processamento de imagens; rede neural convolucional; transferência de aprendizado.

INTRODUÇÃO

O Diabetes Mellitus (DM) é uma doença caracterizada pela elevação da glicose no sangue (hiperglicemia). Esta doença se manifesta em dois tipos, o tipo 1 no qual o pâncreas não produz insulina ou produz muito pouco e o tipo 2 ocorre quando o organismo se torna resistente à insulina ou não produz insulina suficiente. Estima-se que uma em cada dez pessoas com idade entre 20 e 79 anos, têm diabetes. Isso equivale a aproximadamente 537 milhões de pessoas no mundo segundo o Atlas do Diabetes da Federação Internacional de Diabetes (*International Diabetes Federation* - IDF) (IDF, 2021).

Esta doença traz uma série de complicações para o portador e dentre elas a Retinopatia Diabética que de acordo com o Instituto Nacional do Olho (*National Eye Institute* - NEI) (NEI, 2021), pode causar cegueira ao afetar os vasos sanguíneos da retina provocando vazamentos nesta área acarretando em pontos de visão turva (onde o fluido vazado se armazena) e manchas escuras. É uma doença que por si só pode causar severos danos à saúde ocular do portador e em seus estágios mais avançados pode levar à outras complicações como descolamento de retina, glaucoma neovascular e edema macular diabético. A doença acomete uma em cada quinze pessoas diabéticas e é a principal causa de perda de visão.

O Edema Macular Diabético (*Diabetic Macular Edema* - DME) ocorre quando vazamentos de fluidos da retina atingem a região da mácula causando uma visão borrada no paciente (NEI, 2021). Para o diagnóstico do DME são comumente usados exames como a Angiofluoresceinografia (exame que avalia o fluxo dos vasos da retina através de fotografias do fundo do olho), e principalmente a Tomografia de Coerência Óptica ou OCT (*Optical Coherence Tomography*). Na OCT um aparelho provê imagens detalhadas das camadas da retina para que o especialista avalie a incidência da doença verificando a presença ou não de fluidos na região macular (AAO, 2020).

O processamento digital de imagens envolve técnicas de transformação de imagens visando melhorar suas características visuais para a interpretação humana e o processamento de cenas para percepção de máquinas (visão computacional) (Gonzalez e Woods 2008).

As redes neurais convolucionais que são uma das mais importantes estruturas no campo da *deep learning*, tem atraído muita atenção da indústria e da academia nos últimos anos (Li *et al*., 2022). Esta estrutura é um tipo de rede neural útil especialmente em tarefas relacionadas a imagens como classificação e segmentação semântica (Wu, 2017). Uma forma de aprimorar os resultados de uma rede neural convolucional é utilizar o método de transferência de aprendizado. Este procedimento consiste na transferência dos recursos aprendidos por uma rede neural aplicada a outra base de dados em uma nova resultando na melhoria da capacidade de

generalização da nova rede neural (Fawaz *et al*., 2018).

Para o auxílio na classificação das imagens se propõe neste trabalho a aplicação de rede neural convolucional em uma base de exames de OCT com imagens normais e com edema macular diabético. Inicialmente aplica-se técnicas de pré-processamento visando eliminar ruídos para em seguida ser treinada utilizando um modelo de rede neural pré-treinado com transferência de aprendizado. Após o modelo treinado são mensuradas métricas de desempenho como precisão, acurácia, especificidade, *recall* e *F-measure* utilizando as imagens da base de teste para avaliar como o modelo se comportou na classificação.

MATERIAL E MÉTODOS

A Figura 1 apresenta a metodologia proposta para este trabalho. O primeiro passo é a aquisição das imagens de OCT da base de dados. Após ser delimitada as imagens da base de dados, aplica-se um filtro gaussiano que visa suavizar as imagens de modo que a rede consiga performar melhor.

Com as imagens devidamente processadas, o próximo passo é realizar o *fine tunning* do modelo neural com o modelo pré-treinado ResNet-50. Após inserido os pesos da rede pré-treinada e as camadas específicas são realizados o treinamento da rede e a classificação das imagens de teste em duas classes.

