

**INSTITUTO FEDERAL DE EDUCAÇÃO,
CIÊNCIA E TECNOLOGIA DO CEARÁ
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE
TELECOMUNICAÇÕES**

Beatriz Silva Brasil

**Inteligência Artificial aplicada à classificação de doenças oculares em imagens
de Tomografia de Coerência Óptica**

Fortaleza, Ceará

2022

Beatriz Silva Brasil

Inteligência Artificial aplicada à classificação de doenças oculares em imagens de
Tomografia de Coerência Óptica

Dissertação apresentada ao Programa de Pós-graduação em Engenharia de Telecomunicações do Instituto Federal de Educação, Ciência e Tecnologia do Ceará como requisito parcial para a obtenção do título de mestre em Engenharia de Telecomunicações. Área de concentração: Engenharia de Telecomunicações.

Prof. Dr. Auzuir Ripardo de Alexandria

Fortaleza, Ceará

2022

Ficha catalográfica elaborada pelo SIBI/IFCE, com os dados fornecidos pelo(a) autor(a)

B823i Brasil, Beatriz Silva.
Inteligência Artificial aplicada à classificação de doenças oculares em imagens de Tomografia de Coerência Óptica / Beatriz Silva Brasil. - 2022.
67 f. : il. color.

Dissertação (Mestrado) - Instituto Federal do Ceará, Mestrado em Engenharia de Telecomunicações, Campus Fortaleza, 2022.
Orientação: Prof. Dr. Auzuir Ripardo de Alexandria.

1. Classificação. 2. Redes neurais convolucionais. 3. Tomografia de coerência óptica. I. Título.

CDD 621.382

Beatriz Silva Brasil

Inteligência Artificial aplicada à classificação de doenças oculares em imagens de Tomografia de Coerência Óptica

Dissertação apresentada ao Programa de Pós-graduação em Engenharia de Telecomunicações do Instituto Federal de Educação, Ciência e Tecnologia do Ceará como requisito parcial para a obtenção do título de mestre em Engenharia de Telecomunicações. Área de concentração: Engenharia de Telecomunicações.

Trabalho aprovado. Fortaleza, Ceará, 16 de dezembro de 2022:

**Prof. Dr. Auzuir Ripardo de
Alexandria**
Orientador - IFCE

Prof. Dr. Daniel Xavier Gouveia
IFCE

**Prof. Dr. Antônio Augusto Teixeira
Peixoto**
IFCE

Prof. Dr. Daniel da Rocha Lucena
Escola Cearense de Oftalmologia

Fortaleza, Ceará
2022

*Este trabalho é dedicado às crianças adultas que,
quando pequenas, sonharam em se tornar cientistas.*

Agradecimentos

Agradeço ao Instituto Federal do Ceará e ao Programa de Pós Graduação em Engenharia de Telecomunicações pela possibilidade de cursar o mestrado em uma escola de renome no estado do Ceará.

Agradeço a minha família pelo apoio e incentivo, mesmo com todas as dificuldades encontradas e o crescente desânimo com as escassas oportunidades.

Agradeço ao professor Dr. Auzuir Ripardo de Alexandria pela paciência de acompanhar seus orientandos por todo o período de pandemia.

Agradeço a minha irmã Juliana pela disponibilidade de ler este trabalho e sugerir melhorias, pela disponibilidade de ajudar a debugar os códigos utilizados e pelo prazer de dividir um planeta e uma época.

Agradeço a Lyndainês Araújo, colega de graduação e mestrado, pelas inúmeras ajudas e dicas de solução de problemas na programação.

Agradeço também a todos os outros colegas de mestrado a quem, em algum momento, eu pedi ajuda e prontamente fui auxiliada.

“Somos feitos da matéria de que são feitos os sonhos.”
(William Shakespeare, A Tempestade)

Resumo

Dada a importância que a inteligência artificial adquiriu na atualidade e as suas possíveis aplicações na área da medicina, esta dissertação tem o objetivo de classificar a existência das doenças oculares CNV (neovascularização coroidal), DME (edema macular diabético) e DRUSEN (acúmulos lipídicos na retina) ou não (imagens normais) usando um banco de imagens digitais oriundas de exames de Tomografia de Coerência Óptica (OCT - *Optical Coherence Tomography*). Foram empregados neste estudo dois algoritmos de *Deep Learning* com a utilização de duas arquiteturas de Redes Neurais Convolucionais (RNC), uma rede baseada na Cifar-10 e a rede Resnet50. Nas Redes Neurais a extração de características é realizada como uma etapa do próprio algoritmo, que trabalha diretamente sobre o banco de imagens. Foi utilizado um banco de 800 imagens, 200 para cada quadro clínico, oriundas de um banco de imagens disponibilizado na internet. Os valores obtidos para acurácia foram 83,09% para a primeira rede neural apresentada, uma rede sequencial, e 97,78% para a rede Resnet50. Os resultados obtidos condizem com os resultados apresentados nos trabalhos pesquisados que realizam classificação. Esse estudo contribui com o avanço no emprego de inteligência artificial no processo de auxílio ao diagnóstico médico.

