

Detecção De Glaucoma Usando Redes Em Cápsula

João Marcello M. Moreira, João Dallyson S. de Almeida¹, Geraldo Braz Junior¹,
Anselmo Cardoso de Paiva¹

¹Núcleo de Computação Aplicada (NCA) – Universidade Federal do Maranhão
(UFMA) Caixa Postal 65.080-805 – São Luís – MA – Brasil

joao.marcello@discente.ufma.br, jdallyson@nca.ufma.br,
geraldobraz@ufma.br, anselmo.paiva@ufma.br

Abstract. *Glaucoma is a disease that degrades the optic nerve causing irreversible loss of vision and whose symptoms usually appear in advanced stages. The leading resource against glaucoma is its early diagnosis, which shows the importance of computer-aided diagnostic systems, as these provide a second opinion to the specialist, helping in the correct diagnosis. Thus, this work presents a computational method for detecting glaucoma using capsule networks. The tests were performed by joining eight public databases of images from glaucoma exams, totaling 2,086 images. Our method obtained 85.29% accuracy, 81.77% sensitivity, and 88.28% specificity.*

Resumo. *O glaucoma é uma doença que degrada o nervo óptico causando perda irreversível da visão e cujos sintomas costumam aparecer somente em estágios avançados da doença. O principal recurso contra o glaucoma é o seu diagnóstico precoce, o qual mostra a importância dos sistemas de diagnóstico assistido por computador, visto que estes provêm uma segunda opinião ao especialista, ajudando no correto diagnóstico. Assim, este trabalho apresenta um método computacional para detecção de glaucoma utilizando redes em cápsula. Os testes foram realizados juntando-se oito bases públicas de imagens de exames de glaucoma, totalizando 2.086 imagens. Nosso método obteve 85,29% de acurácia, 81,77% de sensibilidade e 88,28% de especificidade.*

1. Introdução

O termo glaucoma engloba um conjunto de doenças ópticas com inúmeras causas, porém com características comuns, que formam um quadro clínico semelhante e que levam à perda total e irreversível da visão. São elas: o aumento da pressão intraocular, perdas campimétricas, aumento da escavação e atrofia do nervo óptico (Lauretti et al., 1997). Comumente assintomáticos, os portadores de glaucoma só manifestam os sintomas em estágios avançados da doença o que demonstra a importância de um diagnóstico precoce.

Entre a população global portadora de cegueira até o ano de 2015, considerando todas as idades (cerca de 36 milhões de pessoas), o glaucoma encontrava-se na terceira posição no ranking das principais causas de cegueira no mundo e apresentou, do ano de 1990 a 2015, um aumento de 2,5 milhões para 3 milhões no número de portadores de cegueira (Flaxman et al., 2017). Nos últimos anos, vem crescendo o número de pesquisas envolvendo diagnósticos assistido por computador (CAD) que são ferramentas que realizam a detecção de anomalias em exames de forma automatizada e servem como uma segunda opinião para o especialista (Silva, 2018). Nesse contexto, surgem as redes

neurais convolucionais (CNN), atualmente o estado da arte em classificação de imagens e utilizadas em vários trabalhos e em diferentes arquiteturas, entre elas VGG, ResNet e U-Net. Contudo, Sabour et al. (2017) apontou problemas em relação às arquiteturas tradicionais de *deep learning*. Segundo os autores, há perda significativa de informações por conta da organização sequencial de múltiplas camadas presentes nessas arquiteturas. Ainda, a presença das camadas de *pooling*, que reduzem a dimensionalidade do volume de entrada, descartam informações importantes que deveriam ser propagadas pela rede. Por fim, várias arquiteturas de CNN são muito dependentes das características espaciais dos dados utilizados para treinamento, fazendo com que rotações ou translações destes dados façam com que sejam classificados incorretamente. Assim, Sabour et al. (2017) propuseram uma nova arquitetura afim de superar as limitações das arquiteturas convencionais, chamada de Capsnet, que, entre outras particularidades, possui camadas contendo cápsulas, que são agrupamentos de neurônios que simbolizam um vetor de n dimensões, e que agregam várias propriedades particulares de uma característica presente na imagem. Os autores obtiveram o estado da arte na base MNIST para reconhecimento de dígitos manuscritos.

Assim, esta pesquisa busca avaliar a utilização de redes em cápsula para detecção de glaucoma em retinografias, contendo como principal contribuição a construção de um modelo treinado para classificar o glaucoma em imagens provenientes de *datasets* e equipamentos de aquisição diferentes.

2. Trabalhos Relacionados

Esta seção apresenta alguns trabalhos recentes disponíveis na literatura que utilizaram *deep learning* para a detecção do glaucoma.

Li et al, (2018) utilizaram em seu trabalho uma rede neural convolucional com a arquitetura Inception-V3 com 22 camadas, com o objetivo de detectar glaucoma em retinografias. Usando 31.745 imagens para treino e 8000 para validação da base online privada *LabelMe*, eles obtiveram uma acurácia de 98,6%, com sensibilidade de 95,6% e especificidade de 92,2%.