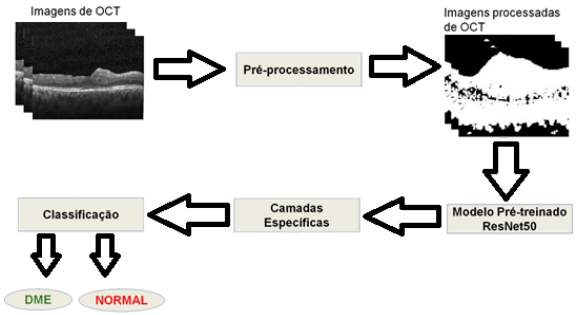


Figura 1. Metodologia proposta.

A base de dados para este trabalho foi obtida de Kermany, Zhang e Goldbaum (2018). A base possui 83.484 imagens divididas entre as seguintes classes: “normal”, “edema macular diabético”, “neovascularização coroidal” e “drusas”. A Figura 2 apresenta exemplos de imagens do *dataset.* Pode-se perceber a diferença entre uma tomografia de retina acometida pelo DME (Figura 2(a)) e outra normal (Figura 2(b)).

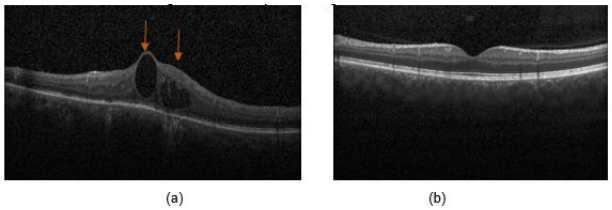


Figura 2. Exemplos de tomografia de coerência óptica (a) Imagem com edema macular diabético; (b) imagem normal.

Como o objetivo deste trabalho é diferenciar os exames que apresentam o edema macular diabético dos normais, foram selecionadas todas as 11.348 imagens do conjunto ‘DME’ mais 11.348 imagens aleatórias do conjunto ‘Normal’ para realizar o treinamento. Para a etapa de teste foram mantidos os dois conjuntos ‘DME’ e ‘Normal’ como aparecem no *dataset* onde cada classe possui um total de 250 imagens. Não foi utilizado o método de *data augmentation*.

Neste experimento, utilizou-se a técnica de transferência de aprendizado para melhorar o desempenho de uma rede neural, aproveitando *features* aprendidas por uma rede pré-existente. O modelo pré-treinado escolhido foi o ResNet-50, que possui 50 camadas e é mais preciso do que os modelos ResNets com menos camadas. Para utilizar o modelo ResNet, foram selecionados os pesos da base de dados ImageNet, que conta com aproximadamente 14 milhões de imagens. As camadas mais gerais do modelo original foram mantidas e apenas as últimas camadas foram manipuladas para abordar as especificidades deste experimento, como o formato de saída. Foi inserida uma camada de *Dropout*, para evitar *overfitting*, uma camada *Flatten*, para transformar o formato da imagem em formato linear, e uma camada Densa, com ativação sigmoidal, como camada de saída para a classificação binária. O otimizador utilizado foi o Adam e a taxa de aprendizado foi definida em 0,001.

No desenvolvimento do projeto foi utilizada a linguagem de programação Python, a biblioteca OpenCV (*Open Source Computer Vision Library*) e o MatLab para manipulação das imagens e o *framework* de redes neurais *Keras*.

RESULTADOS E DISCUSSÃO

O treinamento da rede culminou em um modelo neural. Uma vez obtido o modelo é possível submete-lo a imagens externas (que não fizeram parte do processo de treinamento) para avaliar o desempenho real da rede. Neste caso a base de dados utilizada