Palavras-chaves: Classificação, Redes neurais convolucionais, Tomografia de coerência óptica.

Abstract

Given the importance that artificial intelligence has acquired today and its possible applications in the field of medicine, this dissertation aims to classify the existence of eye diseases CNV (choroidal neovascularization), DME (diabetic macular edema) and DRUSEN (lipid accumulation in the retina) or not (normal images) using a bank of digital images from Optical Coherence Tomography (OCT) exams. Two Deep Learning algorithms were used in this study using two Convolutional Neural Networks (CNN) architectures, a network based on Cifar-10 and the Resnet50 network. In Neural Networks, feature extraction is performed as a step of the algorithm itself, which works directly on the image bank. A bank of 800 images was used, 200 for each clinical condition, from an image bank available on the internet. The results obtained for accuracy were 83.09% for the first neural network presented, a sequential network, and 97.78% for the Resnet50 network. The results obtained are consistent with the results presented in the researched works that perform classification. This study contributes to the advancement in the use of artificial intelligence in the process of aiding medical diagnosis.

Key-words: Classification, Convolutional Neural Network, Optical Coherence Tomography.

Lista de ilustrações

Figura 1 – Fluxograma da estrutura do trabalho capítulo a capítulo.	21
Figura 2 – Exemplo de exame de OCT revelando a existência de drusas no olho do paciente.	25
Figura 3 – Exemplo de exame de OCT revelando a existência de neovascularização coroidal no olho do paciente.	26
Figura 4 – Exemplo de exame de OCT revelando a existência de edema macular diabético no olho do paciente.	27
Figura 5 – Assimetrias negativa e positiva.	28
Figura 6 – Tipos de curtose.	28
Figura 7 – Instâncias de cada método computacional no conjunto de técnicas de Inteligência Artificial.	30
Figura 8 – Ilustração da estrutura de uma rede neural artificial. Para redes mais complexas de deep learning haverá um número bem maior de camadas de neurônios ocultas.	32
Figura 9 – Ilustração da estrutura de um neurônio artificial. Na figura, as entradas são multiplicadas por seus respectivos pesos e somadas. Essa soma é avaliada com base no valor limite, ou viés, estabelecido para o neurônio para que a saída atribuída.	32
Figura 10 – Minimização da função de custo.	34
Figura 11 – Esquematização das camadas que compõem uma Rede Neural Convolutacional. Há na ilustração duas regiões distintas: A região de extração de características, delimitada em vermelho, e a região de classificação, destacada em azul. Na camada de classificação estão apenas as camadas totalmente conectadas, enquanto na camada de extração de características pode-se identificar as camadas de convolução e de <i>pooling</i>	36

Figura 12 – Formatação de uma matriz de confusão 2x2. Na figura está exemplificada uma matriz de confusão para uma classificação em duas categorias: uma denominada "Sim" e outra denominada "Não". Essas exemplos de classes podem ser interpretados como ter ou não ter determinada doença. No eixo vertical temos os valores reais, que são os rótulos dos dados de entrada no algoritmo. No eixo horizontal temos os valores preditos, que serão calculados a partir dos pesos obtidos com o treinamento. Comparando os valores desses dois eixos tem-se: os Verdadeiros Positivos, que são os valores cuja predição concorda com o valor real, ou seja, a pessoa tem a doença assim como é afirmado pelo rótulo; os Verdadeiros negativos, que também significam uma concordância dos valores reais e preditos, mas dessa vez em favor da segunda classe considerada, a pessoa não tem patologia e o rótulo é confirmado pela inteligência artificial.	41
Figura 13 – Formatação de uma matriz de confusão 4X4. Abaixo da matriz é possível observar as combinações que compõem os Verdadeiros Positivos, Verdadeiros Negativos, Falsos Positivos e Falsos Negativos para cada uma das 4 classes.	42
Figura 14 – Normal.	43
Figura 15 – Edema macular diabético.	43
Figura 16 – Neo-vascularização coroidal.	44
Figura 17 – Drusas.	44
Figura 18 – Mapa de Ideias da metodologia utilizada neste trabalho.	45
Figura 19 – Esquema da rede sequencial apresentada no artigo Hu et al. (2019). Neste diagrama tem-se um patch de 33x33 de uma imagem sendo submetido a três ciclos de camadas de Convolução seguidas de <i>Pooling</i> . Logo em seguida há duas camadas sequenciais totalmente conectadas e por último o classificador <i>softmax</i> .	45
Figura 20 – Formatação genérica do bloco residual, também chamado conexão de salto, característica da rede Resnet50.	47
Figura 21 – Conforme os dados sejam discretos, pode-se plotar os histogramas de frequências para os atributos máximo, mínimo, média, variância, assimetria e curtose oriundos da extração de características das imagens. Nesta figura é possível observar os histogramas para mínimo e variância.	52
Figura 22 – Conforme os dados sejam discretos, pode-se plotar os histogramas de frequências para os atributos máximo, mínimo, média, variância, assimetria e curtose oriundos da extração de características das imagens. Nesta figura é possível observar os histogramas para máximo e assimetria.	53

