

# Análise comparativa de técnicas de Machine Learning para classificação de imagens de fundo de olho

Rosana Aurélio de Jesus

## Abstract

Glaucoma é uma doença crônica que pode causar cegueira irreversível se não tratada precocemente. Segundo a Organização Mundial da Saúde (OMS) 60,5 milhões de pessoas no mundo já foram afetadas pela doença e estima-se que 80 milhões de pessoas serão afetadas em 2020 e 111,5 milhões em 2040. Atualmente os médicos especialistas fazem o diagnóstico através de análise da imagem de fundo de olho do paciente, porém esta é uma avaliação subjetiva e que pode requerer um tempo considerável, mesmo de um médico experiente. Sendo assim, faz-se necessário automatizar o diagnóstico de glaucoma de forma que a ferramenta possa auxiliar a avaliação do médico. O trabalho proposto utilizou duas técnicas de machine learning para classificar se uma imagem é de fundo de olho ou não e posteriormente a ferramenta será evoluída para detecção de glaucoma na imagem de fundo de olho. Os algoritmos utilizados foram deep learning e Convolutional Neural Network (CNN). Para o problema proposto o algoritmo deep learning apresentou menor tempo de processamento e ambos obtiveram bom valor acurácia, chegando próximo a 1.

O glaucoma é uma doença ocular crônica caracterizada pela lesão das células do nervo óptico, sendo inicialmente assintomática na maioria dos casos (KRIŽAJ, 2019). Quando a pessoa começa perceber os sintomas a doença já está em estágios mais avançados. Estima-se que aproximadamente 50% dos acometidos não sabem que são portadores da doença. O glaucoma compromete primeiramente a visão periférica e se não tratada, o campo visual vai se estreitando, transformando-se numa visão tubular, até a perda total da visão. As pessoas com idade entre 40 e 80 anos são as mais afetadas. Os principais fatores de risco são: histórico familiar, pressão intraocular elevada, idade acima de 50 anos, diabetes, uso prolongado de corticoide, presença de lesões oculares, espessura da córnea e origem étnica. Existem várias técnicas para detectar se o paciente está com suspeita de glaucoma, como a tonometria, que mede a pressão intraocular, a gonioscopia, que verifica se o ângulo entre a íris e a córnea está aberto ou fechado, a tomografia de coerência óptica (OCT), que é um método não invasivo que permite a

realização de cortes transversais da retina, gerando imagens tomográficas de alta resolução e fundoscopia ou imagem de fundo de olho, que consiste em examinar as artérias, veias e nervos da retina através dos meios transparentes do olho (Carrillo et al., 2019). A imagem de fundo de olho é uma das técnicas não invasivas mais utilizadas pelos oftalmologistas. Devido ao diagnóstico subjetivo feito pelos médicos, pelo tempo considerável de profissionais experientes ao avaliar várias imagens e por haver regiões onde o número de especialistas é pequena, como em zonas rurais, várias pesquisas estão sendo realizadas para automatizar os diagnósticos e assim auxiliar os médicos na detecção da doença em seu estágio inicial. Este trabalho possui um escopo reduzido em relação ao detector de glaucoma, pois o propósito dos algoritmos de machine learning utilizados será o de classificar se uma imagem é de fundo de olho ou não. Tendo em vista que o classificador de glaucoma espera receber uma imagem de fundo de olho, preferencialmente somente com a região de interesse (nervo óptico), mostra-se válido garantir que somente imagens de fundo de olho constam na base de dados. Será utilizado o classificador deep learning por ser um algoritmo que vem apresentando bons resultados no reconhecimento de imagens.

## Revisão Bibliográfica

A imagem de fundo de olho, mostrada na figura 1, é utilizada para medir a espessura da camada de fibra nervosa da retina (RNFL). É uma das técnicas não invasivas mais utilizadas pelos oftalmologistas.



Figura 1: Imagem de Fundo de Olho

A espessura da RNFL é calculada medindo a proporção entre o tamanho da região do disco óptico (OD) e a escavação do disco óptico (CDO), que é uma região mais interna

do disco, e na delimitação das bordas dessas duas regiões. Um dos indicadores da presença de glaucoma é o aumento do CDO. Esse aumento pode ser percebido através da relação corpo para disco (CDR), que é a razão entre o DO e CDO (Ahmad et al., 2014). A figura 2 mostra um exemplo da obtenção do CDR, onde estão destacados o OD e CDO.