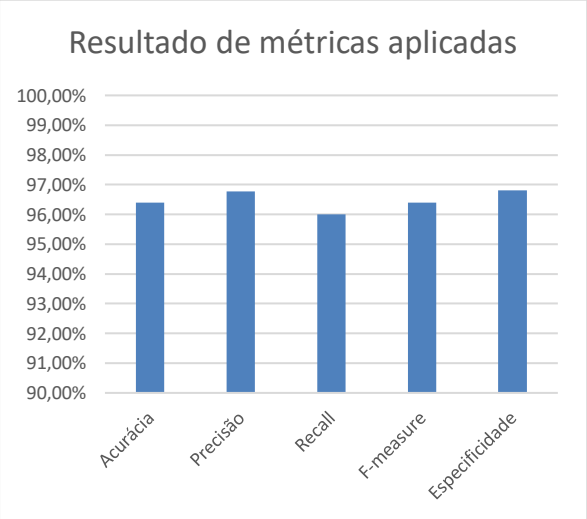
possui 250 imagens de teste para cada classe, estas portanto, não foram incluídas no treinamento. O Quadro 1 apresenta a matriz de confusão obtida após a aplicação do modelo na base de teste.

Quadro 1. Matriz de confusão

Classificação	Real "DME"	Real "Normal"
Predito "DME"	240 (VP)	8 (FN)
Predito "Normal"	10 (FP)	242 (VN)

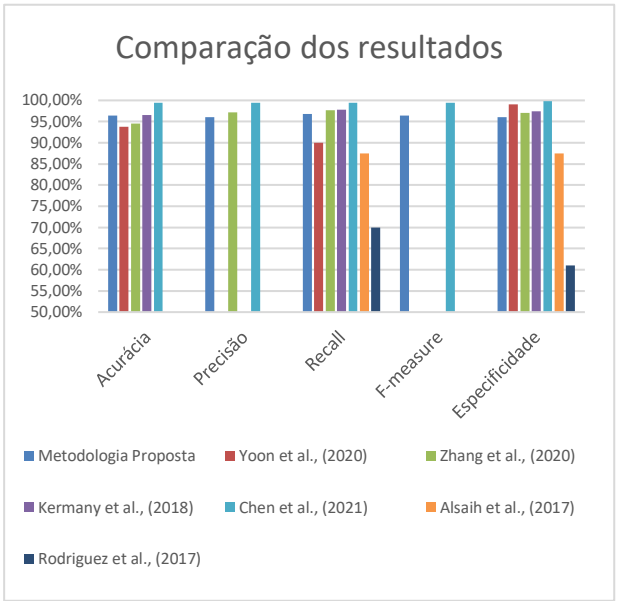
Observa-se no Quadro 1 que apenas 18 imagens foram preditas incorretamente, sendo 8 imagens da classe normal que obtiveram a predição de 'DME' e 10 imagens da classe DME que obtiveram predição 'normal'. A partir desta matriz foi aplicado algumas métricas de desempenho obtendo os resultados mostrados no Gráfico 1.

Gráfico 1. Comparação dos resultados.



O Gráfico 1 mostra que após a aplicação das métricas de desempenho na matriz de confusão obteve-se uma acurácia de 96,40%, uma precisão de 96,77%, *recall* de 96,00%, *F-measure* 96,39% e especificidade de 96,80%. O Gráfico 2 apresenta uma comparação destes resultados com os resultados disponíveis em outros trabalhos relacionados.

Gráfico 2. Comparação dos resultados.



No Gráfico 2 pode-se observar que em alguns trabalhos como o de Chen *et al.*, (2021) apresentaram um resultado superior próximo de 100%, enquanto outros mantiveram uma pequena margem de diferença em algumas métricas.

De qualquer forma um resultado preliminar acima de 95% é um bom indicador da capacidade que estas técnicas têm de fornecer um resultado muito conciso sobre um paciente ter ou não edema macular diabético.

Durante o desenvolvimento algumas alternativas foram testadas como redução do *Learning Rate*, alteração do otimizador para outras opções como SGD (*Stochastic Gradient Descent*) e RMSProp (*Root Mean Squared Propagation*) e aumento e diminuição de camadas e neurônios, porém sem sucesso. A rede permaneceu no estado o qual foi obtido a maior acurácia dentre o que foi testado. Observou-se que a ResNet-50 se comporta muito bem com estes parâmetros para o problema de identificação de edema macular, tendo acertado grande maioria de suas predições de forma similar ao que é abordado pela literatura.

