



UNIVERSIDADE FEDERAL DO MARANHÃO

Programa de Pós-Graduação em Ciência da Computação

Marcos Melo Ferreira

***Redes Neurais Multinível para Classificação do
Ângulo da Câmara Anterior Utilizando Imagens
OCT-SA***

**São Luís
2021**

Marcos Melo Ferreira

Redes Neurais Multinível para Classificação do Ângulo da Câmara Anterior Utilizando Imagens OCT-SA

Dissertação apresentada como requisito parcial para obtenção do título de Mestre em Ciência da Computação, ao Programa de Pós-Graduação em Ciência da Computação, da Universidade Federal do Maranhão.

Programa de Pós-Graduação em Ciência da Computação
Universidade Federal do Maranhão

Orientador: Prof. Dr. Geraldo Braz Junior

São Luís - MA

2021

Ficha gerada por meio do SIGAA/Biblioteca com dados fornecidos pelo(a) autor(a).
Núcleo Integrado de Bibliotecas/UFMA

Ferreira, Marcos Melo.

Redes Neurais Multinível para Classificação do Ângulo da Câmara Anterior Utilizando Imagens OCT-SA / Marcos Melo Ferreira. - 2021.

59 p.

Orientador(a): Geraldo Braz Junior.

Dissertação (Mestrado) - Programa de Pós-graduação em Ciência da Computação/ccet, Universidade Federal do Maranhão, São Luis, 2021.

1. Arquitetura Multinível. 2. Glaucoma. 3. Redes Neurais Convolucionais. 4. Transferência de Aprendizado.
I. Braz Junior, Geraldo. II. Título.

Marcos Melo Ferreira

Redes Neurais Multinível para Classificação do Ângulo da Câmara Anterior Utilizando Imagens OCT-SA

Dissertação apresentada como requisito parcial para obtenção do título de Mestre em Ciência da Computação, ao Programa de Pós-Graduação em Ciência da Computação, da Universidade Federal do Maranhão.

Trabalho Aprovado. São Luís - MA, 18 de Março de 2021:

Prof. Dr. Geraldo Braz Junior

Orientador

Universidade Federal do Maranhão

Prof. Dr. Anselmo Cardoso de Paiva

Examinador Interno

Universidade Federal do Maranhão

Prof. Dr. João Dallyson Sousa de Almeida

Examinador Interno

Universidade Federal do Maranhão

Prof. Dr. Sidnei Alves de Araújo

Examinador Externo

Universidade Nove de Julho

São Luís - MA

2021

Aos meus pais, irmãos e amigos que sempre me apoiaram e deram todo suporte necessário para a conclusão desta etapa.

Agradecimentos

Esta dissertação de mestrado é resultado de muito trabalho e dedicação e é importante demonstrar os meus sinceros agradecimentos aos que me ajudaram nesta etapa da minha vida.

Primeiramente à Deus, pela saúde e determinação para superar todas as dificuldades que se apresentaram.

Aos meus pais, Dimas e Lenir, pois sempre acreditaram na minha capacidade, sempre me incentivaram e me apoiaram, estando ao meu lado, me dando amor e suporte durante toda a minha vida.

Aos meus pais biológicos, Neilson e Vitória, que sempre demonstraram muito amor e sempre me apoiaram.

A todos os meus irmãos, sou muito grato pela amizade e confiança.

Aos professores do programa de pós graduação, em especial ao meu orientador, professor Geraldo, pela oportunidade, por estar sempre disponível para orientar e ajudar, pela paciência, pelo incentivo e por tudo que me ensinou.

Aos meus colegas de mestrado, principalmente Micael que me ajudou muito no início desta caminhada.

Aos meus companheiros de pesquisa, sempre prontos a ajudar e sugerir soluções para os problemas que se apresentaram.

Aos meus colegas do Instituto Federal de Educação, Ciência e Tecnologia do Maranhão - Campus São José de Ribamar, pelo apoio e compreensão que demonstraram, principalmente a Letícia, que sempre atendeu quando precisei de ajuda, e a Flávyia que sempre me apoiou bastante.

E por fim, a todos que contribuíram para o término deste trabalho e acreditaram em mim.

*"A persistência é amiga da conquista. Se você quer chegar aonde a maioria não chega,
faça o que a maioria não faz."*

(Bill Gates)

Resumo

O glaucoma é apontado como sendo uma das principais causas de comprometimento da visão, e a principal causa de cegueira irreversível. As principais formas da doença são o glaucoma primário de ângulo aberto e o glaucoma primário de ângulo fechado. Em pessoas com glaucoma de ângulo fechado, ocorre o estreitamento do ângulo da câmara anterior, consequentemente acarretando o aumento da pressão intraocular provocando danos ao nervo óptico, causando perda parcial ou total da visão. Como os danos são irreversíveis, um diagnóstico precoce é essencial, porém é dificultado devido ao fato da doença ser assintomática nos estágios iniciais. Para detecção precoce da doença, são recomendados exames de imagem de rotina, sendo um deles a Tomografia de Coerência Óptica do Segmento Anterior, que permite a classificação do ângulo, fundamental para o diagnóstico. A análise deste tipo de imagem requer um grau de interpretação por parte dos especialistas, devido a isso, a avaliação de muitas imagens demanda muito tempo, podendo levar a fadiga do profissional. A utilização de métodos automáticos para auxiliar na interpretação das imagens contribuiria para obtenção de diagnósticos mais rapidamente. Neste trabalho é proposto um método automático para classificação do ângulo da câmara anterior, presente em imagens de Tomografia do Segmento Anterior, baseado em aprendizagem profunda, utilizando redes neurais convolucionais. Inicialmente, cinco modelos pré-treinados de redes convolucionais foram ajustados para extração de características e classificação das imagens. A seguir, os modelos foram combinados em uma arquitetura multinível, com o objetivo de se aumentar a capacidade de classificação. Como melhor resultado foi alcançado um valor AUC (do inglês, *Area Under the Curve*) de 0,999.

Palavras-chave: Glaucoma, Redes Neurais Convolucionais, Arquitetura Multinível, Transferência de Aprendizado.

Abstract

Glaucoma is identified as one of the main causes of visual impairment, and the main cause of irreversible blindness. The main forms of the disease are primary open-angle glaucoma and primary angle-closure glaucoma. In people with angle-closure glaucoma, the anterior chamber angle narrows, consequently causing an increase in intraocular pressure causing damage to the optic nerve, causing partial or total vision loss. As the damage is irreversible, an early diagnosis is essential, but it is hampered due to the fact that the disease is asymptomatic in early stages. For early detection of the disease, routine imaging tests are recommended, one of which is Anterior Segment Optical Coherence Tomography, which allows an angle classification, which is essential for diagnosis. An analysis of this type of image requires a degree of interpretation on the part of specialists, because of this, the evaluation of many images requires a lot of time, which can lead to professional fatigue. The use of automated methods to assist in the interpretation of images would contribute to get diagnoses more quickly. In this work, an automated method is proposed to classify the anterior chamber angle, present in Anterior Segment images, based on deep learning, using convolutional neural networks. Initially, five pre-trained models of convolutional networks were adjusted to perform feature extraction and classify images. Next, the models were combined in a multilevel architecture, with the objective of increasing the classification capacity. As best result achieved an AUC value (Area Under the Curve) of 0.999.

Keywords: Glaucoma, Convolutional Neural Networks, Multilevel Architecture, Transfer Learning.

Lista de ilustrações

Figura 1 – Exemplo de OCT-SA de um caso de ângulo aberto	17
Figura 2 – Exemplo de OCT-SA de um caso de ângulo fechado	17
Figura 3 – Contato entre íris e região trabecular	17
Figura 4 – Simulação da progressão da perda visual devido ao glaucoma	26
Figura 5 – Anatomia do olho humano	27
Figura 6 – Estruturas presentes na região do ângulo da câmara anterior	28
Figura 7 – Morfologia dos principais mecanismos de fechamento angular. Bloqueio pupilar (Imagem superior esquerda), íris em platô (Imagem superior direita), fechamento induzido pelo cristalino (Imagem inferior esquerda) e aumento do diâmetro da lente (Imagem inferior direita).	29
Figura 8 – Convolução Discreta	31
Figura 9 – Arquitetura de uma rede convolucional	33
Figura 10 – Arquitetura da rede VGG	34
Figura 11 – Módulo Inception	34
Figura 12 – Bloco Residual	35
Figura 13 – Conexões do Modelo DenseNet	36
Figura 14 – Exemplo de Arquitetura de uma Rede Multinível	37
Figura 15 – Fluxograma da metodologia proposta.	40
Figura 16 – Imagem com dois ângulos presentes na câmara anterior.	41
Figura 17 – Imagem com o ângulo esquerdo da câmara anterior.	42
Figura 18 – Arquitetura Multinível	44
Figura 19 – Curvas ROC.	48
Figura 20 – Imagens OCT e Mapas de Ativação de uma classificação correta.	49
Figura 21 – Regiões usadas pelo modelo VGG19 para classificação das imagens (Classificação correta).	49
Figura 22 – Regiões usadas pelo modelo ResNet50 para classificação das imagens (Classificação correta).	50
Figura 23 – Exemplo de classificação incorreta.	50
Figura 24 – Mapas de Ativação - Modelo DenseNet121 (Classificação correta).	51

Lista de tabelas

Tabela 1 – Comparação entre diferentes métodos.	24
Tabela 2 – Matriz de Confusão.	38
Tabela 3 – Resultados obtidos por modelos com um nível, alimentados com imagens com dois ou um ACA.	46
Tabela 4 – Resultados obtidos por modelos multinível formados pela mesma CNN.	47
Tabela 5 – Resultados obtidos por modelos Multinível formados por CNNs diferentes.	47
Tabela 6 – Comparação dos resultados obtidos por diferentes métodos.	51
Tabela 7 – Artigos aceitos para publicação.	54

Lista de abreviaturas e siglas

ACA	<i>Ângulo da Câmara Anterior</i>
ACU	<i>Acurácia</i>
AGE	<i>Angle Closure Evaluation</i>
AUC	<i>Area Under Curve</i>
BIF	<i>Bio Inspired Features</i>
CNN	<i>Convolutional Neural Network</i>
ESP	<i>Especificidade</i>
FN	<i>False Negative</i>
FP	<i>False Positive</i>
GPAA	<i>Glaucoma Primário de Ângulo Aberto</i>
GPAF	<i>Glaucoma Primário de Ângulo Fechado</i>
HOG	<i>Histogram of Oriented Gradients</i>
ILSVRC	<i>ImageNet Large-Scale Visual Recognition Challeng</i>
KNN	<i>K-Nearest Neighbors</i>
LSTM	<i>Long Short Term Memory</i>
MLP	<i>MultiLayer Perceptron</i>
OCT	<i>Optical Coherence Tomography</i>
OCT-SA	<i>OCT do Segmento Anterior</i>
OMS	<i>Organização Mundial da Saúde</i>
PIO	<i>Pressão Intraocular</i>
PRE	<i>Precisão</i>
ReLu	<i>Rectified Linear Unit</i>
ROC	<i>Receiver Operating Characteristic</i>

SBG	<i>Sociedade Brasileira de Glaucoma</i>
SEN	<i>Sensibilidade</i>
SS	<i>Scleral Spur</i>
SVM	<i>Support Vector Machine</i>
TIC	<i>Trabecular-iris contact</i>
TN	<i>True Negative</i>
TP	<i>True Positive</i>
UBM	<i>Ultrasound Biomicroscopy</i>
VGG	<i>Visual Geometry Group</i>

Sumário

1	INTRODUÇÃO	15
1.1	Objetivos	18
1.1.1	Objetivos Específicos	18
1.1.2	Contribuições	18
1.2	Organização do Trabalho	19
2	TRABALHOS RELACIONADOS	20
2.1	Técnicas que utilizam Aprendizado de Máquina	20
2.2	Técnicas que utilizam Aprendizagem Profunda	22
3	FUNDAMENTAÇÃO TEÓRICA	26
3.1	Glaucoma	26
3.2	Tomografia de Coerência Óptica do Segmento Anterior	28
3.3	Redes Neurais Convolucionais	30
3.4	Arquiteturas de CNN	33
3.5	Transferência de Aprendizado	35
3.6	Arquiteturas Multinível	36
3.7	Métricas de Avaliação	37
3.8	Considerações Finais	39
4	METODOLOGIA	40
4.1	Aquisição das Imagens	40
4.2	Extração de características visuais	41
4.3	Redes Multinível	43
4.4	Avaliação de Resultados	44
4.5	Considerações Finais	45
5	RESULTADOS	46
5.1	Estudo de Casos	48
5.2	Discussão	50
6	CONCLUSÃO	53
6.1	Trabalhos Futuros	54
6.2	Produções Científicas	54
	REFERÊNCIAS	55

1 Introdução

O glaucoma é apontado como sendo uma das principais causas de comprometimento da visão, e a principal causa de cegueira irreversível (QUIGLEY; BROMAN, 2006; LI et al., 2017; XU et al., 2019; FU et al., 2019a). Segundo estimativas publicadas pela Organização Mundial de Saúde (OMS), feitas no ano de 2013, o número de pessoas com glaucoma em 2020 seria de 76 milhões (THAM et al., 2014; OMS, 2019). Deste total, estimativas de 2018 apontam que cerca de 7,7 milhões de pessoas, possuem comprometimento moderado ou severo da visão, ou cegueira resultante de formas mais graves da doença (BOURNE et al., 2021). Segundo a mesma publicação da OMS, uma projeção feita em 2017 aponta que aproximadamente 2,2 bilhões de pessoas possuem algum tipo de comprometimento da visão, e que deste total de casos, ao menos 1 bilhão ainda não foi identificado ou poderia ter sido evitado, caso houvesse o tratamento das doenças nos estágios iniciais (FRICKE et al., 2018; OMS, 2019). Os principais tipos de glaucoma são o glaucoma primário de ângulo aberto (GPAA) e o glaucoma primário de ângulo fechado (GPAF) (QUIGLEY; BROMAN, 2006; THAM et al., 2014; KUMAR; CHAUHAN; DAHIYA, 2016; NIWAS et al., 2016) que é menos comum e tende a ser mais agudo.

Do total de pessoas com glaucoma no Brasil, aproximadamente 1 milhão segundo a OMS (SOUSA et al., 2017), a Sociedade Brasileira de Glaucoma (SBG) estima que 25% dos casos sejam de GPAF (SBG, 2012). Em pessoas com GPAF, o ângulo formado pela junção da íris e da córnea, ângulo iridocorneano, diminui. A diminuição desse ângulo dificulta o fluxo de humor aquoso, essencial para a lubrificação dos olhos, causando aumento da pressão intraocular (PIO), provocando danos irreversíveis ao nervo óptico (XU et al., 2019). Estes danos podem levar o paciente à uma perda parcial ou total da visão, que são irreversíveis. Procedimentos como a Iridotomia à laser (FU et al., 2019a), podem ser empregados para diminuir a pressão, no entanto para que este procedimento tenha eficácia é essencial o diagnóstico precoce da doença.

Para evitar que pacientes com GPAF tenham o nervo óptico gravemente danificado, causando deficiência visual ou cegueira, é essencial um diagnóstico precoce da doença (SARHAN; ROKNE; ALHAJJ, 2019). No entanto, isso é dificultado por alguns fatores, como a ausência de sintomas nos estágios iniciais e o grande número de pacientes em potencial desassistidos em relação à quantidade de oftalmologistas especializados capazes de fazer o diagnóstico. O diagnóstico é feito através de exames de imagem, que possibilitam ao especialista detectar alterações nas estruturas presentes na região da câmara anterior. A gonioscopia, principal exame de imagem utilizado para este fim (XU et al., 2019; PHAM et al., 2019), é altamente dependente da experiência do especialista em identificar algumas estruturas anatômicas, além de ser um exame que requer contato com os olhos, o que

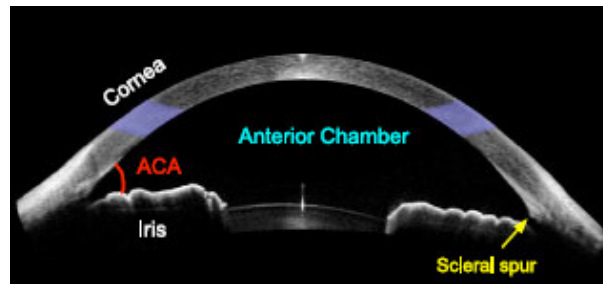
deixa os pacientes desconfortáveis e pode ser considerado inadequado para alguns pacientes (PHAM et al., 2019).