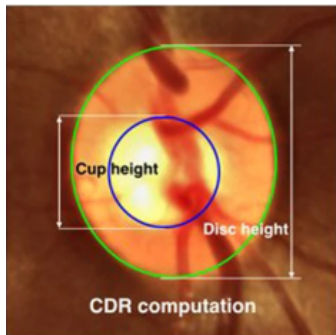


Figura 2: Regiões para obter o CDR

A técnica de aprendizagem profunda - Deep Learning - é uma subárea das técnicas de aprendizado de máquina e vem ganhando espaço para realizar processamento e classificação de imagens. Existem muitos fatores que favorecem a utilização dessas técnicas em áreas como visão computacional e processamento de linguagem natural, como o fato de não chegar a um ponto de saturação com relação a quantidade de dados. A arquitetura das Deep Learning é formada pela composição de múltiplas transformações lineares e não-lineares dos dados com o objetivo de produzir representações mais abstratas e mais úteis. Redes neurais convolucionais (CNNs) são arquiteturas de aprendizagem profunda, e recentemente foram empregadas com sucesso para tarefas de segmentação e classificação de imagens por levar em consideração sua estrutura espacial. Ela subdivide os dados para tentar extrair características de cada conjunto. A arquitetura CNN geralmente é constituída de três camadas: camada convolucional (CL), camada de pooling (PL), e totalmente conectadas (camada de classificação). A figura 3 mostra um exemplo da arquitetura de uma CNN com quatro classes possíveis como saída.

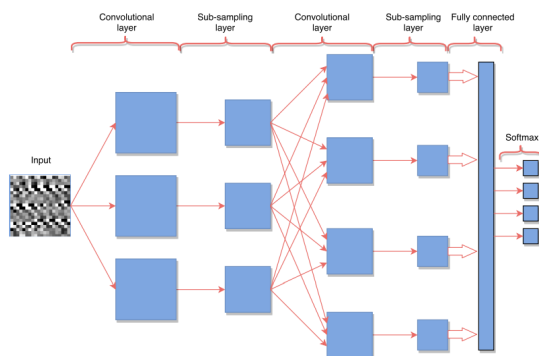


Figura 3: Camadas da arquitetura CNN

Panse em 2015 combinou conceitos de processamento de

imagens e mineração de dados onde características são extraídas das imagens e são utilizadas como entrada de um classificador de imagens (Panse; Ghorpade; Jethani, 2015). Foram utilizadas algoritmos de machine learning para para detectar doença na retina, como glaucoma e Retinopatia diabética (RD). O sistema proposto classifica a imagem como normal, com glaucoma e com RD. Foram utilizados os classificadores KNN (K-Nearest Neighbors), SVM (Support Vector Machine), Naive Bayes e RBF (Radial-Basis Function) usando weka para classificação das doenças. O estudo apresentou bons resultados comparados com trabalhos anteriores, sendo o algoritmo KNN o que obteve melhor resultado. Kim desenvolveu em seu trabalho um detector de glaucoma utilizando 100 casos de dados como teste e 399 casos para treinar os algoritmos de machine learning que utilizou para processar a imagem e verificar se a mesma tem glaucoma: C5.0, Random Forest (RF), SVM and KNN. O modelo RF obteve o melhor desempenho e C5.0, SVM e KNN mostraram desempenhos semelhantes (KIM; CHO; OH, 2017). Em 2015 Chen realizou uma classificação entre normal e glaucoma usando CNN (Convolutional Neural Network) em (Chen et al., 2015). A arquitetura utilizada em seu trabalho pode ser observada na figura 4. Chen projetou o AlexNet-style (KRIZHEVSKY; SUTSKEVER; HINTON, 2012) CNN, avaliou com as bases de dados ORIGA (ZHANG et al., 2010) e SCES (SNG LI-LIAN FOO; AUNG., 2012) e obteve 0.831 e 0.887 de AUC (area under the ROC curve), respectivamente.