O método apresentado por Raghavendra et al., (2018) utiliza uma rede neural convolucional com 18 camadas treinada em 70% de uma base com 589 imagens de pacientes saudáveis e 837 de pacientes com glaucoma obtida da Faculdade de Medicina Kasturba, Índia, alcançando 98,13% de acurácia, 98% de sensibilidade e 98,30% de especificidade.

Com o objetivo de escolher a melhor arquitetura de CNN que seja capaz de auxiliar no diagnóstico do glaucoma através de técnicas evolutivas, Lima (2019) utilizou a base RIM-ONE versão 2, conseguindo 96,63% de acurácia, 94,87% de sensibilidade, 98% de especificidade e 97,37% de precisão.

Utilizando a arquitetura LeNet e a base RIM-ONE versão 2 para detecção de glaucoma, Pereira (2017) conseguiu 88,57% de acurácia, com o desvio padrão de 1,79%. O método apresentado por Reis et al. (2018) utiliza o *Cup-to-Disc-Ratio* (CDR), que é a razão entre os diâmetros verticais do disco óptico e de sua cavidade, a arquitetura de CNN U-Net e a base pública DRISHTI-GS1, obtendo a acurácia média de 88,7%, sensibilidade média de 94% e especificidade média de 53,3% na detecção do glaucoma.

O método desenvolvido por Claro et al. (2019) teve como objetivo determinar o melhor conjunto de características para representação de imagens. Para tal, utilizou o descritor de textura GLCM em conjunto com redes neurais pré-treinadas e as bases de dados RIM-ONE, HRF, ACRIMA, DRISHTI e JSIEC num total de 1675 imagens, obtendo 93,61% de acurácia.

3. Redes Em Cápsula

Idealizadas por Sabour et al. (2017), as redes em cápsula são redes neurais que possuem como principal singularidade a presença de camadas contendo cápsulas, que são agrupamentos de neurônios que representam um vetor que define parâmetros de instanciação (por exemplo, cor e textura) e de pose (como posição, orientação e escala) para as diferentes partes de um objeto (Moraes, 2018). Surgiram da necessidade de superar problemas comuns existentes nas arquiteturas convencionais de CNN, entre os quais, a presença das camadas de *pooling*, que descartam a informação sobre a exata localização da característica detectada o que faz a rede ser incapaz de reconhecer objetos quando estes são sujeitos a rotações ou outros tipos de modificações (Afshar et al, 2019). O módulo do vetor de saída de uma cápsula representa a probabilidade que a característica representada por essa cápsula esteja presente no dado de entrada. Para tal, é utilizada uma nova função de ativação chamada de “*squashing*” que faz com que o vetor tenha seu módulo ajustado para 0 ou 1 de modo que vetores com módulos curtos fiquem próximos de zero e vetores longos fiquem próximos de 1. A Equação 1 mostra como é calculada a função de achatamento (*squashing*).

$$v_j = \frac{\|s_j\|^2}{1 + \|s_j\|^2} \frac{s_j}{\|s_j\|} \quad (1)$$

Onde v_j é a probabilidade da presença da característica da cápsula j e s_j é a soma total da entrada. O vetor de saída de uma cápsula é enviado para a cápsula apropriada da camada seguinte em um processo chamado de “*routing-by-agreement*” (que substitui a necessidade das camadas de *pooling*), no qual a cápsula computa um “vetor de previsão” multiplicando sua própria saída com uma matriz de pesos. Assim, se uma cápsula de nível baixo “concorda” com uma cápsula de nível alto, forma-se uma relação de “parte-todo” em que as cápsulas de nível baixo detectam a presença de parte de uma característica enquanto as de nível alto detectam o todo (Patrick et al, 2019).

A arquitetura de rede em cápsula utilizada foi baseada no trabalho de Sabour et al. (2017), e possui duas partes: *encoder* e *decoder*. O *encoder* (Figura 1) é composto por 3 camadas convolucionais cuja função de ativação, tamanho e quantidade dos filtros são otimizáveis e são responsáveis por extrair características da entrada. A camada *PrimaryCaps* é formada por 32 cápsulas 8D enquanto que na *DigitCaps* são 2 cápsulas 16D. O algoritmo “*routing-by-agreement*” foi utilizado entre as *PrimaryCaps* e as *DigitCaps*. O *decoder* (Figura 2) tenta reconstruir a entrada a partir das cápsulas finais, o que força a rede a preservar o máximo possível de informações da entrada através da rede. Isto funciona como um regularizador que reduz o risco de *overfitting* e ajuda na generalização para novos dados (Mobiny et al, 2018).