O diagnóstico de glaucoma também pode ser feito com imagens de fundo de olho, tomografia de coerência óptica (OCT, do inglês *Optical Coherence Tomography*) e biomicroscopia ultra-sônica (UBM, do inglês *Ultrasound Biomicroscopy*) (PRIYANKA; JULIET; DEVI, 2019). Dentre os tipos de exames, a OCT é utilizada para diagnóstico de doenças da córnea, avaliação de pré, pós-operatório de pacientes que necessitam ou fizeram cirurgias na córnea e diagnóstico de glaucoma primário de ângulo fechado, através da avaliação do sistema de fluxo de saída do humor aquoso e do ângulo da câmara anterior (ACA) (ANG et al., 2018). O sistema de saída do fluido humor aquoso, é um importante componente para a manutenção da pressão intraocular. Este líquido flui através da pupila, da câmara posterior para a câmara anterior, sendo então drenado pela região trabecular através do canal de Schlemm (ANG et al., 2018), o que mantém a pressão constante. Um aumento da resistência ao fluxo, devido a alterações no trajeto, leva ao aumento da pressão, uma vez que ocorre um desequilíbrio entre a produção e drenagem do fluido.

Através da tomografia de coerência óptica do segmento anterior, Figura 1, (OCT-SA, do inglês *OCT de Segmento Anterior*) é possível avaliar o ângulo da câmara anterior (ACA). A avaliação deste ângulo tem ganhado importância para o diagnóstico de glaucoma primário de ângulo fechado. Além do ângulo é possível avaliar a câmara anterior, a íris e as lentes. O fechamento do ângulo é determinado por qualquer contato entre a íris e a parte superior do ângulo, antes do esporão escleral (SS, do inglês *Scleral Spur*). Sendo assim, a identificação do SS é um marco importante para a avaliação (ANG et al., 2018). Com este exame também é possível determinar parâmetros quantitativos, como a abertura angular, área da câmara anterior (*Anterior Chamber*), área e curvatura da íris (FU et al., 2017). Alterações nestes parâmetros são considerados fatores de risco. Além de possibilitar a avaliação qualitativa e quantitativa da câmara anterior, a OCT-SA é um exame não invasivo, não requer contato com os olhos, é rápido e indolor (CHANSANGPETCH; ROJANAPONGPUN; LIN, 2018). A Figura 1 apresenta um exemplo de OCT-SA de ângulo aberto, e a Figura 2 um exemplo de ângulo fechado, onde é possível visualizar que o ângulo formado pela junção da íris com a córnea está obstruído, o que leva ao bloqueio dos canais de drenagem do humor aquoso.

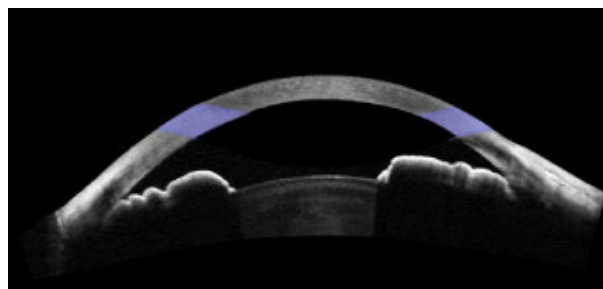
A avaliação do ângulo da câmara anterior, requer um grau de interpretação dos médicos especialistas. O contato entre a íris e a região trabecular (TIC, do inglês *trabecular-iris contact*), ilustrado na Figura 3 é o principal indicador da doença (FU et al., 2019a; SHANG et al., 2019). No entanto, em imagens OCT-SA a região trabecular não pode ser confiavelmente identificada devido a limitações da resolução (FU et al., 2019a). O esporão escleral, *landmark* utilizado para determinar se há ou não o contato entre íris e a região trabecular, não é facilmente identificado nas imagens. Devido a isso, a análise manual de

Figura 1 – Exemplo de OCT-SA de um caso de ângulo aberto



Fonte: (FU et al., 2019).

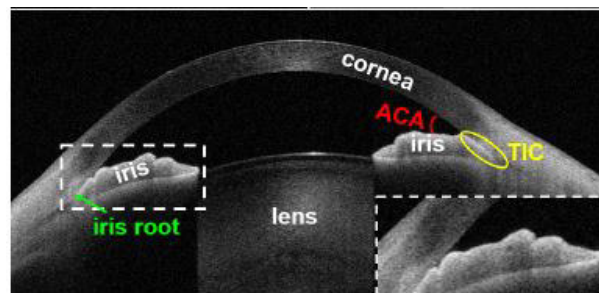
Figura 2 – Exemplo de OCT-SA de um caso de ângulo fechado



Fonte: (FU et al., 2019).

muitos exames exige muito tempo e depende da experiência do profissional.

Figura 3 – Contato entre íris e região trabecular



Fonte: (SHANG et al., 2019)

Devido ao fato de o comprometimento visual devido ao glaucoma ser irreversível (FU et al., 2020), é necessário que intervenções sejam feitas nos estágios iniciais da doença para evitar a perda total da visão (FU et al., 2019a). Com um número estimado de 23 milhões de pessoas afetadas pelo glaucoma de ângulo fechado mundialmente em 2020, a doença é um desafio para os sistemas de saúde preventiva.

Neste contexto, um método automático que possa auxiliar especialistas a detectar a doença utilizando exames de imagem pode vir a ter grande potencial para uso clínico. Métodos baseados em técnicas que utilizam aprendizagem profunda têm mostrado bons

resultados em tarefas de classificação de imagens médicas. As redes neurais convolucionais (CNN, do inglês *Convolutional Neural Network*), principal tipo de rede profunda empregada em classificação de imagens, extrai características das imagens que podem ser consideradas irrelevantes para os seres humanos, que necessitam fazer extração de parâmetros quantitativos. Logo, um método que independa da localização precisa de qualquer estrutura da câmara anterior, como o esporão escleral, teria um grande potencial para uso clínico.

1.1 Objetivos

O objetivo geral deste trabalho consiste em utilizar uma arquitetura multinível de redes neurais convolucionais com o propósito de possibilitar a classificação automática do ângulo da câmara anterior em aberto ou fechado de forma rápida e eficaz, possibilitando aos médicos uma sugestão de diagnóstico precoce a partir de exames de imagem de tomografia de coerência óptica do segmento anterior.

1.1.1 Objetivos Específicos

Especificamente, este trabalho busca os seguintes objetivos aplicados ao problema de classificação do ângulo da câmara anterior usando imagens OCT-SA:

- Utilizar CNNs para extração de características de imagens de OCT-SA;
- Propor e avaliar uma arquitetura multinível que possibilite combinar características extraídas de diferentes regiões de uma imagem de OCT-SA;
- Verificar se redes multinível híbridas, com níveis formados por diferentes CNNs atingem desempenho superior a redes com níveis formados pela mesma CNN;
- Analisar as vantagens e limitações do método proposto;
- Comparar o desempenho do método proposto com outras arquiteturas que utilizam CNNs.

1.1.2 Contribuições

Destacam-se como principais contribuições:

- Um método para classificação do ângulo da câmara anterior baseado em aprendizagem profunda e transferência de aprendizagem, que não necessita da localização de nenhuma estrutura da câmara anterior;

- Avaliação de diferentes CNNs, combinadas em uma arquitetura multinível, capaz de extrair características de regiões relevantes para a classificação;

Os resultados obtidos neste trabalho com uso de aprendizagem profunda, utilizada como auxílio ao diagnóstico de glaucoma de ângulo fechado, mostrou-se efetiva ao utilizar a região das imagens onde encontra-se o ângulo da câmara anterior, formado pela junção da íris com a córnea, para classificação dos exames de imagem.

1.2 Organização do Trabalho

Este trabalho está estruturado da seguinte forma:

- O Capítulo 2 descreve trabalhos relacionados ao tema, apresentando seus métodos e resultados obtidos, para classificação dos ângulos em imagens de tomografia de coerência óptica do segmento anterior.
- O Capítulo 3 trata da fundamentação teórica das técnicas utilizadas. São abordados conceitos fundamentais, além da teoria em que se baseia as técnicas e algoritmos utilizados neste trabalho.
- O Capítulo 4 apresenta as etapas adotadas que compõem a metodologia proposta para este trabalho. São detalhados todos os passos do método proposto para classificação dos ângulos presente nas imagens em aberto ou fechado.
- O Capítulo 5 trata sobre os resultados obtidos e discussões em relação aos experimentos realizados. É feita uma análise sobre os casos em que foi obtido êxito e em que houve falhas, do método proposto.
- O Capítulo 6 apresenta as considerações finais sobre os resultados e trabalhos futuros e os artigos científicos desenvolvidos.

2 Trabalhos Relacionados

Neste capítulo são apresentados os trabalhos publicados na literatura nos últimos 10 anos, com técnicas que utilizam aprendizagem profunda, com o propósito de classificar as imagens e segmentar estruturas presentes na câmara anterior.

Os métodos propostos para classificação do ACA (Ângulo da Câmara Anterior) em imagens de Tomografia de Coerência Óptica do Segmento Anterior (OCT-SA) podem ser organizados em duas abordagens. O primeiro calcula parâmetros quantitativos que têm relevância clínica (por exemplo, área da câmara anterior e espessura da íris). Nesta abordagem, é necessário segmentar as estruturas presentes na câmara anterior. A segunda abordagem consiste da extração de características visuais, que pode ser feita através de técnicas de aprendizado de máquina ou com redes profundas e as redes neurais convolucionais.

2.1 Técnicas que utilizam Aprendizado de Máquina

Os trabalhos citados a seguir utilizaram técnicas de aprendizado de máquina clássico, como forma de classificar o ângulo da câmara anterior e auxiliar na detecção de glaucoma primário de ângulo fechado (GPAF). Estes métodos consistem no uso de técnicas para cálculos de parâmetros clínicos ou para extração de características com a finalidade de obter uma descrição da imagem, por exemplo, histograma de gradientes orientados (HOG, do inglês *Histogram of Oriented Gradients*). Estas representações são utilizadas para classificação das imagens, utilizando alguma técnica de aprendizado de máquina.

Xu et al. (2012) apresentam um método para classificação automática do ângulo da câmara anterior com o propósito de identificação do subtipo de glaucoma, utilizando histogramas de gradientes orientados. Inicialmente a região onde encontra-se o ângulo é localizada, e então o vértice do ângulo é estimado. Tendo o vértice como centro, uma região de tamanho 150x150 pixels é capturada e utilizada para extração de características. Um SVM é utilizado como classificador. Para realização de experimentos foi utilizado um dataset composto de 2040 imagens OCT-SA obtidas de 8 pacientes, 4 diagnosticados com GPAA (Glaucoma Primário de Ângulo Aberto) e 4 diagnosticados com GPAF (Glaucoma Primário de Ângulo Fechado). A validação cruzada foi utilizada para avaliar o método, sendo que o conjunto de teste foi formado pelas imagens de dois pacientes, um paciente de cada classe, e as demais imagens foram utilizadas para treinamento. O método proposto obteve resultado de AUC (*do inglês, Area Under the Curve*) de 0,835.

Em Xu et al. (2013) foi proposto um método para classificação automática do ângulo

da câmara anterior, utilizando-se equalização de histograma como técnica para extração de características juntamente com um SVM (Support Vector Machine) para classificação. Primeiro a região onde encontra-se o ângulo é detectada, e a seguir o ângulo localizado. O próximo passo consiste da extração de características, treinamento e classificação. Um *dataset* composto de 2048 imagens foi utilizado para realização dos experimentos. As imagens foram obtidas de 8 pacientes, sendo 4 diagnosticados com glaucoma primário de ângulo aberto e 4 com glaucoma primário de ângulo fechado. O *dataset* foi dividido em treino e teste a nível de paciente, sendo que a cada sequência de experimentos, as imagens de dois pacientes formavam o conjunto de teste, sendo um com glaucoma primário de ângulo aberto e um com glaucoma primário de ângulo fechado. Foi realizada uma comparação com outros descritores, HOG e BIF (*Bio Inspired Features*), sendo que o método proposto obteve melhores resultados alcançando um valor AUC de 0,921.

Ni et al. (2014) propuseram um método automático para classificação automática do ângulo da câmara anterior baseado no cálculo de parâmetros clínicos. Inicialmente é utilizada uma técnica de pré-processamento para remoção de saturações verticais que aparecem em algumas imagens. A próxima etapa consiste em segmentar as estruturas da câmara anterior para o cálculo de parâmetros. Ao todo são calculados seis parâmetros, três biométricos e três relacionados ao formato da câmara anterior na região onde se encontra o ângulo. A classificação foi realizada usando o algoritmo *fuzzy* KNN. Os experimentos foram um *dataset* de 264 imagens, 135 imagens com ângulos fechados e 129 imagens com ângulos abertos. Para avaliar a generalização do método foi realizada a validação cruzada, dividindo-se o dataset em 5 subconjuntos. O método também foi testado utilizando-se outros três classificadores, sendo que o melhor resultado foi obtido pelo classificador *fuzzy* KNN, alcançando um valor AUC de 0,98.

Fu et al. (2017) empregaram um método de segmentação e quantificação de parâmetros para detecção de fechamento do ângulo. Após a segmentação das estruturas presentes na câmara anterior, íris e córnea, 13 parâmetros clínicos são calculados e utilizados para classificação das imagens OCT-SA. Ao todo, 5 classificadores foram experimentados e os melhores resultados foram obtidos utilizando-se um SVM. Além de experimentos com parâmetros clínicos, também foram realizados experimentos com características visuais. Foi utilizado o descritor HOG para extração de características da região do ângulo. A utilização de características visuais alcançou melhores resultados em dois *datasets*. No primeiro com 4135 imagens, foi obtido um valor AUC de 0.91 e no segundo com 1402 imagens, foi obtido um valor AUC de 0,8727.

2.2 Técnicas que utilizam Aprendizagem Profunda

Os trabalhos citados a seguir utilizaram técnicas de aprendizagem profunda como forma de auxiliar na tarefa de classificar o ângulo da câmara anterior, presente em imagens OCT-SA, essencial para o diagnóstico de glaucoma de ângulo fechado. Alguns trabalhos propõem métodos para segmentação das estruturas e cálculos de parâmetros clínicos, enquanto outros propõem métodos para extração de características.

Em [Fu et al. \(2018\)](#) foi proposto um método para classificação do ângulo da câmara anterior em 4135 imagens, coletadas de exames de 2113 pacientes. O *dataset* é privado. Cada imagem foi dividida em duas, resultando em 8270 imagens, sendo 7375 imagens com ângulo aberto e 895 com ângulo fechado. O método utiliza duas abordagens. Na primeira foi utilizada segmentação das estruturas da câmara anterior para localizar a região onde se encontra o ângulo, e para cálculo de parâmetros clínicos. Esses parâmetros clínicos alimentam um SVM que tem como saída as probabilidades das imagens que pertenceram a cada uma das classes. Na outra abordagem foi utilizada uma arquitetura multicontexto, que contém dois níveis de extração de características, formados por blocos convolucionais, camadas de convolução e *Max pooling*. O primeiro nível recebe como entrada o recorte da região da câmara anterior, e o segundo recebe a imagem completa. As características extraídas por cada nível são concatenadas e alimentam um classificador, formado por camadas totalmente conectadas. O método utilizando aprendizagem profunda obteve um resultado melhor, com AUC de 0,930, que o método que utiliza os parâmetros clínicos, com AUC de 0,924.

No trabalho de [Fu et al. \(2019a\)](#) é proposta uma CNN para detecção de ângulos fechados em um *dataset* privado composto por 4135 imagens coletadas de exames de 2113 pacientes. Cada imagem foi dividida da mesma forma que a descrita no parágrafo anterior. Duas abordagens foram realizadas. A primeira calculou parâmetros quantitativos que possuem relevância clínica (ex: espessura da íris). Para que o cálculo de cada parâmetro pudesse ser feito, foi necessário realizar a segmentação das estruturas da câmara anterior. Um total de 6 parâmetros globais e 6 parâmetros locais (específicos da região do ângulo) foram calculados. Esses parâmetros foram utilizados como entrada para uma máquina de vetor de suporte (SVM, do inglês *Support Vector Machine*), utilizada como classificador. Na outra abordagem foi utilizada uma CNN, o modelo VGG16. Para agilizar a aprendizagem, foi utilizado um modelo pré-treinado e para aumentar os dados de treinamento foi realizado *data augmentation*. O método que utilizou uma CNN obteve um resultado melhor, com AUC de 0,96, que o método que utiliza os parâmetros clínicos, com AUC de 0,90.

O trabalho de [Xu et al. \(2019\)](#) comparou o desempenho de 3 CNNs para classificação de ângulos em um *dataset* privado composto por 4036 imagens, com cada imagem sendo dividida em duas, resultando sendo 1943 com ângulo aberto e 2093 com ângulo fechado, coletadas de 791 pacientes. As 3 CNNs foram treinadas para classificar os ângulos presentes

em cada imagem como aberto ou fechado. Para cada imagem de entrada, cada rede produz como resultado a probabilidade da imagem pertencer a uma das classes da escala de Shaffer. Esta escala, classifica o ângulo em cinco classes, p_0 , p_1 , p_2 , p_3 , p_4 . Como resultado final as probabilidades p_0 e p_1 são somadas para indicar ângulo fechado e as probabilidades p_2 , p_3 e p_4 são somadas para indicar ângulo aberto. O método obteve como melhor resultado um valor AUC de 0,933.