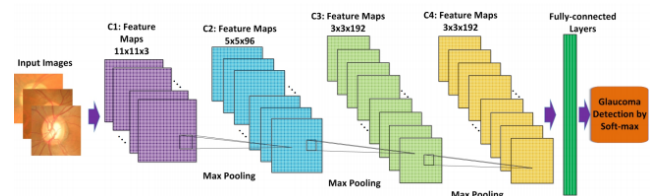


Figura 4: Arquitetura CNN do classificador

## Metodologia

Os classificadores receberão como entrada imagens de fundo de olho e imagens da parte externa do olho humano (FUSEK, 2018), exemplificada na figura 5. As imagens foram normalizadas e colocadas em escala de cinza antes da classificação.



Figura 5: Imagem externa olho

As bases de dados com imagens de fundo de olho utilizadas foram: DRIONS-DB (Digital Retinal Images for Op-

tic Nerve Segmentation Database) (DRIONS-DB, 2009), DRIVE (Digital Retinal Images for Vessel Extraction) (STAAL M.D. ABRAMOFF, 2004), BinRushed e Magrabi (ALMAZROA et al., 2018). No total, foram utilizadas 1000 imagens, onde 700 foram utilizadas para treino e o restante para teste. As técnicas de machine learning utilizadas foram Deep Learning e Deep Learning CNN. O algoritmo Deep Learning utilizado possui quatro camadas escondidas e duas saídas possíveis ('fundo de olho' ou 'outro'). Foram utilizadas a função de ativação 'relu' na primeira camada e nas camadas escondidas e 'softmax' na camada de saída. Foram implementados dois algoritmos CNN. O primeiro possui 12 camadas convolucionais, onde a cada quatro camadas foi colocada uma de pooling. Na camada de classificação (totalmente conectada), foram utilizadas as funções de ativação 'relu' e 'softmax'. O segundo possui quatro camadas convolucionais, onde cada uma é seguida de uma camada de pooling. Na camada de classificação foram utilizadas as funções de ativação 'relu' e 'softmax'. Para reduzir a chance de ocorrer overfitting e melhorar o desempenho da rede, foi utilizada, nos três algoritmos, a técnica de dropout, que consiste em desativar aleatoriamente neurônios com probabilidade  $p$  de ocorrência durante a fase de treinamento e da fase de propagação dos dados pela rede.

## Resultados

Foram realizados testes para 10 e 100 épocas para verificar o desempenho dos algoritmos.

- Configuração do algoritmo Deep Learning para 10 e 100 épocas:
  - batch size: 63
  - kernel size: (3,3)
  - Percentual para validação = 10%
- Configuração da CNN 1 para 10 épocas:
  - batch size: 14
  - kernel size: (3,3)
  - Percentual para validação = 10%
- Configuração da CNN 1 para 100 épocas:
  - batch size: 35
  - kernel size: (3,3)
  - Percentual para validação = 10%
- Configuração da CNN 2 para 10 e 100 épocas:
  - batch size: 14
  - kernel size: (5,5)
  - Percentual para validação = 10%

No CNN 1 o tamanho do batch para as épocas ficou diferente, porque no caso de 100 épocas o mesmo obteve melhor resultado com 35 amostras do que com 14.

Como resultado dos testes para 10 épocas, o algoritmo Deep Learning obteve acurácia de 1.0, o CNN 1 obteve acurácia de 0.8833 e o CNN 2 obteve 0.9967 de acurácia.

Como resultado dos testes para 100 épocas, o algoritmo Deep Learning e CNN 2 obtiveram acurácia de 1.0 e o CNN 1 obteve acurácia de 0.79.

O algoritmo deep learning conseguiu alcançar acurácia próxima de 1.0 durante o treinamento, como mostra figura 6, sem variar o tamanho do batch. A figura 7 mostra o gráfico de perdas no treinamento durante as 10 épocas.