Os próximos passos serão realizar testes aplicando outras técnicas de pré-processamentos nas imagens de treino e também a utilização de outros modelos pré-treinados como o *Inception* V3, *AlexNet*, *GoogleNet* e VGG-19. Dessa forma, será possível gerar uma comparação entre a eficácia de cada rede para assim determinar as configurações que apresentem os melhores resultados nas classificações de imagens de edema macular diabético.

CONCLUSÃO

A proposta do trabalho é avaliar a utilização de redes neurais convolucionais, com transferência de aprendizado. Com o auxílio de modelos pré-treinados e métodos de pré-processamento obteve-se uma acurácia de 96,40% na classificação das imagens de exames de tomografia de coerência óptica contendo exames normais e exames que apresentam DME. O *dataset* inicial foi delimitado em 11.348 imagens de cada classe para evitar que a rede ficasse com viés para a classe de maior número de imagens.

REFERÊNCIAS

Alsaih, K., Lemaitre, G., Rastgoo, M., Massich, J., Sidibé, D., & Meriaudeau, F. Machine learning techniques for diabetic macular edema (DME) classification on SD-OCT images. Biomedical engineering online, 16(68), 1-12, 2017.

American Academy of Ophthalmology (AAO). Macular Edema Diagnosis. National Eye Institute. Retrieved from https://www.nei.nih.gov/learn-about-eye-health/eye-conditions-and-diseases/diabetic-retinopathy, 2020.

Chen, Y. M., Huang, W. T., Ho, W. H., & Tsai, J. T. Classification of age-related macular degeneration using convolutional-neural-network-based transfer learning. BMC Bioinformatics, 22(Suppl 5, No. 99), 2021.

Fawaz, H. I., Forestier, G., Weber, J., Idoumghar, L., & Muller, P. A. Transfer learning for time series classification. IRIMAS, Universite de Haute Alsace, Mulhouse, France. arXiv:1811.01533v1 [cs.LG], 2018.

Gonzalez, R. C., & Woods, R. E. Processamento de imagens digitais. 3ª edição, Editora Pearson, 2008.

IDF Diabetes Atlas. International Diabetes Federation, 10th Edition Committee. Retrieved from https://diabetesatlas.org/idfawp/resource-files/2021/07/IDF_Atlas_10th_Edition_2021.pdf, 2021.

Kermany, D., Zhang, K., & Goldbaum, M. Labeled optical coherence tomography (OCT) and Chest X-Ray Images for Classification. Mendeley Data, V2, 2018.

Li, Z., Yang, W., Peng, S., & Liu, F. A survey of convolutional neural networks: analysis, applications, and prospects. IEEE Transactions on Neural Networks and Learning Systems, 33(12), 6999-7019, 2022.

National Eye Institute (NEI). At a glance: Diabetic Retinopathy. National Eye Institute. Retrieved from https://www.nei.nih.gov/learn-about-eye-health/eye-conditions-and-diseases/diabetic-retinopathy, 2021.

National Eye Institute (NEI). At a glance: Macular Edema. National Eye Institute. Retrieved from https://www.nei.nih.gov/learn-about-eye-health/eye-conditions-and-diseases/macular-edema, 2019.

Rodriguez, F. J. J., Perdomo, O., Gomez, F. E., Rosenstiehl, S. M., & Rios, H. A. Neural network for detection of diabetic macular edema in fundus color images. Investigative Ophthalmology & Visual Science, 58(8), 688-688, 2017.

Wu, Jianxin. Introduction to convolutional neural networks. National Key Lab for Novel Software Technology, 2017.

Yoon, J., Han, J., Park, J. I., Hwang, J. S., Han, J. M., Sohn, J., Park, K. H., & Hwang, D. D. Optical coherence tomography-based deep-learning model for detecting central serous chorioretinopathy. Scientific Reports, 10(18852), 1-9, 2020.

Zhang, Q., Liu, Z., Li, J., & Liu, G. Identifying diabetic macular edema and other retinal diseases by optical coherence tomography image and multiscale deep learning. Diabetes, metabolic syndrome and obesity: targets and therapy, 13, 4787-4800, 2020.