O trabalho [Fu et al. \(2019\)](#) utiliza uma arquitetura multinível, formada por 3 níveis, para detectar ângulos fechados em um *dataset* privado composto por 4135 imagens coletadas de exames de 2113 pacientes. As imagens foram divididas da mesma forma utilizada em trabalhos anteriores, resultando em 8270 imagens. O primeiro nível recebe como entrada a imagem completa, o segundo nível recebe metade da imagem e o terceiro nível recebe a região da imagem que contém o ângulo. É utilizado um método para detecção do SS, a partir da segmentação da extremidade inferior da córnea. Uma vez que o SS é encontrado, é realizada a extração da região onde se encontra o ângulo. As características extraídas por cada nível, formados por camadas convolucionais pré-treinadas, são concatenadas e alimentam um classificador. O método proposto obteve como resultado um valor AUC de 0.961. O método também foi testado em um *dataset* menor, com 701 imagens, alcançando como resultado um valor AUC de 0,952.

[Hao et al. \(2019\)](#) propôs um método para classificação das imagens composto por dois estágios: detecção automática da região onde se encontra o ângulo e uma CNN multi-escala para classificação das imagens. Inicialmente, as 4864 imagens OCT-SA pertencentes a um *dataset* privado, são divididas em duas, totalizando 9728 imagens. Estas imagens foram denominadas como escala 1. A partir destas imagens, foram localizadas as extremidades inferior e superior, e foi extraída uma região grosseira do ângulo. Estas imagens foram denominadas escala 2. A partir destas imagens foi localizada a linha de *Schwalbes*, a extremidade inferior da córnea. A partir desta linha foram extraídas imagens que contém a região exata do ângulo. Estas imagens foram denominadas escalas 3. Cada uma destas escalas foi utilizada para alimentar uma arquitetura que contém três níveis de extração de características, formados pelas camadas de convolução do modelo ResNet50. O método classifica os ângulos em abertos, estreitos e fechados. Foi obtido um valor AUC de 0,961.

No trabalho de [Yuan et al. \(2019\)](#) foi proposto um método para classificação do ângulo da câmara anterior das 1600 imagens pertencentes à base pública AGE ([FU et al., 2020](#)). Cada imagem foi dividida em duas resultando em 3200 imagens, sendo 7375 de imagens com ângulo aberto e 640 com ângulo fechado. O modelo utilizado para classificação das imagens, foi a CNN ResNet152. O modelo original sofreu algumas alterações, como modificação do tamanho de filtros usados em convoluções e do passo do *stride*, para aumentar a quantidade de dados resultantes dos blocos residuais. Alcançou como resultado um valor AUC de 1,000.

Hao et al. (2020) propôs uma rede multi-escala sequencial para classificação do ângulo em aberto, estreito e fechado (*synechia*), baseado em sequências de imagens OCT-SA. O primeiro passo após a divisão de cada imagem em duas, consiste de localizar e extrair a região do ângulo de cada imagem. Então uma sequência de imagens (regiões ACA) é obtida e alimenta os blocos que fazem a extração das características da sequência de imagens. O último passo consiste de utilizar um bloco LSTM (*Long Short Term Memory*) para estudar a dinâmica temporal dos mapas de características em nível sequencial e de camadas totalmente conectadas para realizar a predição. O método foi testado em uma base privada (3 classes), alcançando valor AUC de 0.82, e na base pública AGE (2 classes), alcançando valor AUC de 1,000.

A Tabela 1 apresenta um resumo dos principais trabalhos relacionados. Foram incluídos os trabalhos que propuseram métodos baseados em aprendizagem profunda, por serem os mais recentes e terem apresentado os melhores resultados até o momento da escrita deste documento.

Tabela 1 – Comparação entre diferentes métodos.

Trabalho	DataSet	Técnica	AUC	Sensibilidade
(FU et al., 2018)	Particular	Rede Multicontexto	0,945	0,888
(FU et al., 2019)	Particular	Rede Multinível	0,961	0,929
(FU et al., 2019)	Particular	Rede Multinível	0,952	0,873
(XU et al., 2019)	Particular	CNNs	0,933	-
(FU et al., 2019a)	Particular	CNN VGG16	0,960	0,900
(HAO et al., 2019)	Particular	Rede Multiescala	0,914	0,750
(YUAN et al., 2019)	AGE	CNN Residual	1,000	1,000
(HAO et al., 2020)	AGE	Rede Sequencial	1,000	1,000

Como as imagens OCT-SA fornecem uma visualização completa da câmara anterior, é possível utilizar tanto a imagem completa, que engloba os dois ângulos da câmara anterior, quanto uma região da imagem que engloba apenas um dos ângulos, para análise e cálculo de parâmetros clínicos. Este fato contribui para o emprego de métodos que utilizam redes neurais com uma arquitetura multinível, uma vez que este tipo de arquitetura tem como principal característica, a possibilidade de se utilizar mais de uma região da mesma imagem como entrada da rede. Devido a isso, com a utilização deste tipo de arquitetura mais características são extraídas das imagens, obtendo-se uma melhor representação das mesmas. Alguns dos trabalhos relacionados empregaram redes multinível para detecção de fechamento angular, obtendo resultados significativos. No entanto, os métodos utilizaram técnicas de segmentação de estruturas presentes nas imagens, com o propósito de capturar a região do ângulo iridocorneano. Além disso, pode-se considerar, no momento da escrita deste documento, que as redes utilizadas em alguns dos trabalhos relacionados não são tão profundas.

O presente trabalho propõe um método baseado em aprendizagem profunda para classificação automática do ângulo da câmara anterior usando imagens OCT-SA, com o objetivo de auxiliar no diagnóstico precoce da doença, possibilitando intervenções para revertê-la. Não é utilizado nenhuma técnica de pré-processamento, e nem de segmentação. O método proposto não se baseia na localização prévia de nenhuma estrutura presente nas imagens. É proposto uma rede neural convolucional multinível, treinada com diferentes partes da imagem e que utilize transferência de aprendizagem. Estima-se que, com essa abordagem, seja possível codificar as características com mais precisão e aumentar o desempenho do método.

3 Fundamentação Teórica

Este capítulo apresenta detalhes importantes sobre a anatomia do olho humano, sobre o glaucoma, com ênfase para o glaucoma primário de ângulo fechado (GPAF), o uso de imagens de Tomografia de Coerência Óptica do Segmento Anterior (OCT-SA) para avaliação de estruturas presentes na câmara anterior, além modelos de Redes Neurais Convolucionais (CNN, do inglês *Convolutional Neural Network*) que podem ser utilizadas em tarefas de classificação de imagens. Este trabalho visa a classificação do ângulo da câmara anterior presente em imagens OCT-SA em aberto ou fechado.

3.1 Glaucoma

O glaucoma é uma condição que tem como principal fator de risco uma elevada pressão intraocular (PIO), que provoca danos progressivos ao nervo óptico, levando inicialmente à cegueira bilateral, que pode evoluir para uma perda total da visão, Figura 4. Além da PIO elevada, outros fatores de risco como idade mais avançada, histórico familiar e raça são outros fatores que devem ser levados em consideração (SBG, 2009). Apesar de ser irreversível, a perda da visão pode ser evitada, caso procedimentos médicos sejam realizados em estágios iniciais da doença para evitar o seu progresso (OMS, 2019). No entanto, como o glaucoma é assintomático nos estágios iniciais, o diagnóstico precoce é dificultado, sendo que a doença geralmente é diagnosticada quando os efeitos do aumento da PIO já provocaram danos à visão (SARHAN; ROKNE; ALHAJJ, 2019).

Figura 4 – Simulação da progressão da perda visual devido ao glaucoma

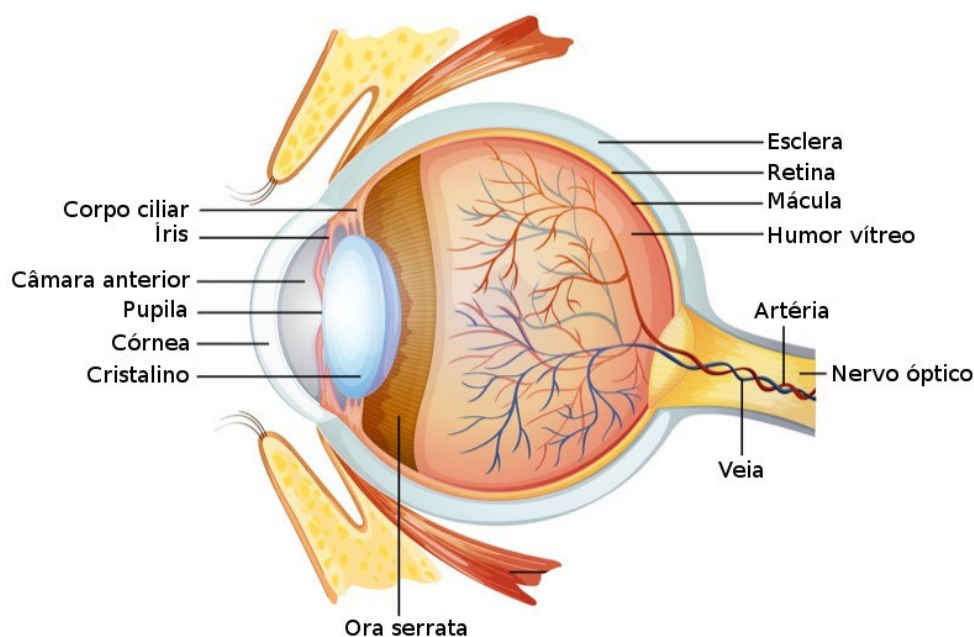


Fonte: Sociedade Brasileira de Glaucoma.

O olho humano, Figura 5 contém um líquido denominado humor aquoso, que flui de forma contínua pelo seu interior. Esse fluido tem como principal função a nutrição da córnea e do cristalino. O líquido é produzido e escoado através de uma região denominada câmara anterior (SBG, 2009), que é uma região delimitada anteriormente pela córnea, e posteriormente pela íris, por uma pequena porção do corpo ciliar e por parte do cristalino (SBG, 2012). Em pacientes com glaucoma, ocorre a diminuição do fluxo deste líquido,

o que provoca o acúmulo do mesmo, levando a um aumento da PIO. Esse aumento da pressão provoca, no decorrer de alguns anos, danos as fibras nervosas do nervo óptico, causando danos à visão (LI et al., 2020).

Figura 5 – Anatomia do olho humano



Fonte: (ARAÚJO et al., 2017)

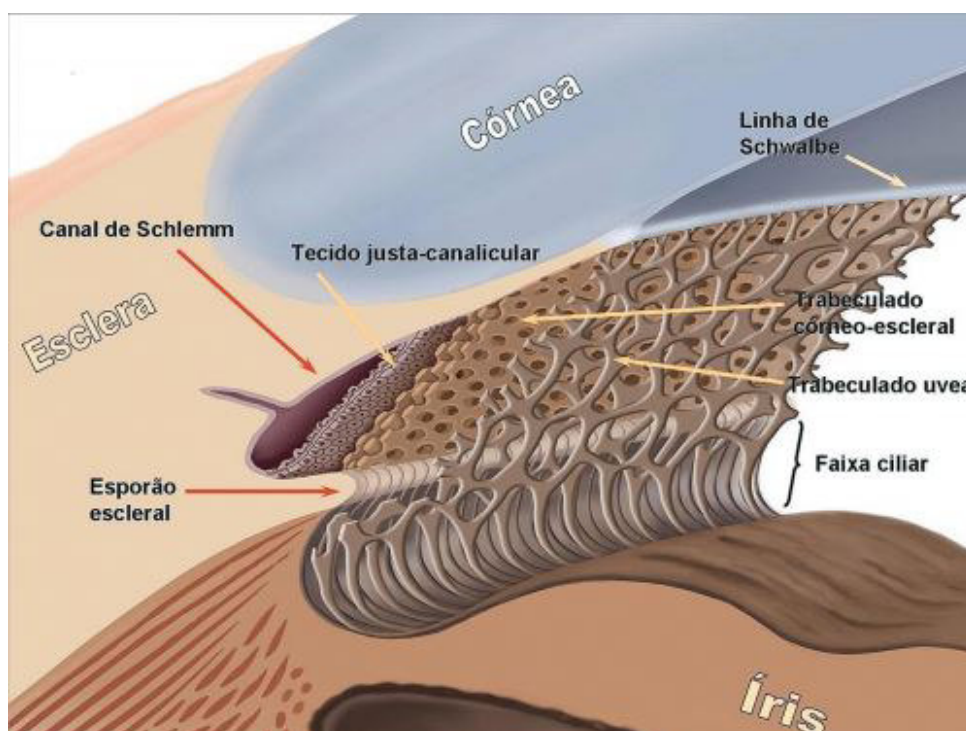
O glaucoma têm como principais tipos o glaucoma primário de ângulo aberto (GPAA) e o glaucoma primário de ângulo fechado (GPAF) (QUIGLEY; BROMAN, 2006). Embora o GPAF atinja um terço do total de pacientes com glaucoma, esta condição causa cegueira em um número semelhante ao GPAA (SBG, 2012). Isso deve-se ao fato do GPAF ser uma forma mais agressiva da doença, sendo seu curso mais severo, levando a maiores morbidades (CHANSANGPETCH et al., 2018). Devido à diferenças em termos de fisiopatologia entre as formas da doença, as estratégias de tratamento são substancialmente diferentes. Logo, a diferenciação entre os dois tipos é crítica para o desenvolvimento da estratégia apropriada de tratamento, para que seja possível evitar o perdas irreversíveis da visão (CHANSANGPETCH; ROJANAPONGPUN; LIN, 2018).

A principal característica do GPAF é o processo de fechamento do ângulo formado pela junção da íris com a córnea, conforme ilustrado na Figura 6. Este mecanismo é caracterizado pelo fechamento parcial ou completo do ângulo, decorrente de condições anatômicas que propiciam o surgimento de goniossinéquias ¹, com frequente elevação da PIO SBG (2012). A elevação da pressão pode ocorrer de forma aguda, manifestação mais

¹ Aposição ou adesão da periferia da íris à parede externa do ângulo da câmara anterior, o que indica a presença de contato iridotrabecular (TIC, do inglês *Trabecular Iris Contact*) SBG (2012).

agressiva e sintomática, intermitente, caracterizada por manifestações mais brandas e recorrentes ou crônica, onde a evolução ocorre de forma assintomática. A classificação dos pacientes baseia-se na evolução da doença, sendo que existem três possíveis casos: suspeito de fechamento angular, fechamento angular primário, onde já existe aposição iridotrabecular e/ou aumento da pressão intraocular, e glaucoma primário de ângulo fechado, onde além das alterações encontradas em pacientes com fechamento angular, existe a presença de neuropatia óptica glaucomatosa (SBG, 2012).

Figura 6 – Estruturas presentes na região do ângulo da câmara anterior



Fonte: (SBG, 2009)

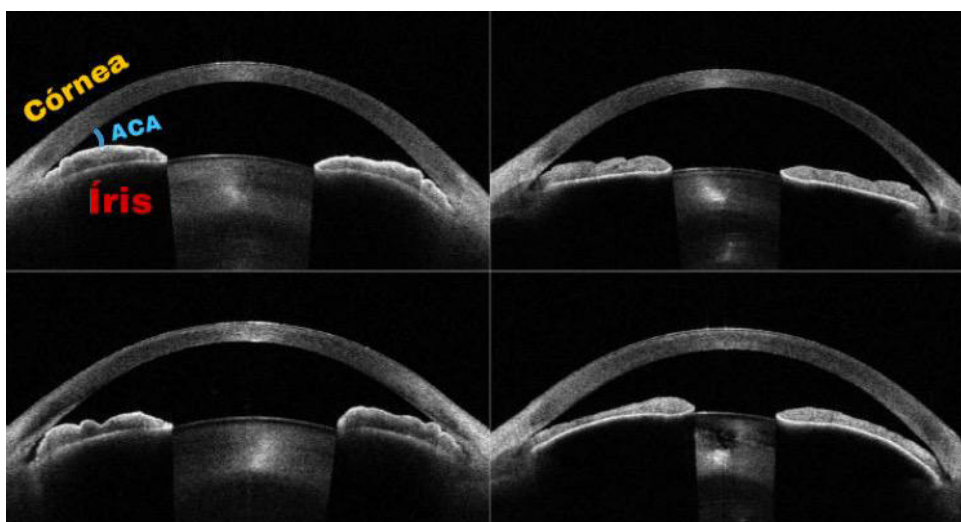
3.2 Tomografia de Coerência Óptica do Segmento Anterior

Para avaliação do ângulo da câmara anterior, diferentes exames de imagem têm sido utilizados, podendo ser citados a gonioscopia, biomicroscopia ultra-sônica (UBM) e tomografia de coerência óptica do segmento anterior (ANG et al., 2018). A gonioscopia é considerado o exame padrão para avaliar a anatomia do ângulo da câmara anterior (ACA) e determinar a presença de contato entre íris e malha trabecular (FU et al., 2019a). No entanto, o exame apresenta algumas desvantagens, entre elas dificuldades técnicas, vulnerabilidade a erros e baixa reprodutibilidade (CHANSANGPETCH; ROJANAPONG-PUN; LIN, 2018). A Tomografia de Coerência Óptica (OCT, do inglês *Optical Coherence Tomography*) por sua vez, é um exame rápido e que não necessita que haja contato entre o

equipamento e o os olhos dos pacientes, o que facilita consideravelmente a aquisição das imagens (CHANSANGPETCH et al., 2018).