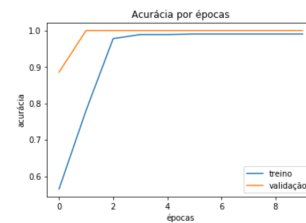


Figura 6: Treinamento Deep Learning para 10 épocas

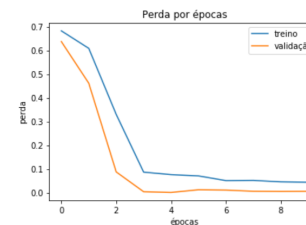


Figura 7: Perdas no treinamento Deep Learning para 10 épocas

A figura 8 mostra o desempenho do algoritmo Deep Learning nas 100 épocas durante o treinamento e a figura 9 mostra como foi a perda:

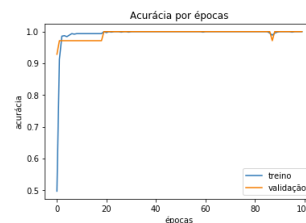


Figura 8: Treinamento Deep Learning para 100 épocas

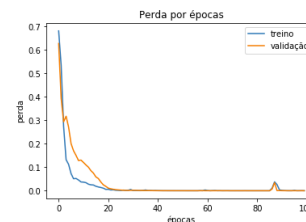


Figura 9: Perdas no treinamento Deep Learning para 100 épocas

A figura 10 mostra o desempenho do CNN 1 durante 10 épocas e a figura 11 mostra a perda do CNN 1 durante o treinamento com 10 épocas.

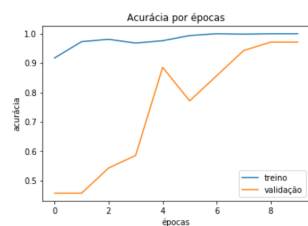


Figura 10: Treinamento CNN 1 para 10 épocas

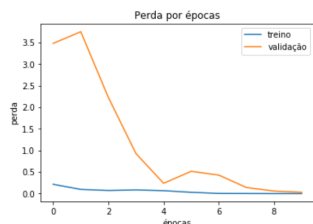


Figura 11: Perdas no treinamento CNN 1 para 10 épocas

A figura 12 mostra o desempenho do CNN 1 durante 100 épocas e a figura 13 mostra a perda do CNN 1 durante o treinamento com 100 épocas.

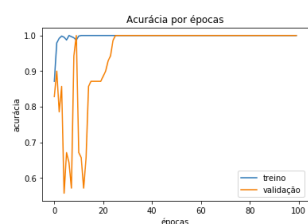


Figura 12: Treinamento CNN 1 para 100 épocas

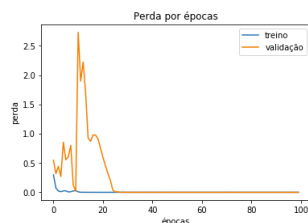


Figura 13: Perdas no treinamento CNN 1 para 100 épocas

A figura 14 mostra o desempenho do CNN 2 durante 10 épocas e a figura 15 mostra a perda do CNN 2 durante o treinamento com 10 épocas.

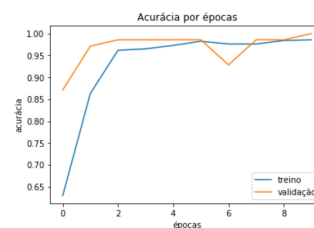


Figura 14: Treinamento CNN 2 para 10 épocas

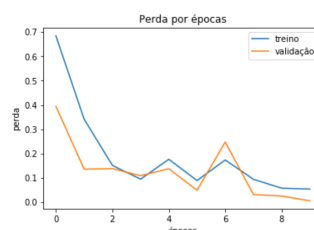


Figura 15: Perdas no treinamento CNN 2 para 10 épocas

A figura 16 mostra o desempenho do CNN 2 durante 100 épocas e a figura 17 mostra a perda do CNN 2 durante o treinamento com 100 épocas.