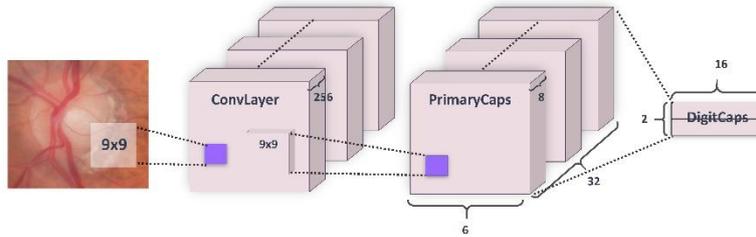


Figura 1. Arquitetura da Capsnet baseado em (Sabour et al, 2017).

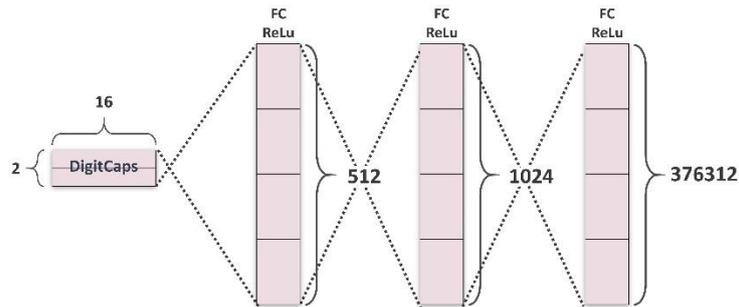


Figura 2. Arquitetura da Capsnet – Decoder (Sabour et al, 2017).

4. Materiais e Método

Esta seção apresenta os materiais utilizados no trabalho que inclui, as bases de dados, *softwares* e o método proposto. Utilizou-se da biblioteca *Keras* (usando como *backend* o *Tensorflow*), que permite uma fácil e rápida implementação de algoritmos de aprendizado de máquina. Para a execução dos experimentos, foi utilizado um computador com placa de vídeo Nvidia GeForce GTX 1660Ti e 8 GB de RAM. O método está organizado em quatro etapas (Figura 3): (1) Aquisição de imagens, (2) Pré-processamento, (3) Definição e treinamento da Capsnet, e (4) Classificação e Validação dos resultados.

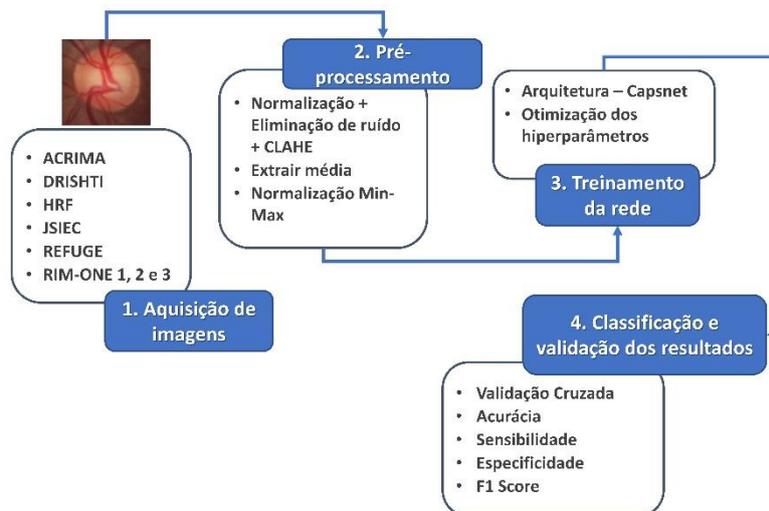


Figura 3. Etapas do método.

4.1. Aquisição de Imagens

Nesta primeira etapa, buscou-se bases de imagens de retinografias gratuitas na internet que estivessem rotuladas como glaucomatosa ou saudável. O conjunto de imagem utilizado incluiu imagens das bases: RIM-ONE versão 1, 2 e 3, JSIEC, ACRIMA, DRISHTI, HRF e REFUGE, todas com a máscara da região do disco óptico. Para a base RIM-ONE 1, as imagens rotuladas como *Moderate*¹ e *OHT*² foram incluídas nesse trabalho na classe Glaucoma. Da base JSIEC, que possui 1000 retinografias divididas em 30 classes, foram utilizadas somente as classificadas como *normal* e *glaucoma e suspeita*. A base REFUGE possui no total 1200 imagens, contudo, somente 400 são rotuladas e somente estas foram utilizadas. A base DRISHTI é geralmente utilizada para trabalhos de segmentação do nervo óptico e originalmente não possui classificação e foi obtida do trabalho de Claro et al. (2019). No total, foram reunidas 2.086 imagens de retinografias, sendo 848 de pacientes com glaucoma e 1227 de pacientes saudáveis. A Figura 4 apresenta alguns exemplos de imagens da base de dados.