Através de OCTs podem ser obtidas imagens transversais de apenas um ângulo da câmara anterior (ACA), ou de todo o segmento anterior. Desta forma, uma variedade de parâmetros clínicos quantitativos podem ser obtidos de imagens OCT-SA (OCT de Segmento Anterior), referentes ao ACA (ex. distância de abertura do ângulo), ou à câmara anterior (ex. área da câmara anterior) (CHANSANGPETCH; ROJANAPONGPUN; LIN, 2018). Por meio da análise quantitativa desses parâmetros, é possível determinar aspectos anatômicos do segmento anterior, além do ACA, como íris, lentes e corpo ciliar (FU et al., 2017). Além da análise quantitativa, as imagens OCT-SA permitem uma análise qualitativa. Utilizando o esporão escleral (SS, do inglês *scleral spur*) como *landmark*, é possível determinar a presença de contato iridocorneano anterior ao SS, característica presente em pacientes com fechamento angular primário. Além disso, a OCT permite a visualização da morfologia do segmento anterior em quatro dos principais mecanismos de fechamento angular primário: bloqueio pupilar, íris em platô, fechamento angular induzido pelo cristalino e aumento do diâmetro da lente, Figura 7. Deste modo, é possível afirmar que a OCT-SA têm como vantagens em relação à gonioscopia, uma análise que não se restringe ao ACA, visualização de características da íris, como sua configuração, predição e monitoramento e subclassificação de casos de fechamento angular (CHANSANGPETCH; ROJANAPONGPUN; LIN, 2018).

Figura 7 – Morfologia dos principais mecanismos de fechamento angular. Bloqueio pupilar (Imagem superior esquerda), íris em platô (Imagem superior direita), fechamento induzido pelo cristalino (Imagem inferior esquerda) e aumento do diâmetro da lente (Imagem inferior direita).



Fonte: (CHANSANGPETCH; ROJANAPONGPUN; LIN, 2018)

3.3 Redes Neurais Convolucionais

As redes neurais convolucionais (CNN, do inglês *Convolutional Neural Network*) são um modelo de rede neural profunda. Enquanto redes neurais simples são formadas por poucas camadas, redes profundas são formadas por várias camadas, uma camada de entrada, uma de saída e várias camadas escondidas. Outra característica deste modelo de rede é ser do tipo *feedforward*, ou seja, não existe realimentação. Os dados percorrem a rede em uma única direção, da entrada para a saída (GOODFELLOW; BENGIO; COURVILLE, 2016). Modelos de CNN têm alcançado resultados expressivos em tarefas de classificação e segmentação de diferentes tipos de imagens médicas (SHANG et al., 2019), (SHI et al., 2019), (RAMAN et al., 2019).

A principal característica deste modelo é a realização da operação matemática denominada convolução. (Equação 3.1).

$$s(t) = (x * w)(t) \quad (3.1)$$

onde x representa o valor de uma variável em um instante t , e w representa uma função de ponderação. Em redes convolucionais, o primeiro argumento da equação refere-se a um dado de entrada, por exemplo os pixels que formam uma região de uma imagem, e o segundo refere-se a um filtro (*kernel*). O resultado é geralmente denominado mapa de características (*feature maps*). Redes Convolucionais são redes neurais que utilizam a operação matemática convolução em pelo menos uma de suas camadas (GOODFELLOW; BENGIO; COURVILLE, 2016). Estas camadas são denominadas convolucionais, e fazem parte do extrator de características visuais. Como os computadores utilizam dados discretos, pode-se definir a convolução discreta como sendo a Equação 3.2.

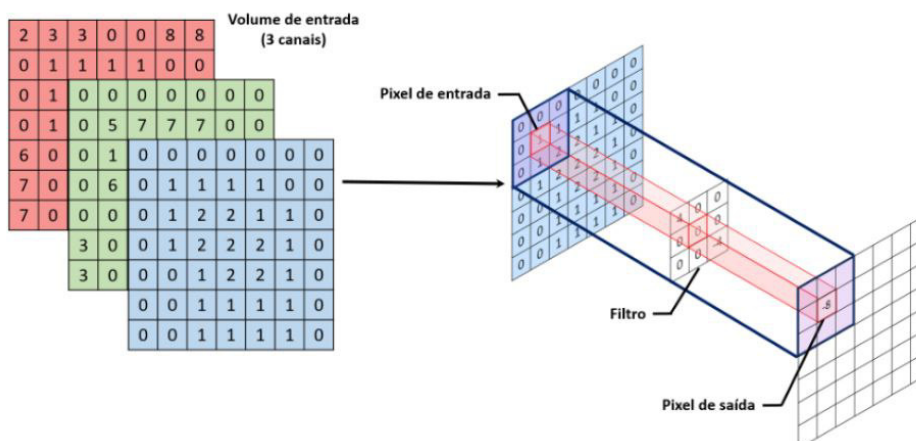
$$s(t) = (x * w)(t) = \sum_{a=0}^N x(a)w(t - a) \quad (3.2)$$

onde x e w serão definidos apenas para valores inteiros de t .

A Figura 8 ilustra a operação de convolução entre uma região de uma imagem e um filtro. Esta região da imagem é chamada de campo receptivo local (ACADEMY, 2019). Este campo receptivo é deslizado por toda a imagem, a um passo medido em *pixels* denominado *stride length*. Apesar da Figura 8 ilustrar a convolução sendo realizada com os pixels variando entre 0 e 255, antes do cálculo acontece uma normalização. Cada pixel é normalizado para um valor entre zero e um.

Em redes neurais que utilizam multiplicação de matrizes, o resultado (saída) de um neurônio é utilizado como entrada para todos os neurônios da próxima camada. Como essa conectividade “total” entre neurônios não acontece em uma camada de convolução, essas redes possuem interações esparsas. Isso acontece devido ao filtro possuir dimensões menores

Figura 8 – Convolução Discreta



Fonte: (ARAÚJO et al., 2017)

que a entrada (GOODFELLOW; BENGIO; COURVILLE, 2016), como representado na Figura 8. Com isso é necessário a realização de menos operações matemáticas diminuindo o custo computacional. Através da convolução é possível detectar pequenas, porém relevantes características que serão utilizadas para classificação das imagens. Além disso, em uma rede convolucional existe o compartilhamento de pesos. Um mesmo filtro é utilizado para realizar a convolução em toda a imagem, o que não acontece em uma rede totalmente conectada, onde existe um peso para cada conexão entre o neurônio de uma camada e os neurônios da camada seguinte. Isso implica em dizer que existem menos parâmetros para serem analisados e ajustados, o que diminui o custo computacional. Outra vantagem dessas Redes é o fato de necessitarem de pouco pré-processamento. Enquanto em outros métodos o design dos filtros é feito por algum especialista, com o treinamento necessário, as Redes Convolucionais são capazes de aprender a ajustar os filtros (SAHA, 2018).

As redes convolucionais possuem, além de camadas que realizam a operação de convolução, camadas que realizam uma operação denominada *pooling*. Em geral, o extrator de características deste tipo de rede é formado por blocos que contém duas ou mais camadas de convolução seguidas de uma camada de *pooling*. Esta função substitui a saída de uma camada por um índice estatístico. Por exemplo, a operação *Max Pooling*, predominantemente utilizada, reporta a saída máxima dentre as saídas existentes em uma vizinha retangular (GOODFELLOW; BENGIO; COURVILLE, 2016). Assim como a camada de convolução, a camada de *Max Pooling* também reduz o tamanho de uma imagem, uma vez que somente retorna um valor de uma determinada região da imagem de entrada. Além disso, a utilização do *Max Pooling* produz invariância à pequenas translações da imagem de entrada. Mesmo que haja alguma variação deste tipo, sendo que são imagens de um mesmo objeto, essa invariância evita que a rede classifique as variações de uma mesma imagem como imagens diferentes.

Em CNNs utilizadas para classificação de imagens, após os blocos formados pelas camadas de Convolução/Max Pooling, que tem por finalidade extrair características de imagens de entrada, são utilizadas camadas que irão utilizar estas características para gerar o resultado final da classificação. Os classificadores dos modelos são formados por camadas totalmente conectadas, denominadas camadas densas, exatamente iguais a uma rede MLP (do inglês, *Multi Layer Perceptron*) (ARAÚJO et al., 2017). Isto significa que todos os neurônios de uma camada estão conectados aos neurônios da camada seguinte (GOODFELLOW; BENGIO; COURVILLE, 2016). Em termos matemáticos, um neurônio pode ser descrito pelas Equações 3.3 e 3.4

$$h_j = \sum_{i=1}^m w_{ij}x_i \quad (3.3)$$

$$y_j = \varphi(h_j + b_j) \quad (3.4)$$

onde x_i são os m sinais de entrada, w_{ij} são os pesos sinápticos do neurônio j , e b_j corresponde ao *bias*, responsável por realizar o deslocamento da função de ativação definida por φ (ARAÚJO et al., 2017).

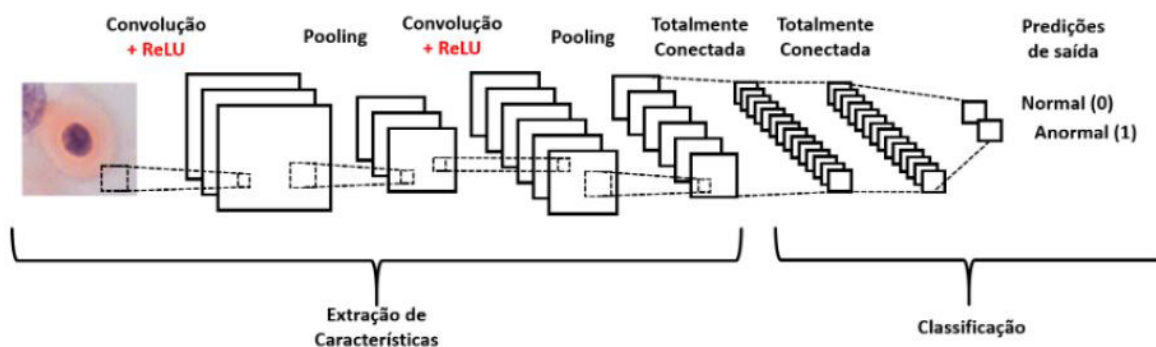
Para que seja possível utilizar os mapas de características como entrada do classificador é necessário alterar o formato matricial dos dados para vetorial. Esse processo é denominado *flattening*. Em geral, as primeiras camadas densas utilizam a função ReLu (*Rectified Linear Unit* Equação 3.5), como função de ativação, e a última camada do classificador utiliza a função de ativação *softmax* (BISHOP, 2006), definida pela Equação 3.6

$$f(x) = \max(0, x) \quad (3.5)$$

$$\sigma(Z)_i = \frac{e^{z_i}}{\sum_{j=1}^K e^{z_j}} \quad (3.6)$$

onde z é um vetor de números reais que é normalizado em uma distribuição de K probabilidades. Esta função tem como saída uma distribuição de probabilidades, que é utilizada para classificação das imagens, e o número de neurônios da camada é igual ao número de classes em que as imagens podem ser classificadas. Outro parâmetro importante dos classificadores é o percentual de *dropout*, que consiste na remoção aleatória de um percentual de neurônios a cada iteração de treinamento, com o propósito de evitar *overfitting* e reduzir o tempo de treinamento (GOODFELLOW; BENGIO; COURVILLE, 2016). A arquitetura básica de uma CNN utilizada para classificação de imagens é ilustrada na Figura 9.

Figura 9 – Arquitetura de uma rede convolucional



Fonte: (ARAÚJO et al., 2017)

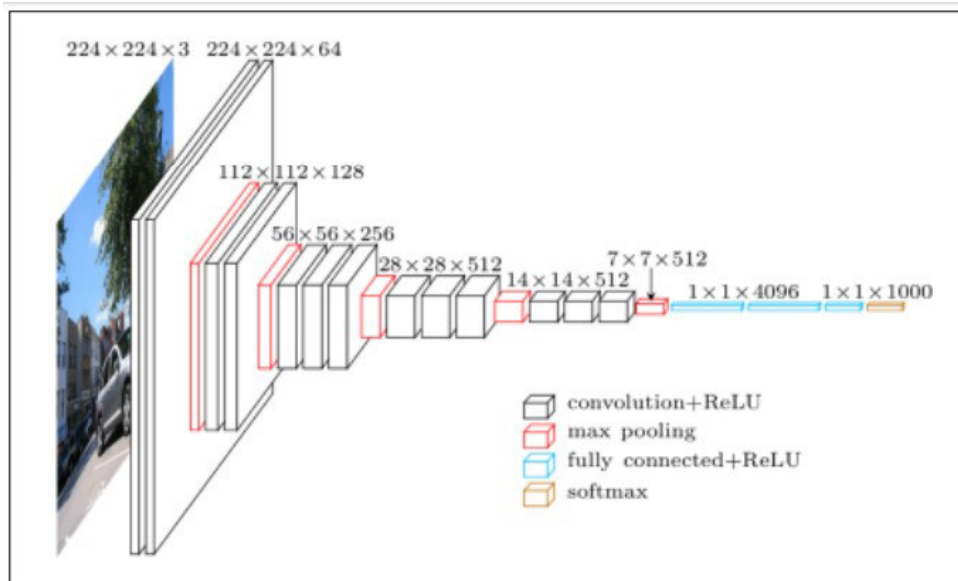
3.4 Arquiteturas de CNN

Existem atualmente diferentes arquiteturas de Redes Convolucionais. Os desenvolvedores têm um objetivo claramente definido, que é alcançar o melhor desempenho possível, sendo a meta classificar corretamente diferentes imagens. Anualmente é realizado um desafio denominado ImageNet Large-Scale Visual Recognition Challenge (ILSVRC) (RUSSAKOVSKY et al., 2015), que tem servido como teste para novas gerações de sistemas de classificação de imagem em larga escala (SIMONYAN; ZISSERMAN, 2014).

A rede *Visual Geometry Group* (VGG), (SIMONYAN; ZISSERMAN, 2014), tem como principal característica o fato de utilizar filtros pequenos, o que possibilitou uma rede com muitas camadas. Um dos grandes problemas ao se aumentar o número de camadas é o desaparecimento do gradiente. A solução encontrada foi utilizar uma pilha de camadas convolucionais, onde são utilizados filtros com um campo receptivo muito pequeno (3x3). A utilização desses filtros reduz o custo computacional, sendo esta, a grande contribuição da VGG. Além disso esta rede utiliza cinco camadas de *Max Pooling* com uma janela de tamanho 2x2. Após as camadas Convolucionais há três camadas totalmente conectadas: as duas primeiras possuem 4096 neurônios cada e a última possui 1000. A Figura 10 apresenta a arquitetura.

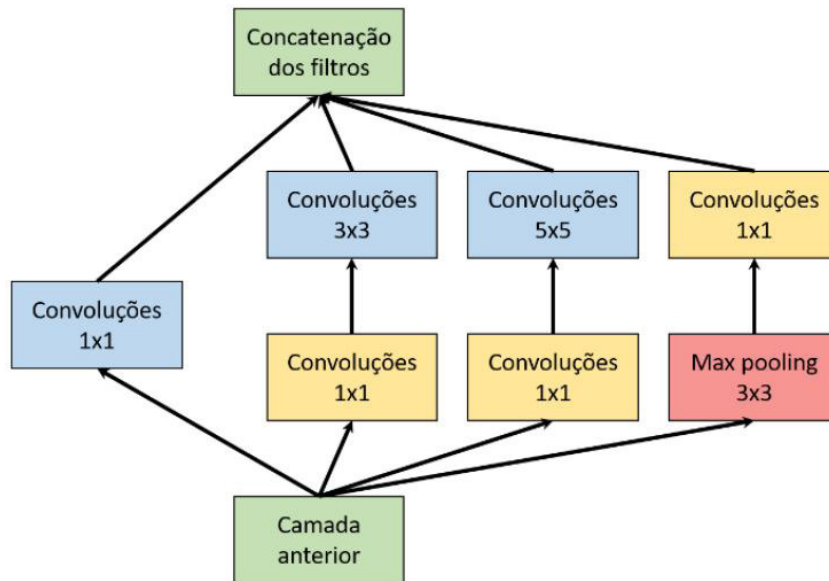
A rede Inception (SZEGEDY et al., 2014), é uma Rede Convolucional Profunda que disputou o ILSVRC em 2014, atingindo o estado da arte naquele ano para classificação de imagens. Esta arquitetura introduziu a ideia de que as camadas não precisavam ser executadas de forma sequencial, mas que podiam ser executadas de forma paralela. Esta é a principal característica desse modelo (Figura 11). A grande vantagem dos módulos Inception é o uso com filtros 1x1 para reduzir o número de características dos blocos antes de realizar convoluções com filtros maiores (ARAÚJO et al., 2017). A utilização destes blocos permitiu o aumento da largura e profundidade da rede, sem aumento considerável do custo computacional.