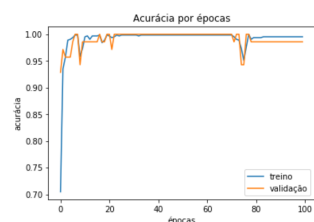


Figura 16: Treinamento CNN 2 para 100 épocas

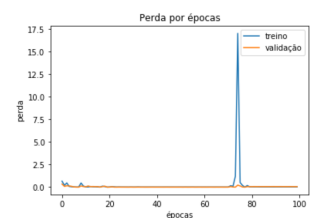


Figura 17: Perdas no treinamento CNN 2 para 100 épocas

## Conclusão

O algoritmo Deep Learning alcançou 1.0 de acurácia nos dois casos (10 e 100 épocas). O CNN 1 teve acurácia melhor para 10, que obteve 0.8833 épocas do que para 100, onde obteve 0.79 de acurácia. O CNN 2 teve acurácia 0.9967 para 10 épocas obteve 1.0 de acurácia para 100 épocas. Os algoritmos CNN levaram um tempo maior para realizar o treinamento, o que deve-se à maior complexidade de sua arquitetura. Como trabalho futuro, será implementado um

classificador que irá receber imagens de fundo de olho e diagnosticar se a mesma possui ou não padrões de glaucoma. Será utilizado o algoritmo CNN, uma base de dados maior para treinamento e teste e serão pesquisadas outras formas para otimização dos parâmetros, hiperparâmetros e arquitetura da rede, para realizar a classificação da forma mais otimizada.

ZHANG, Zhuo et al. Origa(-light): An online retinal fundus image database for glaucoma analysis and research. *Conference proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference*, v. 2010, p. 3065–8, 08 2010.

## References

- Ahmad, H. et al. Detection of glaucoma using retinal fundus images. p. 321–324, April 2014.
- ALMAZROA, Ahmed et al. Retinal fundus images for glaucoma analysis: the riga dataset. In: . [S.l.: s.n.], 2018. p. 8.
- Carrillo, J. et al. Glaucoma detection using fundus images of the eye. p. 1–4, April 2019. ISSN 2329-6259.
- Chen, X. et al. Glaucoma detection based on deep convolutional neural network. p. 715–718, Aug 2015. ISSN 1094-687X.
- DRIONS-DB. *Digital Retinal Images for Optic Nerve Segmentation Database*. 2009. Disponível em: <[http://www.ia.uned.es/~ejcarmona/DRIONS-DB/\\_2.html#Ref](http://www.ia.uned.es/~ejcarmona/DRIONS-DB/_2.html#Ref)>. Acesso em: 19 set. 2019.
- FUSEK, R. Pupil localization using geodesic distance. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, v. 11241 LNCS, p. 433–444, 2018.
- KIM, Seong Jae; CHO, Kyong Jin; OH, Sejong. Development of machine learning models for diagnosis of glaucoma. *PLOS ONE*, Public Library of Science, v. 12, n. 5, p. 1–16, 05 2017. Disponível em: <<https://doi.org/10.1371/journal.pone.0177726>>.
- KRIZHEVSKY, Alex; SUTSKEVER, Ilya; HINTON, Geoffrey E. Imagenet classification with deep convolutional neural networks. *Neural Information Processing Systems*, v. 25, 01 2012.
- KRIŽAJ, Ph.D. David. *What is glaucoma?* University of Utah Health Sciences Center, 2019. Disponível em: <<https://www.ncbi.nlm.nih.gov/books/NBK543075/>>.
- Panase, N. D.; Ghorpade, T.; Jethani, V. Retinal fundus diseases diagnosis using image mining. In: *2015 International Conference on Computer, Communication and Control (IC4)*. [s.n.], 2015. p. 1–4. Disponível em: <<http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7375721&isnumber=7374772>>.
- SNG LI-LIAN FOO, Ching-Yu Cheng John C Allen Mingguang He Gita Krishnaswamy Monisha E Nongpiur David S Friedman Tien Y Wong Chelvin C; AUNG., Tin. Determinants of anterior chamber depth: the singapore chinese eye study. p. 1143–1150, 2012. Disponível em: <<http://oar.a-star.edu.sg:80/jspui/handle/123456789/1060>>.
- STAAL M.D. ABRAMOFF, M. Niemeijer M.A. Viergever B. van Ginneken J.J. Ridge based vessel segmentation in color images of the retina. v. 23, p. 501–509, 2004. Disponível em: <<https://www.isi.uu.nl/Research/Databases/DRIVE/>>.