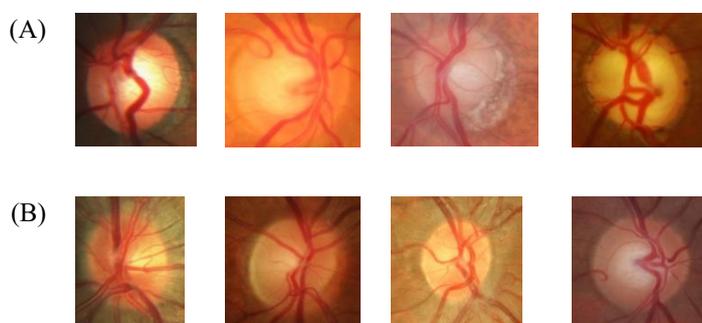


Figura 4. Exemplos de imagens da base. (A) Glaucoma, (B) Saudável.

4.2. Pré-processamento

A etapa de pré-processamento teve como objetivo remover ruídos e destacar aspectos da textura da retinografia. Para tanto, foram feitos testes utilizando duas técnicas de pré-processamento a fim de comparar os resultados: especificação de histograma e uma técnica similar à que foi aplicada em (Shanti et al., 2018), utilizando a normalização *min-max* para manter os valores dos pixels limitados a um determinado intervalo, a remoção de ruídos que, resumidamente, substitui os pixels de uma imagem com a média das cores de pixels similares (Buade et al, 2011), e por último a aplicação do CLAHE, uma equalização de histograma baseada em regiões da imagem. A Figura 5 mostra alguns exemplos do resultado da aplicação das funções de normalização, remoção de ruído (Buade et al, 2011) e CLAHE, baseado em (Shanti, R., 2018).

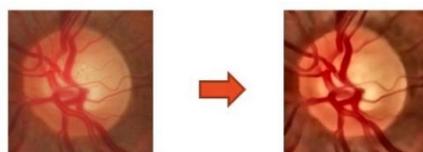


Figura 5. Aplicação da normalização, remoção de ruído e Clahe. À esquerda, a imagem original e à direita, o resultado do processamento.

¹ Moderado.

² *Ocular Hypertension*, ou hipertensão ocular.

Por serem provenientes de diferentes fontes, as imagens da base de dados utilizada apresentam uma grande variação de cores. Portanto, após a realização do processo descrito anteriormente, cada imagem foi submetida ao mesmo pré-processamento utilizado em (Simonyan et al., 2015) que consiste em subtrair o valor RGB médio, computado no conjunto de treinamento, de cada pixel da imagem antes de passá-la para a rede, com o intuito de padronizar as cores das imagens. Para eliminar os possíveis valores negativos originados nessa etapa, utilizou-se a normalização *min-max*. O resultado é ilustrado na Figura 6. Pelo fato da média dos canais de cores das retinografias apresentarem um menor valor no canal azul (tornando a subtração menos significativa nesse canal), as imagens adquirem um tom azulado, tornando esta uma característica em comum entre as imagens.

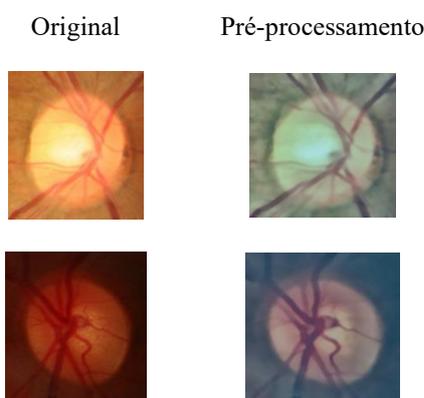


Figura 6. Resultado da extração da média.

Além da extração da média, testou-se ainda a especificação de histograma em uma tentativa de mitigar o problema das diferentes cores das imagens da base. A especificação de histograma é uma técnica usada para modificar um dado de modo que o seu histograma combine com o histograma de uma referência pré-definida (Zhang et al., 1990). Pode ser usada para normalização no processamento de imagens, principalmente em caso onde as imagens foram obtidas de diferentes fontes ou situações. Nesta pesquisa, testes foram feitos escolhendo como referência duas imagens da base de dados: uma de um paciente com glaucoma e outra de um paciente saudável. As referências foram selecionadas levando em consideração a visibilidade e a qualidade dos detalhes que estas apresentavam.

4.3. Definição e Treinamento da Capsnet

O objetivo desta etapa foi avaliar diferentes combinações de hiperparâmetros para a Capsnet para obtenção de uma arquitetura que apresente o modelo mais robusto. Os hiperparâmetros escolhidos focaram nas camadas convolucionais da rede, ou seja, no extrator de características, e em aspectos da configuração do treinamento da rede. Após alguns testes iniciais, foi definido um número limite de três camadas convolucionais para as Capsnets testadas. Afim de otimizar os hiperparâmetros de forma automática utilizou-se a implementação do trabalho de Bergstra et al. (2015), o *hyperopt*. O trabalho do autor faz uso de um algoritmo batizado como *Tree Parze Estimator* (TPE) (Bergstra et al., 2011), que busca minimizar o resultado da função de minimização, retornando no final a configuração de hiperparâmetros entre o espaço de busca que obteve o melhor resultado. A função para minimização passada ao método incluiu as seguintes ações: definir uma

Capsnet baseada no espaço de busca informado, configurar o treinamento e treinar a rede. Entre os hiperparâmetros, foram escolhidos: o tamanho do *batch*, taxa de aprendizado, queda da taxa de aprendizado, otimizador, função de ativação, número de épocas, tamanho do filtro das camadas convolucionais, multiplicador do filtro (utilizado para definir a quantidade de filtros das camadas convolucionais) e função de perda.