Figura 10 – Arquitetura da rede VGG



Fonte: (SIMONYAN; ZISSERMAN, 2014)

Figura 11 – Módulo Inception

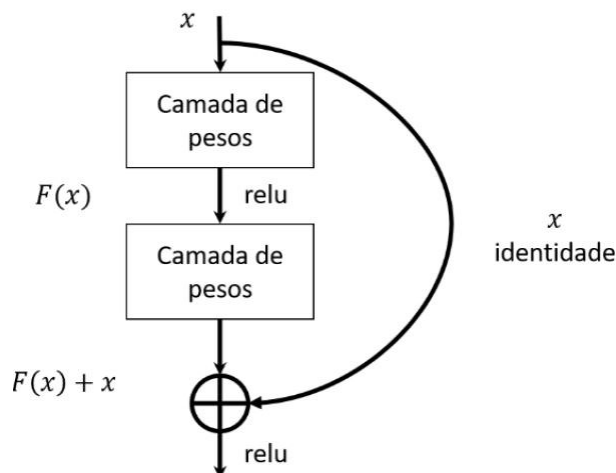


Fonte: (ARAÚJO et al., 2017)

As redes residuais (ResNet) (HE et al., 2016) trouxeram como grande contribuição conexões residuais entre as camadas de uma CNN. Em um bloco residual, um volume de entrada passa por uma sequência de convoluções e de ativações Relu. O resultado dessas operações é somado ao valor da entrada como ilustrado na Figura 12. Em CNNs com conexões apenas *feedforward*, o volume de saída de um bloco convolucional é completamente diferente do volume de entrada x . Em relação ao modelo ResNet, o volume de saída $H(x)$, é

somente uma alteração do volume de entrada (ARAÚJO et al., 2017). Com estas conexões residuais foi possível implementar um modelo com até 152 camadas.

Figura 12 – Bloco Residual



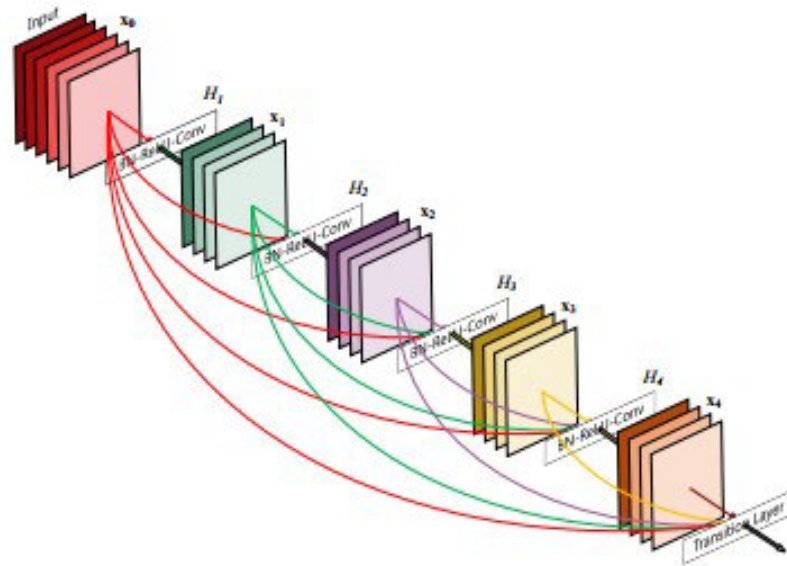
Fonte: (ARAÚJO et al., 2017)

A DenseNet (*Dense Convolutional Network*) (HUANG; LIU; WEINBERGER, 2016), é uma rede convolucional que possui como principal característica conexões densas entre os blocos convolucionais. Redes Convolucionais tradicionais conectam a saída de uma camada à entrada da próxima camada. Essa característica é uma das causas do desaparecimento do gradiente em redes muito profundas. As redes residuais, adicionaram uma conexão diferente das conexões até então utilizadas. A vantagem da conexão residual é que o gradiente pode fluir diretamente das últimas camadas para as primeiras. Já o modelo DenseNet propõe conexões diretas de qualquer camada para todas as camadas subsequentes. O modelo dessa conexão densa é ilustrado na Figura 13. Além dos blocos com conexões densas, existem camadas com conexões apenas *feedforward* que formam o classificador.

3.5 Transferência de Aprendizado

Um problema recorrente em tarefas de classificação de imagens utilizando aprendizagem profunda, deve-se ao fato de ser necessário um conjunto muito grande de imagens para que seja possível realizar um treinamento eficiente de uma CNN (ARAÚJO et al., 2017). Logo, raramente é realizado o treinamento de uma CNN com inicialização aleatória de pesos, pois em geral conjuntos de imagens médicas não são suficientemente grandes. Uma das técnicas utilizadas para sanar esta limitação é a transferência de aprendizado. Esta técnica consiste de se utilizar o que foi aprendido na classificação de um conjunto de dados para aumentar a capacidade de generalização em outro conjunto (GOODFELLOW;

Figura 13 – Conexões do Modelo DenseNet



Fonte: (HUANG; LIU; WEINBERGER, 2016)

BENGIO; COURVILLE, 2016). Neste caso, o primeiro conjunto é muito maior que o segundo. Uma prática comum é a utilização dos pesos de uma CNN previamente treinada na base *ImageNet*, que possui mais de 1 milhão de imagens e 1000 classes (ARAÚJO et al., 2017).

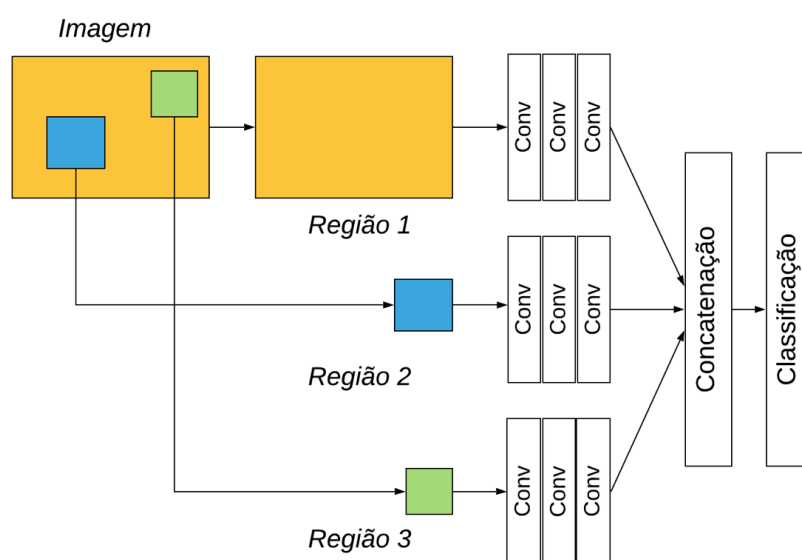
Redes pré-treinadas podem ser utilizadas de diferentes formas em tarefas de classificação de imagens. Em geral, apenas as camadas que formam o extrator de características de uma CNN são aproveitadas de modelos pré-treinados, sendo que um novo classificador, ajustado a nova tarefa, é adicionado a rede. A estratégia *fine-tuning* consiste de se dar continuidade ao treinamento de uma rede previamente treinada. É possível a realização de *fine-tuning* de todas as camadas de uma CNN, ou somente das últimas camadas. Isto deve-se ao fato de que as primeiras camadas de uma rede convolucional possuem extratores mais genéricos, como detectores de bordas, enquanto as camadas mais profundas possuem detalhes específicos da base com a qual a rede foi previamente treinada (ARAÚJO et al., 2017). Também é possível utilizar uma CNN pré-treinada como um extrator de características. Neste caso, o classificador é removido, e as características extraídas são utilizadas como entrada para um classificador que é mais rapidamente treinado (ARAÚJO et al., 2017), por exemplo, *Random Forest* (HO, 1995).

3.6 Arquiteturas Multinível

Redes neurais convolucionais tem alcançado resultados significativos em tarefas de classificação de imagens, através da extração de características discriminativas. Com o

objetivo de se obter uma representação hierárquica ainda mais discriminativa, métodos que utilizam redes multiescala ou multi-regiões têm sido propostos (FU et al., 2019). Uma rede convolucional multinível, ilustrada na Figura 14, é uma rede que possui em sua estrutura mais de um nível de extração de características visuais, sendo que cada nível recebe como entrada uma imagem. Devido a isso, é possível utilizar diferentes regiões de uma mesma imagem como entrada da rede, com cada nível realizando a extração de características de cada região, aumentando desta forma o conjunto de características que representa as imagens. Após os níveis de extração, há uma camada que realiza a concatenação das características que são utilizadas como entrada para o classificador.

Figura 14 – Exemplo de Arquitetura de uma Rede Multinível



Fonte: Acervo do Autor

3.7 Métricas de Avaliação

Com o propósito de avaliar se os resultados obtidos por um algoritmo de classificação são bons ou ruins, existem funções matemáticas que permitem avaliar a capacidade de erro e acerto do algoritmo. A Tabela 2 apresenta a matriz de confusão. Essa matriz é utilizada para organizar os resultados obtidos pelo algoritmo. Cada linha da matriz representa as instâncias de uma classe prevista, enquanto cada coluna representa instâncias da classe real. Essa matriz permite calcular as métricas utilizadas para verificar se o modelo está fazendo previsões confusas na tarefa de classificação.

Tabela 2 – Matriz de Confusão.

		Resultado Real	
		Classe A	Classe B
Resultado Previsto	Classe A	TP	FP
	Classe B	FN	TN

- Verdadeiros Positivos (TP): referem-se aos casos pertencentes a classe positiva (classe A) e que foram preditos corretamente como pertencentes a esta classe;
- Falsos Negativos (FN): referem-se aos casos pertencentes a classe positiva e que foram preditos incorretamente como sendo da classe negativa (classe B);
- Falsos Positivos (FP): referem-se aos casos pertencentes a classe negativa e que foram preditos incorretamente como sendo da classe positiva;
- Verdadeiros Negativos (TN): referem-se aos casos pertencentes a classe negativa e que foram preditos corretamente como sendo pertencentes a esta classe.

Uma das métricas mais utilizadas, a acurácia (ACU) é definida como a soma do número de verdadeiros positivos e verdadeiros negativos dividida pelo total de exemplares, Equação 3.7. No entanto, utilizar apenas a acurácia pode levar a uma interpretação errada, principalmente para conjuntos de dados desbalanceados.

$$ACU = \frac{TP + TN}{TP + TN + FP + FN} \quad (3.7)$$

A sensibilidade, também chamada de taxa de verdadeiros positivos, é a razão entre o número de exemplares classificados corretamente como positivos e o número total de exemplares positivos reais, Equação 3.8. A especificidade, também chamada de taxa de verdadeiros negativos, é a razão entre o número de exemplares classificados corretamente como negativos e o número total de exemplares negativos, Equação 3.9. A precisão é a razão entre o número de exemplares classificados corretamente como positivos e o número total de exemplares classificados como negativos 3.10.

$$SEN = \frac{TP}{TP + FN} \quad (3.8)$$

$$ESP = \frac{TN}{TN + FP} \quad (3.9)$$

$$PRE = \frac{TP}{TP + FP} \quad (3.10)$$

Além das funções citadas anteriormente, uma métrica bastante utilizada para avaliar o desempenho de modelos, principalmente para tarefas em que a base de dados é desbalanceada, é a área sob a curva de característica de operação do receptor, ou área sob a curva ROC (ROC, do inglês *Receiver operating characteristic*). A curva ROC indica o quanto o modelo pode distinguir as classes, e é o resultado da taxa de verdadeiros positivos em função da taxa de falsos positivos. O valor desta área é denominado valor AUC, (AUC, do inglês *Area Under Curve*), e quanto maior este valor, melhor o modelo realiza a distinção correta entre as classes.

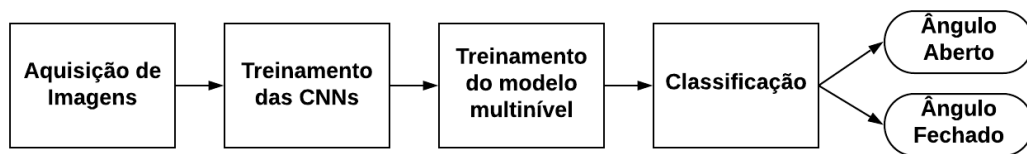
3.8 Considerações Finais

Este capítulo apresentou os principais conceitos em que se baseia a metodologia proposta neste trabalho, abordando as principais características do glaucoma de ângulo fechado, descrevendo as imagens de tomografia de coerência óptica do segmento anterior, detalhando a arquitetura das Redes Neurais utilizadas e apresentado o método para avaliação dos resultados obtidos.

4 Metodologia

Este capítulo apresenta a metodologia proposta para classificação do ângulo da câmara anterior em imagens OCT-SA. O método proposto para classificação é dividido em quatro etapas: aquisição e preparação de imagens, treinamento de extratores de características alimentados por diferentes regiões das imagens, Redes Multinível formadas a partir da combinação de CNNs e finalmente, classificação das imagens usando um modelo estimado. A Figura 15 apresenta as etapas do método proposto.

Figura 15 – Fluxograma da metodologia proposta.



Acervo do autor.

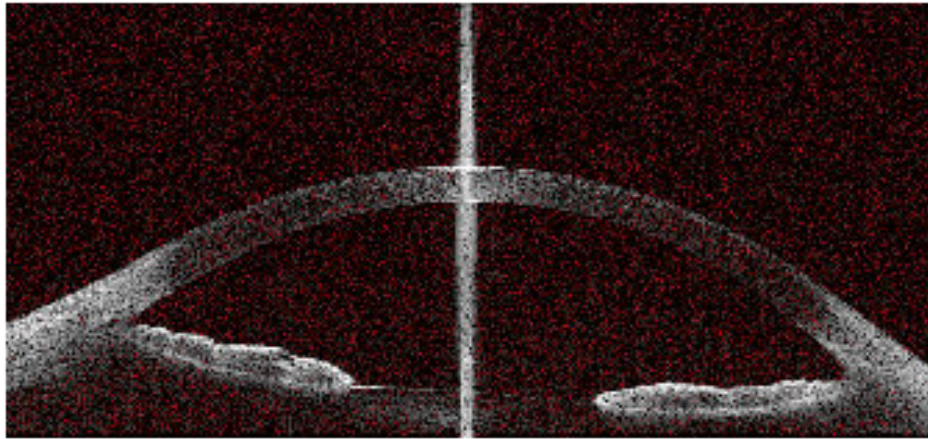
4.1 Aquisição das Imagens

As imagens OCT-SA utilizadas neste trabalho foram disponibilizadas para a realização do Desafio de Avaliação de Glaucoma de Ângulo Fechado (AGE) (FU et al., 2020). Este desafio foi idealizado tendo como um dos principais propósitos, fornecer uma base de dados para melhor avaliação de métodos existentes para classificação, sendo até o momento da escrita deste trabalho, a única base pública de imagens OCT-SA, utilizadas para detecção de glaucoma de ângulo fechado, com base na classificação do ângulo da câmara anterior. Esta base fornece um grande conjunto de imagens, com anotações confiáveis para localização de uma estrutura presente na câmara anterior, denominada esporão escleral, e para a classificação do ângulo da câmara anterior. Devido ao fato de ser pública, esta base permite que diferentes abordagens sejam comparadas adequadamente, e que técnicas de aprendizagem profunda para a detecção de ângulo fechado sejam desenvolvidas, uma vez que a maioria dos modelos redes neurais convolucionais precisa de um grande número de imagens para treinamento.

A base disponibilizada é composta por volumes de imagens, sendo que cada volume possui 16 imagens, juntamente com a anotação da classificação de cada ângulo presente na câmara anterior, sendo que cada ângulo pode ser classificado como aberto ou fechado. A classe ângulo aberto é representada pelo rótulo 0 (zero) e a classe ângulo fechado é representado pelo rótulo 1 (um). Ao total a base possui 100 volumes, sendo 20 pertencentes

à classe de ângulo fechado (320 imagens), e 80 pertencentes à classe de ângulo aberto (1280 imagens), totalizando 1600 imagens. Um exemplo é apresentado na Figura 16.

Figura 16 – Imagem com dois ângulos presentes na câmara anterior.



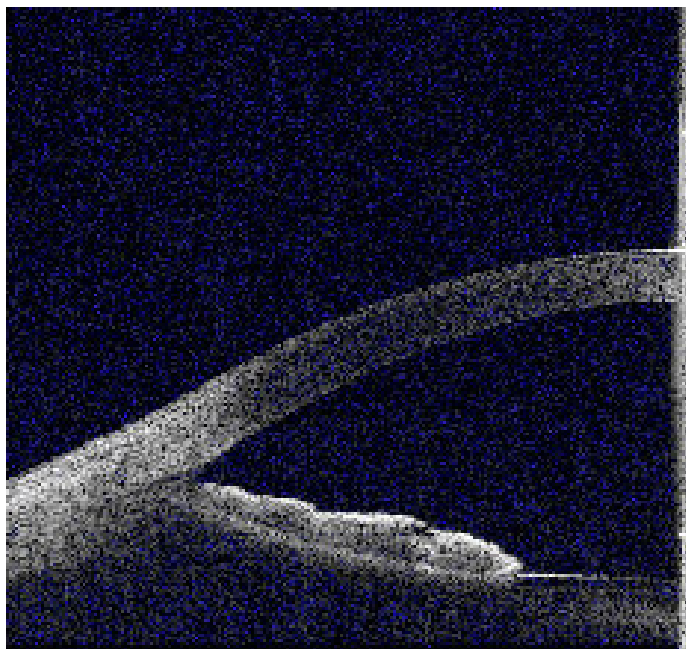
Acervo do autor.