4.4. Classificação

O objetivo desta etapa é classificar as retinografias em saudáveis e não saudáveis. Além de extrair características da base de dados, as redes definidas com o *hyperopt* também classificam as imagens recebidas. As últimas camadas da Capsnet consistem de camadas completamente conectadas, que conectam todos os neurônios de uma camada com cada neurônio da próxima camada, e são responsáveis pela classificação. Utilizou-se a função sigmoïdal como função de ativação na última camada FC.

5. Resultado

Esta seção apresenta os resultados obtidos neste estudo para cada abordagem utilizada. Primeiramente, apresentam-se os resultados dos primeiros experimentos realizados. Foram utilizadas três diferentes combinações de técnicas de pré-processamento: (1) Especificação de histograma utilizando a imagem de um paciente saudável como referência, (2) Especificação de histograma utilizando a imagem de um paciente com glaucoma como referência e (3) Normalização + remoção de ruído + CLAHE. A Tabela 1 mostra o espaço de busca utilizado para o *hyperopt*. Nesses experimentos, utilizou-se *data augmentation*, que incluiu rotações das imagens da classe minoritária, para resolver o desbalanceamento da base e validação cruzada com 5 *folders*. Antes de inseridas na rede, as imagens tiveram suas dimensões alteradas para 112x112. Separou-se 20% da base para validação e 80% para treino. A Tabela 2 mostra os resultados obtidos para as três técnicas de pré-processamento.

Tabela 1. Espaço de busca utilizado.

Hiperparâmetro	Valores
tamanho do batch	entre 10 e 45
taxa de aprendizado	entre 5×10^{-5} e 10^{-3}
queda da taxa de aprendizado	entre 0,1 e 0,9
otimizador	sgd, adadelta, adagrad, adam, adamax e nadam
função de ativação	relu, tanh, elu
épocas	entre 40 e 55
tamanho do filtro	5, 7, 9, 11
multiplicador do filtro	entre 0,5 e 1,2
função de perda	categorical crossentropy, mse, mean squared logarithmic error, squared hinge, categorical hinge

Analisando a Tabela 2, percebe-se que as técnicas de pré-processamento (1) e (2), obtiveram resultados regulares, com a técnica (2) alcançando os melhores resultados indicando que a utilização da imagem de um paciente com glaucoma como referência para a especificação de histograma foi mais significativa. Contudo, a técnica de pré-processamento (3) obteve os melhores resultados gerais e, por isso, somente esta foi utilizada nos testes seguintes.

Tabela 2. Resultado dos experimentos para os três pré-processamentos.

Pré-processamento	ACU. (%)	ESP. (%)	SEN. (%)	F1. (%)
(1)	71,32 (+/- 0,01)	70,17 (+/- 0,02)	72,99 (+/- 0,02)	67,55 (+/- 0,01)
(2)	79,75 (+/- 0,01)	82,96 (+/- 0,02)	75,11 (+/- 0,03)	75,18 (+/- 0,01)
(3)	85,29 (+/- 0,01)	88,28 (+/- 0,01)	81,77 (+/- 0,04)	83,55 (+/- 0,02)

Partindo do fato de que os vasos sanguíneos existentes na região do disco óptico auxiliam no diagnóstico do glaucoma, os próximos experimentos incluíram uma nova abordagem: uma Capsnet com duas entradas, uma para a retinografia e outra para a máscara dos vasos sanguíneos. Contudo, as bases de dados utilizadas neste trabalho não incluem a segmentação dos vasos. Desse modo, buscou-se técnicas de segmentação dos vasos sanguíneos disponíveis na internet. O trabalho escolhido utiliza a VesselNet, uma variação da arquitetura de rede neural convolucional U-Net, baseada no trabalho de Huang et al (2018). Este método obteve AUC de 0,9841 na base DRIVE, que é amplamente utilizada em trabalhos de segmentação de vasos sanguíneos em retinografia. A Figura 7 apresenta o resultado deste método aplicado à base utilizada neste trabalho.

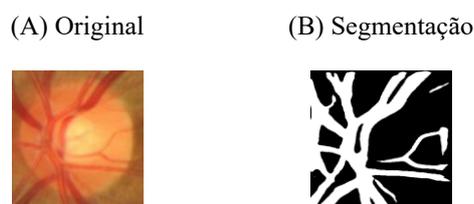


Figura 7. Resultado da segmentação dos vasos sanguíneos.

De posse das máscaras, definiu-se uma arquitetura de Capsnet que permite a entrada de duas imagens, a retinografia e a máscara dos vasos sanguíneos, batizada de Vessel Capsnet. Cada uma das duas camadas de entrada leva a um bloco de 3 camadas convolucionais, num total de 6 camadas, como mostra o exemplo da Figura 8. Estes blocos são conectados no final por uma camada que concatena as características extraídas de cada bloco em um único vetor que alimentará as cápsulas da camada *PrimaryCaps* da rede.

Arquitetura utilizada	
Input(shape=112,112,3)	Input(shape=112,112,1)
Conv2D(256, 3x3)	Conv2D(256, 3x3)
Conv2D(256, 3x3)	Conv2D(256, 3x3)
Conv2D(256, 3x3)	Conv2D(256, 3x3)
Concat	
PrimaryCaps(dim=8)	
CapsuleLayer(dim=16)	
Dense(512)	
Dense(1024)	
Dense(37632)	

Figura 8. Exemplo de arquitetura proposta. O tamanho e a quantidade de filtros são apenas ilustrativos já que estes são parâmetros otimizáveis.

Nos experimentos, os hiperparâmetros foram otimizados utilizando o mesmo espaço de busca da Tabela 1. Utilizou-se ainda *data augmentation* para tratar o desbalanceamento da base e validação cruzada com 5 *folders*. A comparação do resultado desta abordagem em relação à anterior é mostrado na Tabela 3. A Figura 9 compara a evolução da sensibilidade e especificidade de ambas as abordagens no decorrer das épocas.

Tabela 3. Comparação dos resultados.

Método	ACU. (%)	ESP. (%)	SEN. (%)	F1. (%)
Capsnet	85,29 (+/- 0,01)	88,28 (+/- 0,01)	81,77 (+/- 0,04)	83,55 (+/- 0,02)
Vessel Capsnet	83,22 (+/- 0,01)	88,52 (+/- 0,01)	76,87 (+/- 0,02)	80,64 (+/- 0,02)

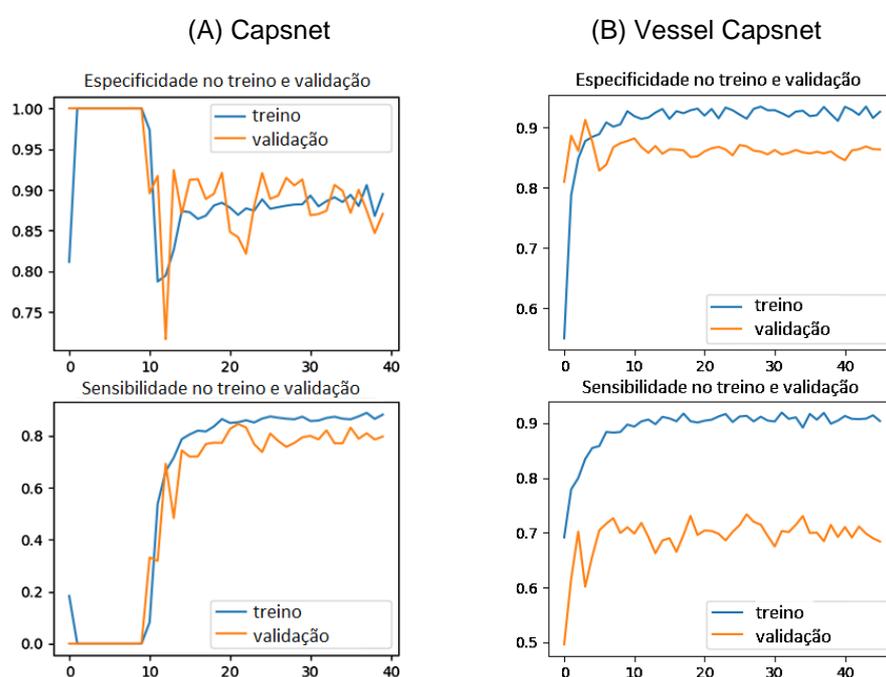


Figura 9. Evolução dos métodos no decorrer das épocas.

Apesar de obter um resultado levemente superior na especificidade de acordo com a Tabela 3, o método que utiliza a Vessel Capsnet não conseguiu um resultado apropriado na sensibilidade por conta do *overfitting*. Mas observando a evolução de ambos os métodos na Figura 9, vemos que a abordagem com Vessel Capsnet obteve no treino valores maiores de especificidade e sensibilidade e o aprendizado da rede se deu ainda nas 10 primeiras épocas do treinamento, o que pode indicar que a combinação dos vasos sanguíneos e retinografia melhora a detecção do glaucoma, apesar do método não conseguir generalizar o conhecimento na validação. Se tratado o *overfitting*, é possível que esta abordagem obtenha valores superiores aos apresentados pela Capsnet tradicional. A Tabela 4 mostra a comparação do método desenvolvido com trabalhos relacionados. Ela inclui ainda uma abordagem testada que utiliza a arquitetura VGG16 e *transfer learning* a fim de avaliar o resultado da classificação do glaucoma utilizando a Capsnet em comparação às redes tradicionais.