As imagens da base foram adquiridas com o tamanho de 2130 pixels de altura e 998 pixels (largura). Não foi aplicada nenhuma técnica de pré-processamento neste trabalho, apenas o redimensionamento das imagens para o tamanho 224x224 pixels, para ser possível utilizá-las como entrada para as arquiteturas de CNN que foram treinadas para classificação das imagens, já que as mesmas têm sua camada de entrada projetada para receber imagens com essas dimensões. Além disso, a diminuição das imagens tem como consequência a diminuição do custo computacional. O próximo passo foi realizar a captura do ângulo esquerdo, gerando imagens que contém apenas um ângulo. Um exemplo é apresentado na Figura 17. Estas novas imagens foram geradas com o propósito de se realizarem testes utilizando-se arquiteturas multiníveis. A escolha de utilização de imagens com apenas um ângulo, deve-se ao fato de existirem parâmetros clínicos relacionados à região local do ângulo. Como os dois ângulos presentes na câmara anterior possuem a mesma classificação, seria possível utilizar tanto o ângulo esquerdo quanto o ângulo direito. Também foi possível analisar com qual imagem, com um ou dois ângulos, os modelos de redes neurais teriam melhor desempenho.

4.2 Extração de características visuais

Para realização da extração de características visuais das imagens OCT-SA, foram utilizadas redes neurais convolucionais. Redes Neurais Convolucionais têm alcançado um desempenho superior em tarefas de classificação de imagens, quando comparadas com técnicas que requerem uma parametrização manual por parte do cientista (*handcrafted features*). Devido ao fato de serem redes neurais profundas, as características extraídas

Figura 17 – Imagem com o ângulo esquerdo da câmara anterior.



Acervo do autor.

pelas redes são denominadas *deep features*. Dentre as grandes vantagens destas redes estão o ajuste dos pesos utilizados nos filtros de forma automática e necessidade de pouco ou nenhum emprego de técnicas de pré-processamento. No entanto, o uso de redes neurais tem um elevado custo computacional e as redes necessitam de um grande número de imagens para treinamento e ajuste de seus parâmetros.

Uma forma de se reduzir o tempo gasto pelas redes para ajustes de parâmetros e de suprimir uma eventual falta de dados para treinamento é a transferência de conhecimento (*transfer learning*). Esta técnica consiste em usar redes neurais previamente treinadas, que tiveram seus pesos ajustados para extrair e reconhecer características de um determinado grupo de imagens. O conhecimento adquirido por redes pré-treinadas pode ser utilizado para extração de características de novos tipos de imagem. Desta forma, evita-se a necessidade de treinamento da rede com valores aleatórios para os pesos. Outra vantagem dessa técnica é que o vetor de características extraídos pelas redes podem ser utilizados como entrada para outros classificadores, diferentes do classificador original utilizado em cada modelo de rede neural.

Neste trabalho, para realizar a transferência de aprendizagem, todos os modelos de CNN utilizados foram carregados com os pesos sinápticos obtidos no *ImageNet Large Scale Visual Recognition Challenge* (RUSSAKOVSKY et al., 2015) (Desafio de reconhecimento visual em grande escala ImageNet). Neste desafio, algoritmos são testados para detecção de objetos e classificação de imagens em larga escala, permitindo que seja verificado o

progresso nestes campos de pesquisa da visão computacional. Uma vantagem é que os modelos são treinados em uma grande variedade de imagens.

Além de carregar os modelos pré-treinados, foi realizado *fine-tuning* (ajuste fino). Esta técnica consiste de ajustar modelos treinados anteriormente para resolução de um novo problema. Desta forma, necessita-se de menos tempo para treinamento dos modelos. Em todos os modelos de CNN utilizados neste trabalho, foi feita a substituição das últimas camadas da rede, que são alimentadas com os vetores de características e realizam a tarefa de classificação. Desta forma um classificador adequado ao problema proposto neste trabalho foi adicionado. As duas primeiras camadas do classificador usam a função Relu como função de ativação, e a última usa a regressão softmax, que tem como resultado a probabilidade da imagem pertencer a uma das classes, sendo a maior probabilidade utilizada como resultado final de classificação, em ângulo aberto ou fechado. A função de entropia cruzada foi usada como uma função de perda.

Neste trabalho, foram utilizadas as seguintes arquiteturas de CNN: VGG16 e VGG19 (SIMONYAN; ZISSERMAN, 2014), InceptionV3 (SZEGEDY et al., 2014), ResNet50 (HE et al., 2016) and DenseNet121 (HUANG; LIU; WEINBERGER, 2016). Cada um desses modelos possui características na sua arquitetura, detalhados no capítulo [Fundamentação Teórica](#), que foram primordiais para o elevado desempenho no desafio Imagenet. Cada uma das CNN foi treinada para extrair características e classificar as imagens que contém dois e um ângulo da câmara anterior. Desta forma foram obtidos dois modelos a partir de cada rede, totalizando dez modelos. Os modelos tiveram seus pesos salvos para serem utilizados na próxima etapa.

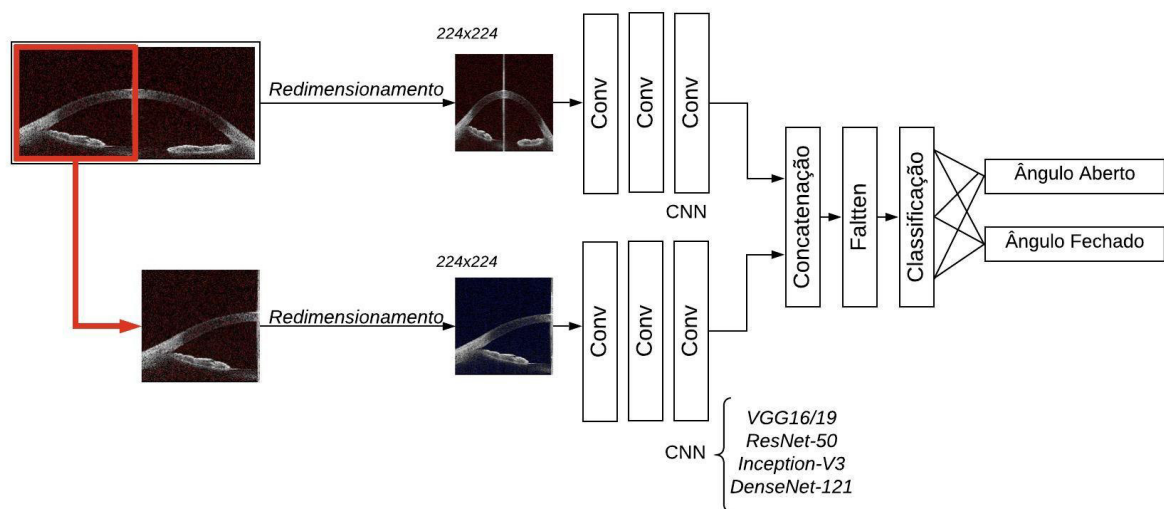
4.3 Redes Multinível

Uma vez que as imagens OCT-SA fornecem uma visualização completa da câmara anterior, possibilitando o cálculo de parâmetros clínicos referentes à região global da câmara (ex. área e comprimento da câmara anterior), e de parâmetros referentes à região onde encontra-se o ângulo formado pela junção da íris com a córnea (ex. medição do ângulo, morfologia da íris), é possível utilizar as duas regiões para classificação das imagens. Com o propósito de extrair características relevantes para classificação das imagens em diferentes níveis, um nível tendo como entrada a região global e o outro a região que contém apenas o ângulo, neste trabalho foi proposta uma arquitetura multinível, formada por dois níveis de extração de características. Os dois níveis são formados por camadas convolucionais de CNNs, sendo que cada uma das CNNs foi previamente treinada para extrair características de imagens com dois ângulos, esquerdo e direito, (primeiro nível de extração), e com apenas o ângulo esquerdo (segundo nível de extração) da câmara anterior.

Nesta etapa, os modelos obtidos após o treinamento realizado na etapa anterior são carregados sem as camadas que formam o seu classificador. O próximo passo foi formar uma arquitetura multinível, conforme mostrado na Figura 18. A arquitetura proposta possui duas camadas de entrada, seguidas das camadas que formam os modelos carregados, que realizam a extração de características. Desse modo, há dois níveis de extração, sendo que o primeiro nível gera mapas de características de imagens de dois ângulos e o segundo gera mapas de imagens com um ângulo. As características extraídas são concatenadas e alimentam uma camada *flatten*. Esta camada alimenta o classificador, formado por camadas totalmente conectadas, que tem como resultado a classificação das imagens OCT-SA como ângulo aberto ou fechado.

O próximo passo foi realizar uma nova etapa de treinamento, para ajuste dos pesos das camadas que formam o classificador. Foram testadas 5 arquiteturas multinível formadas pelo mesmo modelo de CNN em cada nível, denominados modelos Dual, por exemplo Dual VGG16, e 12 arquiteturas formadas por modelos diferentes em cada nível.

Figura 18 – Arquitetura Multinível



Acervo do autor.

4.4 Avaliação de Resultados

A validação cruzada foi escolhida como o procedimento para avaliar a generalização do modelo. O conjunto de dados foi dividido em *k-folds* (subconjuntos), sendo $k=5$. Cada *fold* contém imagens OCT-SA de 20 volumes, 16 contendo imagens da classe de ângulo aberto e 4 da classe de ângulo fechado. Durante a validação, cada *fold* é usado como conjunto de teste até que todos tenham sido testados. Os demais *folds* são combinados para formar os conjuntos de treinamento e validação. Desta forma, o conjunto de treino em cada

etapa da validação cruzada foi formado por 20% do total de volumes. Os demais volumes foram divididos em validação, formada por 16% do total, e teste, formado por 64% do total de volumes. A avaliação de desempenho dos modelos propostos para classificação das imagens foi feita utilizando-se as métricas de acurácia, precisão, especificidade, sensibilidade e valor AUC. Como no presente trabalho, a classificação do ângulo da câmara anterior é binária, optou-se por representar as imagens com ângulos fechados como classe positiva, e as imagens com ângulos abertos como negativa. Espera-se que o melhor modelo tenha uma alta sensibilidade, que no presente trabalho é a capacidade de classificar corretamente os casos de imagens de olhos com ângulo fechado. No entanto, o modelo também deve ter um alto índice de especificidade, que é a classificação correta dos casos de imagens com ângulo aberto. Uma baixa sensibilidade significa que no resultado do modelo há um número alto de imagens com ângulos fechados classificados como ângulo aberto, enquanto que uma baixa especificidade significa que há um número alto de imagens com ângulos abertos classificados como ângulos fechados. Espera-se também um alto valor AUC, visto que a base utilizada neste trabalho é desbalanceada, e que esta métrica foi reportada por todos os trabalhos relacionados, possibilitando uma comparação entre os métodos.

4.5 Considerações Finais

Este capítulo apresentou as etapas da metodologia proposta, expondo as imagens médicas que compõem o *dataset* utilizado, detalhando as etapas de construção e treinamento dos modelos utilizados para classificação das imagens, como foi realizada a avaliação da capacidade de generalização do método e como serão avaliados os resultados obtidos que serão apresentados no próximo capítulo.

5 Resultados

Este capítulo tem como objetivo apresentar os resultados obtidos pela metodologia proposta. Inicialmente, cinco CNNs foram treinadas para classificar imagens OCT-SA. Estas imagens são utilizadas para se avaliar o ângulo da câmara anterior, ACA, que é um dos principais indicadores do GPAF. As redes foram alimentadas com imagens completas e com imagens que continham apenas o ângulo esquerdo. Todas as CNNs foram inicializadas com os pesos obtidos no desafio ImageNet. O próximo passo consistiu de se utilizar uma arquitetura de 2 níveis de extração de características, sendo que diferentes combinações de modelos de CNN foram avaliados. O primeiro nível teve como entrada as imagens com dois ângulos e o segundo nível teve como entrada imagens com apenas o ângulo esquerdo.

Para validação foram consideradas as métricas de acurácia, sensibilidade, especificidade, precisão e valor AUC, para cada modelo testado. A base de dados utilizada é desbalanceada, sendo 1280 imagens pertencentes a classe 0 (ângulo aberto) e 320 imagens pertencentes a classe 1 (ângulo fechado).

A primeira etapa de experimentos constituiu-se de ajuste e treinamento de modelos de CNN. Foi utilizado, em todos os modelos, um classificador diferente do original. Os modelos foram alimentados com imagens que contém os dois e apenas um ângulo da câmara anterior (ACA). Os resultados apresentados por estes modelos, que possuem um nível de extração de características são apresentados na Tabela 3. Como foi realizada a validação cruzada, é apresentada a média dos resultados obtidos nas execuções. Foram testados os modelos VGG16, VGG19, InceptionV3, ResNet50 e DenseNet121, detalhados no capítulo [Fundamentação Teórica](#).

Tabela 3 – Resultados obtidos por modelos com um nível, alimentados com imagens com dois ou um ACA.

Modelo	Imagem	Acurácia	AUC	PRE	SEN	ESP
VGG16	2 ACA	0.942	0.875	0.940	0.762	0.987
VGG16	1 ACA	0.967	0.932	0.957	0.875	0.990
VGG19	2 ACA	0.950	0.908	0.898	0.838	0.978
VGG19	1 ACA	0.966	0.926	0.966	0.859	0.992
InceptionV3	2 ACA	0.951	0.919	0.894	0.866	0.972
InceptionV3	1 ACA	0.935	0.860	0.936	0.734	0.985
ResNet50	2 ACA	0.961	0.934	0.920	0.887	0.980
ResNet50	1 ACA	0.936	0.899	0.862	0.838	0.960
DenseNet121	2 ACA	0.984	0.977	0.956	0.966	0.988
DenseNet121	1 ACA	0.994	0.994	0.976	0.994	0.994

A segunda etapa de experimentos constituiu-se de combinar as CNNs em uma

arquitetura multinível. Inicialmente, foram testadas arquitetura com os dois níveis de extração formados pela mesma CNN, sendo a CNN do primeiro nível treinada para extrair características da imagem com 2 ACAs (Ângulo da Câmara Anterior) e a do segundo nível treinada para extrair características da imagem com apenas 1 ACA. Estes modelos foram nomeados modelos "Dual". Os resultados obtidos por estes modelos, são apresentados na Tabela 4. O próximo passo realizado foi o teste de arquiteturas formadas por CNNs diferentes em cada nível de extração. Devido aos altos resultados obtidos utilizando-se a arquitetura Dual DenseNet, a CNN DenseNet não foi incluída na etapa de testes de arquiteturas com CNNs diferentes em cada nível. Os resultados obtidos por estes modelos são apresentados na Tabela 5. Como foi realizada a validação cruzada, é apresentada a média dos resultados obtidos em cada execução.

Tabela 4 – Resultados obtidos por modelos multinível formados pela mesma CNN.

Modelo	Acurácia	AUC	PRE	SEN	ESP
Dual VGG16	0.968	0.920	1.000	0.841	1.000
Dual VGG19	0.974	0.934	1.000	0.869	1.000
Dual Inception	0.968	0.931	0.973	0.869	0.993
Dual ResNet50	0.983	0.959	0.997	0.919	0.999
Dual DenseNet	0.999	0.999	0.997	1.000	0.999

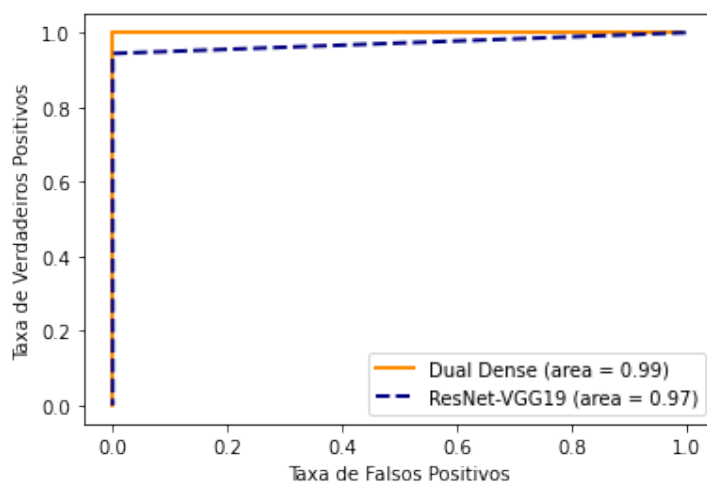
Tabela 5 – Resultados obtidos por modelos Multinível formados por CNNs diferentes.