Tabela 4. Comparação dos resultados com trabalhos relacionados.

Método	Base	Nº de imagens	Performance
Li et al. (2018)	LabelMe (privada)	39.745	ACC: 98,6% SEN: 95,6% ESP: 92,2%
Raghavendra et al. (2018)	Privada	1.426	ACC: 98,13% SEN: 98% ESP: 98,3%
Lima (2019)	Rim-One 2	455	ACC: 96,63% SEN: 94,87% ESP: 98% PREC: 97,37%
Claro et al. (2019)	HRF, JSIEC, Acrima	802	ACC: 93,61% AUC: 0,975
Pereira (2017)	Rim-One 2	455	ACC: 88,57%
VGG16 + <i>Transfer Learning</i>	Acrima, Drishti, HRF, Refuge, JSIEC, Rim-One 1- 2-3	2086	ACC: 88,29% SEN: 84,27% ESP: 91,56% F1: 86,75%
Proposto	Acrima, Drishti, HRF, Refuge, JSIEC, Rim-One 1- 2-3	2086	ACC: 85,29% SEN: 81,77% ESP: 88,28% F1: 83,55%

Os trabalhos de Li et al (2018) e Raghavendra et al. (2018) obtiveram os melhores resultados, com 98,6% e 98,13% de acurácia, respectivamente. O método proposto apresentou resultados inferiores à abordagem envolvendo a arquitetura VGG e *transfer learning*, contudo, novos testes incluindo os vasos sanguíneos e a ampliação do espaço de busca (adicionando, por exemplo, dimensão e quantidade das cápsulas nas camadas) para o *hyperopt* podem aprimorar os resultados obtidos.

6. Conclusão

Esta pesquisa avaliou a utilização de redes em cápsula para a detecção de glaucoma em retinografias. Utilizou-se a implementação do trabalho de Bergstra et al. (2015) para estimar os melhores hiperparâmetros e definir uma versão otimizada da Capsnet. Os resultados utilizando a técnica de processamento baseada em Shanti et al. (2018) mostraram os melhores resultados, alcançando 85,29% de acurácia, 81,77% de sensibilidade e 88,28% de especificidade. Testes incluindo a segmentação dos vasos sanguíneos em conjunto com a retinografia apresentaram resultados promissores e podem futuramente ultrapassar estes resultados. Para trabalhos futuros, sugere-se: (1) ampliar a base de dados, (2) fazer novos experimentos envolvendo a combinação retinografia e vasos sanguíneos (3) testar outras variações na arquitetura da Capsnet, (4) utilizar outras técnicas de pré-processamento.

References

- Acharya, U. R. et al. (2017) “Automated detection of arrhythmias using different intervals of tachycardia ECG segments with convolutional neural network”, <https://doi.org/10.1016/j.ins.2017.04.012>.
- Afshar, P. (2019) “Capsule Networks for Brain Tumor Classification Based on MRI Images and Coarse Tumor Boundaries”, ICASSP.
- Anupama, A. et al. (2019) “Breast Cancer Classification using Capsule Network with Preprocessed Histology Images”, 2019 International Conference on Communication

- and Signal Processing (ICCS), Chennai, Índia, <https://doi.org/10.1109/ICCS.2019.8698043>.
- Araújo, F. H. D. et al. (2017) “Redes Neurais Convolucionais com Tensorflow: Teoria e Prática”, III Escola Regional de Informática do Piauí. Livro Anais - Artigos e Minicursos, v. 1, n. 1, p. 382-406.
- Barreto, J. M. (2002) “Introdução às Redes Neurais Artificiais”, <http://www.inf.ufsc.br/~j.barreto/tutoriais/Survey.pdf>.
- Bergstra, J. et al. (2011) “Algorithms for Hyper-Parameter Optimization”, 25th Annual Conference on Neural Information Processing Systems (NIPS 2011), Granada, Espanha.
- Bergstra, J. et al. (2015) “Hyperopt: A Python Library for Optimizing the Hyperparameters of Machine Learning Algorithms”, <https://doi.org/10.1088/1749-4699/8/1/014008>.
- Buades A. et al. (2011) “Non-Local Means Denoising”, http://dx.doi.org/10.5201/ipol.2011.bcm_nlm.
- Claro, M. et al. (2019) “An hybrid feature space from texture information and transfer learning for glaucoma classification”, <https://doi.org/10.1016/j.jvcir.2019.102597>.
- Ferreira, A. S. (2017) Redes Neurais Convolucionais Profundas na Detecção de Plantas Daninhas em Lavoura de Soja. Dissertação de Mestrado, Universidade Federal do Mato Grosso do Sul.
- Flaxman, S. R. et al. (2017) “Global causes of blindness and distance vision impairment 1990–2020: a systematic review and meta-analysis”, *The Lancet Glob Health*.
- Goodfellow, I. et al. (2016) *Deep Learning*. MIT, ISBN 9780262035613.
- Haykin, S. (2001) *Redes neurais: princípios e prática*. Porto Alegre: Bookman.
- Huang, G. et al. (2016) “Densely Connected Convolutional Networks”, [arXiv:1608.06993v5](https://arxiv.org/abs/1608.06993v5).
- Sabour, et al. (2017) “Dynamic Routing Between Capsules”, [arXiv:1710.09829v2](https://arxiv.org/abs/1710.09829v2).
- Iesmantas, T. e Alzbutas, R. (2018) “. Convolutional capsule network for classification of breast cancer histology images”, https://doi.org/10.1007/978-3-319-93000-8_97.
- Keras (2020). Documentação. Disponível em: <https://keras.io/>.
- Kurup, R. et al (2020) “Effect of Data Pre-processing on Brain Tumor Classification Using Capsulenet”, https://doi.org/10.1007/978-981-13-8461-5_13.
- Lauretto, R. C. e Filho, A. L. (1997) Glaucomas. Simpósio: Oftalmologia para o clínico, 56-65.
- Li, Z. et al. (2018) “Efficacy of a Deep Learning System for Detecting Glaucomatous Optic Neuropathy Based on Color Fundus Photographs”, <https://doi.org/10.1016/j.jophtha.2018.01.023>.
- Lima, A. C. M. (2019) *Aprendizagem Profunda Aplicada ao Diagnóstico de Glaucoma*. Dissertação de Mestrado, Universidade Federal do Maranhão.