Modelo	Acurácia	AUC	PRE	SEN	ESP
VGG19 - VGG16	0.974	0.934	1.000	0.869	1.000
VGG16 - VGG19	0.978	0.946	0.997	0.894	0.999
VGG16 - ResNet50	0.966	0.925	0.976	0.856	0.994
ResNet50 - VGG16	0.985	0.964	0.997	0.928	0.998
VGG19 - ResNet50	0.973	0.934	0.993	0.871	0.998
ResNet50 - VGG19	0.989	0.972	1.000	0.944	1.000
VGG19 - Inception	0.983	0.960	0.994	0.922	0.998
Inception - VGG19	0.979	0.960	0.971	0.928	0.991
VGG16 - Inception	0.956	0.899	0.976	0.803	0.995
Inception - VGG16	0.982	0.961	0.986	0.925	0.996
ResNet50 - Inception	0.980	0.953	0.991	0.909	0.997
Inception - ResNet50	0.978	0.954	0.980	0.912	0.995

Outra forma de análise dos resultados é a curva característica de operação do receptor, curva ROC (do inglês, *Receiver Operating Characteristic*). Através do gráfico que contém as curvas é possível verificar e comparar o desempenho dos modelos. A Figura 19 apresenta as curvas ROC com os resultados médios obtidos pelos modelos Dual DenseNet e ResNet50-VGG19. Na figura é possível visualizar o desempenho superior do modelo

Dual DenseNet, que teve uma taxa maior verdadeiros positivos, o que indica o maior sensibilidade.

Figura 19 – Curvas ROC.



Acervo do autor.

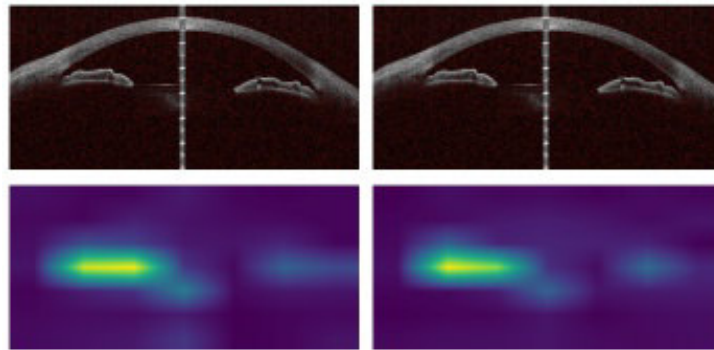
Os resultados obtidos pelo experimento com Redes Multiníveis, criados a partir da combinação de CNNs, mostram que houve um aumento no desempenho em termos de AUC e sensibilidade, quando comparados aos resultados obtidos pelos modelos de um nível. A concatenação de características nos modelos multiníveis, obtida por cada CNN, aumenta os valores de precisão, sensibilidade e valor de AUC. O modelo com melhor desempenho foi resultado da combinação de 2 CNNs DenseNet121, sendo que houve apenas um erro de classificação. No entanto, os modelos que combinam CNNs diferentes em cada nível, também alcançaram bons resultados, sendo o modelo com melhor resultado formado pela CNN ResNet50, que obteve o segundo melhor resultado na classificação de imagens com dois ACA, e o segundo nível pela CNN VGG19, que alcançou o terceiro melhor resultado na classificação de imagens com apenas 1 ACA.

5.1 Estudo de Casos

Para uma melhor análise dos resultados, foram gerados mapas de ativação de classe (Grad-CAM) (SELVARAJU et al., 2016). Essa abordagem usa os gradientes em uma rede de classificação que fluem para a camada convolucional final, para produzir um mapa de localização destacando as regiões críticas das imagens que foram usadas para classificação. A Figura 20 mostra imagens OCT-SA juntamente com seus respectivos mapas de ativação, gerados pelo modelo ResNet50.

Após análise dos mapas gerados pelos modelos VGG19 e ResNet50 treinados com imagens com 2 ACA, Figura 21 e Figura 22, foi verificado que os modelos utilizaram a

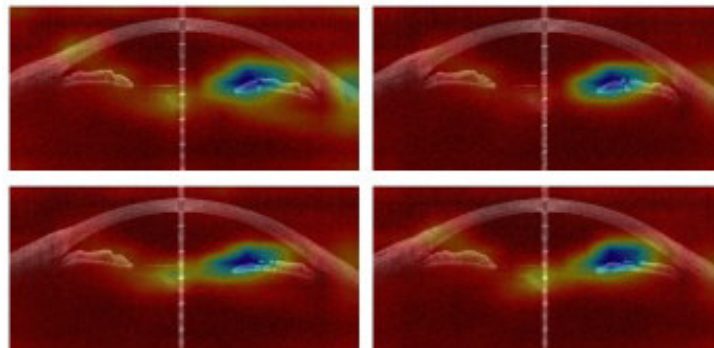
Figura 20 – Imagens OCT e Mapas de Ativação de uma classificação correta.



Acervo do autor.

região da íris para classificar as imagens. Os pacientes com o glaucoma de ângulo fechado apresentam alterações morfológicas nas estruturas da câmara anterior, sendo uma delas o aumento da espessura da íris (FU et al., 2017), (SHANG et al., 2019), (PHAM et al., 2019). Esta alteração permitiu que os modelos classicassem corretamente as imagens da classe ângulo fechado.

Figura 21 – Regiões usadas pelo modelo VGG19 para classificação das imagens (Classificação correta).

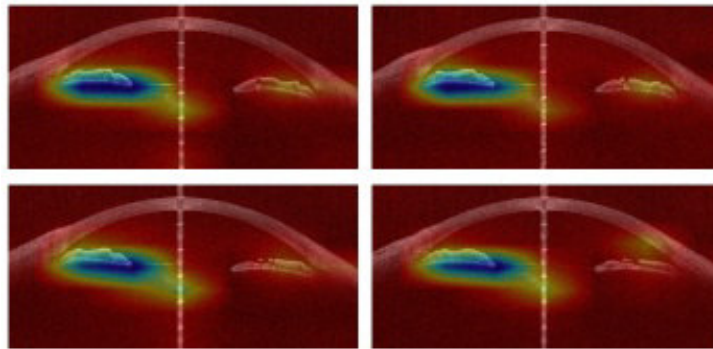


Acervo do autor.

A análise com os mapas de ativação também permitiu verificar possíveis causas para predições incorretas. As diferentes morfologias da íris, para os principais mecanismos de fechamento angular, bloqueio pupilar, íris em platô, fechamento induzido pelo cristalino e aumento do diâmetro da lente, (CHANSANGPETCH; ROJANAPONGPUN; LIN, 2018), parecem ter relação com alguns erros de predição, já que os modelos utilizaram características extraídas da região da íris para a classificação. Casos em que a espessura da íris parece não ter sofrido muita alteração, caso apresentado na Fig. 23, ou onde a extração de características não foi realizada na região da íris, levaram a uma classificação incorreta.

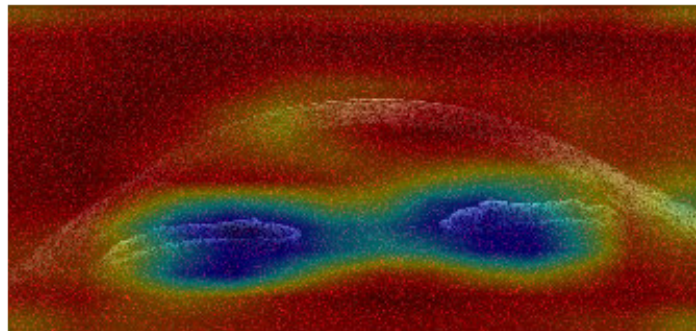
Os melhores resultados foram obtidos pelos modelos que utilizaram a CNN DenseNet como extrator de características. Mapas de ativação, Figura 24, foram utilizados para

Figura 22 – Regiões usadas pelo modelo ResNet50 para classificação das imagens (Classificação correta).



Acervo do autor.

Figura 23 – Exemplo de classificação incorreta.



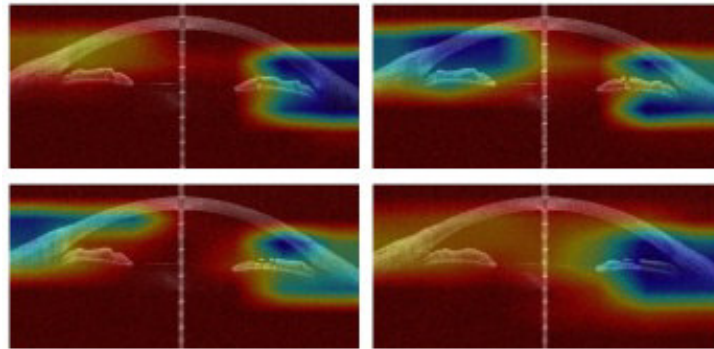
Acervo do autor.

que fosse encontrada uma possível razão do desempenho superior deste modelo em relação aos demais. Os mapas mostram que os modelos DenseNet, utilizaram não apenas a região da íris, mas a região de um ou dois ângulos da câmara anterior, formados pela junção da íris com a córnea. Pacientes diagnosticados com glaucoma primário de ângulo fechado apresentam outras alterações morfológicas, como diminuição do ângulo da câmara anterior, comprimento axial curto (FU et al., 2017), (SHANG et al., 2019), (PHAM et al., 2019), além do aumento da espessura da íris, citado anteriormente.

5.2 Discussão

A análise dos valores de especificidade e sensibilidade mostram que o *dataset* desbalanceado afetou o desempenho de praticamente todos os modelos, com exceção dos modelos que utilizaram a CNN DenseNey121, uma vez que os valores de especificidade foram significativamente mais altos. A exceção foram os modelos que utilizaram a CNN DenseNet121. Como os modelos multinível alcançaram valores maiores de sensibilidade, eles podem ser considerados como uma alternativa para lidar com o desbalanceamento

Figura 24 – Mapas de Ativação - Modelo DenseNet121 (Classificação correta).



Acervo do autor.

entre as classes de um conjunto de dados. A validação cruzada mostrou que a divisão do *dataset* entre treino e teste afeta o desempenho dos modelos, uma vez que resultados significativamente diferentes foram alcançados, quando os modelos foram treinados em conjuntos diferentes. Mais uma vez, a exceção foram os modelos que utilizaram a CNN DenseNet.

A Tabela 6 apresenta os resultados em termos de acurácia, valor AUC e sensibilidade, em comparação com trabalhos relacionados que empregaram métodos baseados em aprendizagem profunda para classificação do ângulo da câmara anterior em imagens OCT-SA. Alguns fatores importantes foram levados em consideração na comparação. Os trabalhos de Fu et al. (2018), Fu et al. (2019) e Hao et al. (2019), reportaram a acurácia balanceada em seus resultados. O método proposto por Hao et al. (2019) realiza a classificação das imagens em três classes. É importante enfatizar também que a maioria dos trabalhos relacionados utilizou *datasets* diferentes, com uma quantidade maior de imagens, em comparação com o *dataset* utilizado neste trabalho, fato que possui relevância quando é realizada uma comparação.

Tabela 6 – Comparação dos resultados obtidos por diferentes métodos.

Trabalho	DataSet	Acurácia	AUC	Sensibilidade
(FU et al., 2018)	Particular	0,892	0,945	0,888
(FU et al., 2019)	Particular	0,918	0,961	0,929
(FU et al., 2019)	Particular	0,912	0,952	0,873
(XU et al., 2019)	Particular	-	0,933	-
(FU et al., 2019a)	Particular	-	0,960	0,900
(HAO et al., 2019)	Particular	0,853	0,914	0,750
(YUAN et al., 2019)	AGE	1,000	1,000	1,000
(HAO et al., 2020)	AGE	1,000	1,000	1,000
Método Proposto	AGE	0,999	0,999	1,000

Os trabalhos de Fu et al. (2018) e Fu et al. (2019a) fazem uma comparação entre

resultados obtidos por métodos baseados em extração de características visuais de forma manual e por meio de CNNs, além do uso de técnicas de segmentação para cálculos de parâmetros clínicos quantitativos. Como esperado os métodos que utilizam aprendizagem profunda obtiveram melhores resultados. Nos métodos propostos por [Fu et al. \(2019\)](#), [Hao et al. \(2019\)](#) e [Hao et al. \(2020\)](#), é necessário capturar a região da imagem onde encontra-se o ACA. Para isso, é necessário localizar o esporão escleral, que é utilizado como centro para captura da região que alimenta a arquitetura proposta. Devido aos métodos propostos por estes trabalhos dependerem desta localização, pode considerar-se que erros nesta etapa podem inviabilizar a classificação. Para avaliação do método proposto por [Fu et al. \(2019a\)](#), o número de imagens classificadas como ângulo aberto foi limitada para balanceamento do *dataset*. Os métodos apresentados por [Yuan et al. \(2019\)](#) e [Hao et al. \(2020\)](#) alcançaram os melhores resultados. Nestes métodos, assim como em outros trabalhos, as imagens são divididas em duas para aumento dos dados de treinamento, sendo que no método proposto por [Hao et al. \(2020\)](#) é utilizada segmentação para captura da região do ângulo da câmara anterior. O método aplicado neste trabalho não teve nenhuma etapa de pré-processamento e não foi realizada segmentação de estruturas presentes nas imagens. Este trabalho avaliou arquiteturas multinível formadas por diferentes CNNs, testando diferentes combinações das mesmas, realizando uma comparação entre o desempenho dos mesmos, buscando verificar quais características foram consideradas relevantes, e utilizadas para classificação das imagens.

Uma possível limitação do método proposto neste trabalho, e possivelmente dos demais trabalhos relacionados, é que ele foi avaliado apenas em exames de imagem obtidas de indivíduos asiáticos. O *dataset* utilizado neste trabalho, AGE, e o utilizado por [Xu et al. \(2019\)](#) é composto por imagens obtidas de indivíduos chineses. As imagens utilizadas por [Fu et al. \(2019a\)](#) foram obtidas de chineses, indianos e malaio. Os demais trabalhos provavelmente utilizaram imagens obtidas de indivíduos asiáticos. Dados obtidos em estudos sugerem que a ocorrência das formas da doença está relacionada com a etnia da população [Quigley e Broman \(2006\)](#), [Tham et al. \(2014\)](#), [OMS \(2019\)](#). Sendo assim, é necessário avaliar o método em diferentes populações, entre elas a brasileira, o que é dificultado pela provável inexistência de um conjunto de dados rotulados. Além disso, o método precisa ser testado em conjuntos maiores de dados, no entanto o único conjunto disponível até o momento da escrita deste documento, é o utilizado neste trabalho. Também é necessário verificar a eficácia do método em classificar as imagens rotuladas em mais de duas classes, aberto, estreito e fechado, como foi feito nos trabalhos de [Hao et al. \(2019\)](#) e [Hao et al. \(2020\)](#).

6 Conclusão

Segundo a OMS, atualmente há cerca de 64 milhões de pessoas com glaucoma mundialmente. Desse total, 6,9 milhões de pessoas têm algum comprometimento da visão devido à doença, que poderia ter sido evitado caso a doença tivesse sido detectada em seus estágios iniciais. Estima-se ainda que do total de pessoas com glaucoma, um terço seja de glaucoma primário de ângulo fechado. Apesar de menos prevalente, o número de indivíduos com comprometimento visual devido ao glaucoma de ângulo fechado é semelhante ao de ângulo aberto. O diagnóstico precoce é essencial, porém é dificultado devido ao fato da maioria dos casos serem assintomáticos nos estágios iniciais. Este fato reforça a necessidade da busca ativa de casos da doença, através de exames periódicos de indivíduos que apresentam fatores de risco, porém muitos não possuem acesso a sistemas de saúde.

As técnicas baseadas em aprendizagem profunda para classificação de imagens, como é o caso da metodologia proposta neste trabalho, apresentam-se como uma alternativa para auxiliar os especialistas na detecção da doença, reduzindo o trabalho manual, permitindo uma otimização do tempo destes profissionais.

Neste trabalho foi proposto um método para classificação automática do ângulo da câmara anterior utilizando-se imagens OCT-SA. O objetivo do trabalho foi combinar modelos de CNN em uma arquitetura multinível, sendo cada nível utilizado para extrair características de diferentes regiões da câmara anterior, que possuem relevância para classificação do ângulo. Cada nível é formado pelas camadas convolucionais de uma CNN previamente treinada e ajustada para extrair características de imagens OCT-SA.

Os resultados mostraram que a arquitetura multinível obteve resultados superiores em relação a modelos tradicionais, sendo que o melhor resultado foi obtido combinando-se duas CNNs DenseNet121, alcançando um valor AUC de 0,999 e um valor de sensibilidade de 1,000. Diferente de outras CNNs, este modelo não foi afetado pelo desbalanceamento do *dataset* e nem pela divisão do mesmo em treino e teste. Os mapas de ativação mostraram que os modelos que utilizaram essa CNN, extraíram características da região onde encontra-se o ângulo formado pela junção da íris com a córnea, diferente de outros modelos que extraíram características apenas da região da íris, o que pode ser a razão do desempenho superior. A análise de resultados incorretos apontou duas prováveis causas de erros. A primeira é que a maioria dos modelos testados extraiu características apenas da região da íris, que pode sofrer diferentes alterações morfológicas, e a segunda é que em algumas imagens a íris não foi localizada. Como os modelos formados pela DenseNet121 não extraíram características apenas da íris, valores superiores de sensibilidade e consequente

de AUC foram alcançados.