- Lopes, U. K. (2017) Redes Neurais Convolucionais Aplicadas ao Diagnóstico de Tuberculose Por Meio de Imagens Radiológicas. Dissertação de mestrado, Universidade do Vale do Rio dos Sinos.
- Mobiny, A. e Nguyen, H. V. (2018) “Fast CapsNet for Lung Cancer Screening”, https://doi.org/10.1007/978-3-030-00934-2_82.
- Moraes, L. A. G. (2018) Utilização de Redes em Cápsula para Classificação de Regiões Promotoras em Sequências de DNA. Dissertação de mestrado, Universidade Federal de Ouro Preto.
- OpenCV (2020). Disponível em: <https://opencv.org/>.
- Patrick, M. K. et al. (2019) “Capsule Network – A Survey”, <https://doi.org/10.1016/j.jksuci.2019.09.014>.
- Pereira, R. M. P. (2017) Estudo do Desempenho de Redes Neurais Convolucionais Aplicada ao Reconhecimento de Símbolos Musicais, Glaucoma e Texto. Monografia, Universidade Federal do Maranhão.
- Pizer, S. M. (1987) “Adaptive Histogram Equalization and its Variations. Computer Vision, Graphics and Image Processing”, vol. 39, 1987, pp. 355–368.
- Raghavendra, U. et al. (2018) “Deep convolution neural network for accurate diagnosis of glaucoma using digital fundus images”, <https://doi.org/10.1016/j.ins.2018.01.051>.
- Reis, N. C. et al. (2018) “Diagnóstico Automático de Glaucoma em Imagens de Retinografia baseado no CDR usando U-Net”, VII Jornada de Informática do Maranhão.
- Ribeiro, J. L. M. (2018) “Meta-aprendizado aplicado ao Problema de Reconhecimento de Expressões Faciais”, Monografia, Universidade Federal do Maranhão.
- Sanchez, A. J. (2018) “Capsule Networks against Medical Imaging Data Challenges”, arXiv:1807.07559v1.
- Scikit-Image (2020). Disponível em: <https://scikit-image.org/>.
- Shanti, R. e Prabakaran, S. (2018) “Detection of Microaneurysms and Hemorrhages in Fundus Image for Glaucoma Diagnosis” International Journal of Engineering and Technology, Vol. 7, <https://doi.org/10.14419/ijet.v7i2.24.12092>.
- Silva, M. G. (2018) “Diagnóstico do glaucoma em retinografias usando descritores de textura”, Centro de Ciências Exatas e Tecnologia, Universidade Federal do Maranhão.
- Simonyan, K. e Zisserman, A. (2014) “Very Deep Convolutional Networks For Large-Scale Image Recognition”, arXiv 1409.1556.
- Tham, Y. C. et al. (2014) “Global prevalence of glaucoma and projections of glaucoma burden through 2040: a systematic review and meta-analysis”, *Ophthalmology*, 2014, 121(11), 2081–2090. <https://doi.org/10.1016/j.ophtha.2014.05.013>.
- World Health Organization (2019). World Report on Vision, ISBN 978-92-4-151657-0.
- Zhang, K. Y. J. e Main, P. (1990) “Histogram Matching as a New Density Modification Technique for Phase Refinement and Extension of Protein Molecules”, <https://doi.org/10.1107/S0108767389009311>.