Apesar dos resultados promissores obtidos pela metodologia proposta, é necessário avaliá-la em outros *datasets*, com um maior número de imagens, e com imagens obtidas de indivíduos de outras etnias além das asiáticas. Além disso, é necessário avaliar o método em um *dataset* que tenha imagens classificadas como ângulo aberto, estreito e fechado. Além de ser um trabalho mais desafiador, possui maior relevância, uma vez que auxiliaria os profissionais na classificação baseada na evolução da doença.

6.1 Trabalhos Futuros

Embora os resultados obtidos sejam promissores, algumas melhorias podem ser realizadas visando elevar o desempenho do método. Abaixo são listadas algumas melhorias.

- Realizar um pós processamento para obtenção do resultado final, de forma a classificar o volume de imagens, não apenas cada imagem;
- Avaliar a metodologia proposta em outros *datasets*, com um maior número de imagens, e com imagens obtidas de indivíduos pertencentes a outras etnias;
- Avaliar a metodologia proposta em *datasets* com imagens rotuladas em ângulo aberto, estreito e fechado;
- Utilizar técnicas de segmentação para captura da região do ângulo da câmara anterior;
- Testar uma abordagem que utilize aprendizagem profunda para classificação e análise da dinâmica temporal das imagens.

6.2 Produções Científicas

A Tabela 7 lista os artigos aceitos, produtos da metodologia proposta, no ano de 2020. Os trabalhos apresentados a seguir foram produzidos como autor principal. O trabalho listado recebeu uma menção honrosa.

Tabela 7 – Artigos aceitos para publicação.

Título do Artigo.	Conferência\Revista	Qualis
Multilevel CNN for Angle Closure Glaucoma Detection using AS-OCT Images	IWSSIP 2020	B1
Multilevel CNN for Anterior Chamber Angle Classification using AS-OCT Images	IJICA	B2

Referências

- ACADEMY, D. S. *Deep Learning Book*. 2019. <<http://deeplearningbook.com.br>>. [Disponível; acessado em 05-Maio-2020]. Citado na página 30.
- ANG, M.; BASKARAN, M.; WERKMEISTER, R. M.; CHUA, J.; SCHMIDL, D.; SANTOS, V. A. D.; GARHÖFER, G.; MEHTA, J. S.; SCHMETTERER, L. Anterior segment optical coherence tomography. *Progress in retinal and eye research*, Elsevier, v. 66, p. 132–156, 2018. Citado 2 vezes nas páginas 16 e 28.
- ARAÚJO, F.; CARNEIRO, A.; SILVA, R.; MEDEIROS, F.; USHIZIMA, D. Redes neurais convolucionais com tensorflow: teoria e prática. v. 1, p. 382–406, 2017. Citado 7 vezes nas páginas 27, 31, 32, 33, 34, 35 e 36.
- BISHOP, C. M. *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Berlin, Heidelberg: Springer-Verlag, 2006. ISBN 0387310738. Citado na página 32.
- BOURNE, R. R. et al. Causes of blindness and vision impairment in 2020 and trends over 30 years, and prevalence of avoidable blindness in relation to vision 2020: the right to sight: an analysis for the global burden of disease study. *The Lancet Global Health*, Elsevier, v. 9, n. 2, p. e144–e160, 2021. Citado na página 15.
- CHANSANGPETCH, S.; ROJANAPONGPUN, P.; LIN, S. C. Anterior segment imaging for angle closure. *American journal of ophthalmology*, Elsevier, v. 188, p. xvi–xxix, 2018. Citado 5 vezes nas páginas 16, 27, 28, 29 e 49.
- CHANSANGPETCH, S.; TRAN, B.; PEREZ, C. I.; SIGUAN-BELL, C.; LAU, K.; NGUYEN, A.-H.; THAI, A.; HE, M.; WANG, D.; NGUYEN, N. et al. Comparison of anterior segment optical coherence tomography parameters among vietnamese, chinese, and whites. *American Journal of Ophthalmology*, Elsevier, v. 195, p. 72–82, 2018. Citado 2 vezes nas páginas 27 e 29.
- FRICKE, T. R.; TAHHAN, N.; RESNIKOFF, S.; PAPAS, E.; BURNETT, A.; HO, S. M.; NADUVILATH, T.; NAIDOO, K. S. Global prevalence of presbyopia and vision impairment from uncorrected presbyopia: systematic review, meta-analysis, and modelling. *Ophthalmology*, Elsevier, v. 125, n. 10, p. 1492–1499, 2018. Citado na página 15.
- FU, H.; BASKARAN, M.; XU, Y.; LIN, S.; WONG, D. W. K.; LIU, J.; TUN, T. A.; MAHESH, M.; PERERA, S. A.; AUNG, T. A deep learning system for automated angle-closure detection in anterior segment optical coherence tomography images. *American journal of ophthalmology*, Elsevier, v. 203, p. 37–45, 2019a. Citado 8 vezes nas páginas 15, 16, 17, 22, 24, 28, 51 e 52.
- FU, H.; LI, F.; SUN, X.; CAO, X.; LIAO, J.; ORLANDO, J. I.; TAO, X.; LI, Y.; ZHANG, S.; TAN, M. et al. Age challenge: Angle closure glaucoma evaluation in anterior segment optical coherence tomography. *arXiv preprint arXiv:2005.02258*, 2020. Citado 3 vezes nas páginas 17, 23 e 40.

- FU, H.; XU, Y.; LIN, S.; ZHANG, X.; WONG, D. W. K.; LIU, J.; FRANGI, A. F.; BASKARAN, M.; AUNG, T. Segmentation and quantification for angle-closure glaucoma assessment in anterior segment oct. *IEEE transactions on medical imaging*, IEEE, v. 36, n. 9, p. 1930–1938, 2017. Citado 5 vezes nas páginas 16, 21, 29, 49 e 50.
- FU, H.; XU, Y.; LIN, S.; WONG, D. W. K.; MANI, B.; MAHESH, M.; AUNG, T.; LIU, J. Multi-context deep network for angle-closure glaucoma screening in anterior segment oct. In: SPRINGER. *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Granada, 2018. p. 356–363. Citado 3 vezes nas páginas 22, 24 e 51.
- FU, H.; XU, Y.; LIN, S.; WONG, D. W. K.; BASKARAN, M.; MAHESH, M.; AUNG, T.; LIU, J. Angle-closure detection in anterior segment oct based on multilevel deep network. *IEEE transactions on cybernetics*, IEEE, 2019. Citado 6 vezes nas páginas 17, 23, 24, 37, 51 e 52.
- GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. *Deep Learning*. Cambridge: MIT Press, 2016. <<http://www.deeplearningbook.org>>. Citado 4 vezes nas páginas 30, 31, 32 e 36.
- HAO, H.; FU, H.; XU, Y.; YANG, J.; LI, F.; ZHANG, X.; LIU, J.; ZHAO, Y. Open-narrow-synechiae anterior chamber angle classification in as-oct sequences. *arXiv preprint arXiv:2006.05367*, 2020. Citado 3 vezes nas páginas 24, 51 e 52.
- HAO, H.; ZHAO, Y.; FU, H.; SHANG, Q.; LI, F.; ZHANG, X.; LIU, J. Anterior chamber angles classification in anterior segment oct images via multi-scale regions convolutional neural networks. In: IEEE. *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. Berlim, 2019. p. 849–852. Citado 4 vezes nas páginas 23, 24, 51 e 52.
- HE, K.; ZHANG, X.; REN, S.; SUN, J. Deep residual learning for image recognition. In: *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Los Alamitos, CA, USA: IEEE Computer Society, 2016. p. 770–778. ISSN 1063-6919. Disponível em: <<https://doi.ieeecomputersociety.org/10.1109/CVPR.2016.90>>. Citado 2 vezes nas páginas 34 e 43.
- HO, T. K. Random decision forests. In: IEEE. *Proceedings of 3rd international conference on document analysis and recognition*. Montreal, 1995. v. 1, p. 278–282. Citado na página 36.
- HUANG, G.; LIU, Z.; WEINBERGER, K. Q. Densely connected convolutional networks. *CoRR*, abs/1608.06993, 2016. Disponível em: <<http://arxiv.org/abs/1608.06993>>. Citado 3 vezes nas páginas 35, 36 e 43.
- KUMAR, B. N.; CHAUHAN, R.; DAHIYA, N. Detection of glaucoma using image processing techniques: A review. In: IEEE. *2016 International Conference on Microelectronics, Computing and Communications (MicroCom)*. Durgapur, India, 2016. p. 1–6. Citado na página 15.
- LI, F.; ZHOU, R.; GAO, K.; JIN, L.; ZHANG, X. Volumetric parameters-based differentiation of narrow angle from open angle and classification of angle configurations: an ss-oct study. *British Journal of Ophthalmology*, BMJ Publishing Group Ltd, v. 104, n. 1, p. 92–97, 2020. Citado na página 27.

- LI, J.; CHEN, Y.; CAI, L.; DAVIDSON, I.; JI, S. Dense transformer networks. *arXiv preprint arXiv:1705.08881*, 2017. Citado na página 15.
- NI, S. N.; TIAN, J.; MARZILIANO, P.; WONG, H.-T. Anterior chamber angle shape analysis and classification of glaucoma in ss-oct images. *Journal of ophthalmology*, Hindawi, v. 2014, 2014. Citado na página 21.
- NIWAS, S. I.; LIN, W.; BAI, X.; KWOH, C. K.; KUO, C.-C. J.; SNG, C. C.; AQUINO, M. C.; CHEW, P. T. Automated anterior segment oct image analysis for angle closure glaucoma mechanisms classification. *Computer methods and programs in biomedicine*, Elsevier, v. 130, p. 65–75, 2016. Citado na página 15.
- OMS. World report on vision. World Health Organization, 2019. Citado 3 vezes nas páginas 15, 26 e 52.
- PHAM, T. H.; DEVALLA, S. K.; ANG, A.; DA, S. Z.; THIERY, A. H.; BOOTE, C.; CHENG, C.-Y.; KOH, V.; GIRARD, M. J. Deep learning algorithms to isolate and quantify the structures of the anterior segment in optical coherence tomography images. *arXiv preprint arXiv:1909.00331*, 2019. Citado 4 vezes nas páginas 15, 16, 49 e 50.
- PRIYANKA, P.; JULIET, V. N.; DEVI, S. S. Detection of glaucoma using anterior segment optical coherence tomography images. In: *Computer Aided Intervention and Diagnostics in Clinical and Medical Images*. New York City: Springer, 2019. p. 293–301. Citado na página 16.
- QUIGLEY, H. A.; BROMAN, A. T. The number of people with glaucoma worldwide in 2010 and 2020. *British journal of ophthalmology*, BMJ Publishing Group Ltd, v. 90, n. 3, p. 262–267, 2006. Citado 3 vezes nas páginas 15, 27 e 52.
- RAMAN, R.; SRINIVASAN, S.; VIRMANI, S.; SIVAPRASAD, S.; RAO, C.; RAJALAKSHMI, R. Fundus photograph-based deep learning algorithms in detecting diabetic retinopathy. *Eye*, Nature Publishing Group, v. 33, n. 1, p. 97–109, 2019. Citado na página 30.
- RUSSAKOVSKY, O.; DENG, J.; SU, H.; KRAUSE, J.; SATHEESH, S.; MA, S.; HUANG, Z.; KARPATY, A.; KHOSLA, A.; BERNSTEIN, M. et al. Imagenet large scale visual recognition challenge. *International journal of computer vision*, Springer, v. 115, n. 3, p. 211–252, 2015. Citado 2 vezes nas páginas 33 e 42.
- SAHA, S. *A Comprehensive Guide to Convolutional Neural Networks*. 2018. <<http://www.towardsdatascience.com>>. [Disponível; acessado em 20-Abril-2019]. Citado na página 31.
- SARHAN, A.; ROKNE, J.; ALHAJJ, R. Glaucoma detection using image processing techniques: A literature review. *Computerized Medical Imaging and Graphics*, Elsevier, p. 101657, 2019. Citado 2 vezes nas páginas 15 e 26.
- SBG. 3º consenso de glaucoma primário de Ângulo aberto. Sociedade Brasileira de Glaucoma, 2009. [Acessado em 12-Janeiro-2021]. Disponível em: <<https://www.sbglaucoma.org.br/wp-content/uploads/2020/06/consenso03-v2.pdf>>. Citado 2 vezes nas páginas 26 e 28.

- SBG. 2º consenso de glaucoma primário de Ângulo fechado. Sociedade Brasileira de Glaucoma, 2012. [Acessado em 12-Janeiro-2021]. Disponível em: <<https://www.sbglaucoma.org.br/wp-content/uploads/2020/06/consenso04-v2.pdf>>. Citado 4 vezes nas páginas 15, 26, 27 e 28.
- SELVARAJU, R. R.; DAS, A.; VEDANTAM, R.; COGSWELL, M.; PARIKH, D.; BATRA, D. Grad-cam: Why did you say that? visual explanations from deep networks via gradient-based localization. *CoRR*, abs/1610.02391, 2016. Disponível em: <<http://arxiv.org/abs/1610.02391>>. Citado na página 48.
- SHANG, Q.; ZHAO, Y.; CHEN, Z.; HAO, H.; LI, F.; ZHANG, X.; LIU, J. Automated iris segmentation from anterior segment oct images with occludable angles via local phase tensor. In: IEEE. *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. Berlin, Germany, 2019. p. 4745–4749. Citado 5 vezes nas páginas 16, 17, 30, 49 e 50.
- SHI, G.; JIANG, Z.; DENG, G.; LIU, G.; ZONG, Y.; JIANG, C.; CHEN, Q.; LU, Y.; SUN, X. Automatic classification of anterior chamber angle using ultrasound biomicroscopy and deep learning. *Translational vision science & technology*, The Association for Research in Vision and Ophthalmology, v. 8, n. 4, p. 25–25, 2019. Citado na página 30.
- SIMONYAN, K.; ZISSERMAN, A. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014. Citado 3 vezes nas páginas 33, 34 e 43.
- SOUSA, J. A. de; PAIVA, A. C. de; ALMEIDA, J. D. Sousa de; SILVA, A. C.; JUNIOR, G. B.; GATTASS, M. Texture based on geostatistic for glaucoma diagnosis from fundus eye image. *Multimedia Tools and Applications*, v. 76, n. 18, p. 19173–19190, Sep 2017. ISSN 1573-7721. Citado na página 15.
- SZEGEDY, C.; LIU, W.; JIA, Y.; SERMANET, P.; REED, S. E.; ANGUELOV, D.; ERHAN, D.; VANHOUCHE, V.; RABINOVICH, A. Going deeper with convolutions. *CoRR*, abs/1409.4842, 2014. Disponível em: <<http://arxiv.org/abs/1409.4842>>. Citado 2 vezes nas páginas 33 e 43.
- THAM, Y.-C.; LI, X.; WONG, T. Y.; QUIGLEY, H. A.; AUNG, T.; CHENG, C.-Y. Global prevalence of glaucoma and projections of glaucoma burden through 2040: a systematic review and meta-analysis. *Ophthalmology*, Elsevier, v. 121, n. 11, p. 2081–2090, 2014. Citado 2 vezes nas páginas 15 e 52.
- XU, B. Y.; CHIANG, M.; CHAUDHARY, S.; KULKARNI, S.; PARDESHI, A. A.; VARMA, R. Deep learning classifiers for automated detection of gonioscopic angle closure based on anterior segment oct images. *American journal of ophthalmology*, Elsevier, v. 208, p. 273–280, 2019. Citado 5 vezes nas páginas 15, 22, 24, 51 e 52.
- XU, Y.; LIU, J.; CHENG, J.; LEE, B. H.; WONG, D. W. K.; BASKARAN, M.; PERERA, S.; AUNG, T. Automated anterior chamber angle localization and glaucoma type classification in oct images. In: IEEE. *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. Osaka, Japan, 2013. p. 7380–7383. Citado na página 20.
- XU, Y.; LIU, J.; TAN, N. M.; LEE, B. H.; WONG, D. W. K.; BASKARAN, M.; PERERA, S. A.; AUNG, T. Anterior chamber angle classification using multiscale

histograms of oriented gradients for glaucoma subtype identification. In: IEEE. *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. San Diego, USA, 2012. p. 3167–3170. Citado na página [20](#).

YUAN, C.; BIAN, C.; KANG, H.; LIANG, S.; MA, K.; ZHENG, Y. Identification of primary angle-closure on as-oct images with convolutional neural networks. *arXiv preprint arXiv:1910.10414*, 2019. Citado 4 vezes nas páginas [23](#), [24](#), [51](#) e [52](#).