Figura 23 – Conforme os dados sejam discretos, pode-se plotar os histogramas de frequências para os atributos máximo, mínimo, média, variância, assimetria e curtose oriundos da extração de características das imagens. Nesta figura é possível observar os histogramas para média e curtose.	54
Figura 24 – Gráficos de densidade para os atributos máximo, mínimo, média, variância, assimetria e curtose calculados sobre os histogramas de extração de características das imagens. Nesta figura é possível observar os gráficos de densidade para mínimo e variância.	55
Figura 25 – Gráficos de densidade para os atributos máximo, mínimo, média, variância, assimetria e curtose calculados sobre os histogramas de extração de características das imagens. Nesta figura é possível observar os gráficos de densidade para máximo e assimetria.	56
Figura 26 – Gráficos de densidade para os atributos máximo, mínimo, média, variância, assimetria e curtose calculados sobre os histogramas de extração de características das imagens. Nesta figura é possível observar os gráficos de densidade para média e curtose.	57
Figura 27 – Evolução de acurácias de treinamento e validação, em 10 treinamentos, para a CNN sequencial e a CNN residual numa divisão do banco de 60%/40%.	58
Figura 28 – Evolução de acurácias de treinamento e validação, em 10 treinamentos, para a CNN sequencial e a CNN residual numa divisão do banco de 70%/30%.	59
Figura 29 – Evolução de acurácias de treinamento e validação, em 10 treinamentos, para a CNN sequencial e a CNN residual numa divisão do banco de 80%/20%.	60

Lista de tabelas

Tabela 1 – Adequação dos métodos à classificação binária e à classificação multiclasse	31
Tabela 2 – Matriz de Confusão da CNN sequencial	48
Tabela 3 – Da matriz de confusão anterior, tem-se:	49
Tabela 4 – Acurácia, precisão e <i>recall</i> :	49
Tabela 5 – Matriz de Confusão da rede Resnet50	50
Tabela 6 – Da matriz de confusão anterior, tem-se:	50
Tabela 7 – Acurácia, precisão e <i>recall</i> :	50
Tabela 8 – Tabela resumo de valores de acurácia para as três partições do banco de dados	60
Tabela 9 – Tabela geral de valores de acurácia de classificação para os artigos pesquisados e para este trabalho	61

Lista de abreviaturas e siglas

IFCE	Instituto Federal de Educação, Ciência e Tecnologia do Ceará
PPGET	Programa de Pós graduação em Engenharia de Telecomunicações
IA	Inteligência Artificial
CNN	<i>Convolutional Neural Network</i>
RNA	Rede Neural Artificial
RNC	Rede Neural Convolucional
OCT	<i>Optical Coherence Tomography</i>
DMRI	Degeneração Macular relacionada à idade
CNV	<i>Choroidal Neovascularization</i>
DME	<i>Diabetic Macular Edema</i>
DRUSEN	Drusas
AMD	<i>Age-related Macular Degeneration</i>
ROC	<i>Receiving Operating Curve</i>
RGB	<i>Red-Green-Blue color system</i>
HSV	<i>Hue-Saturation-Value color system</i>
LBP	<i>Local Binary Pattern</i>
EMD	Modo de Decomposição Empírica
SVM	<i>Support Vector Machine</i>
KNN	<i>K-nearest Neighbours</i>
LDA	<i>Linear Discriminant Analysis</i>
GLM	<i>Generalized Linear Model</i>
SGD	<i>Stochastic Gradient Descent</i>
RF	<i>Random Forest</i>
DT	<i>Decision Tree</i>

LG	<i>Logistic Regression</i>
MLP	<i>Multilayer Perceptron</i>
VPP	Valor Preditivo Positivo
VP	Verdadeiro Positivo
VN	Verdadeiro Negativo
FP	Falso Positivo
FN	Falso Negativo
MAE	<i>Mean Absolute Error</i>
MSE	<i>Mean Squared Error</i>
RMSE	<i>Root Mean Squared Error</i>

Sumário

1	INTRODUÇÃO	17
1.1	Justificativa	18
1.2	Objetivo geral	19
1.2.1	Objetivos específicos	19
1.3	Limites do trabalho	19
1.4	Limitações do trabalho	19
1.5	Estrutura do trabalho	20
2	FUNDAMENTOS TEÓRICOS	22
2.1	Revisão bibliográfica	22
2.2	Tomografia de Coerência Óptica	24
2.3	Patologias oculares	24
2.3.1	Degeneração Macular relacionada à idade	24
2.3.1.1	Degeneração Macular relacionada à idade (seca)	25
2.3.1.2	Degeneração Macular relacionada à idade (úmida)	25
2.3.2	Edema Macular Diabético	26
2.4	Processamento de Imagens	26
2.4.1	Sistema de cores RGB	26
2.4.2	Sistema de cores HSV	27
2.4.3	Extração de Características	27
2.5	Inteligência Artificial	29
2.5.1	Aprendizado de Máquina (<i>Machine Learning</i>) e Aprendizado Profundo (<i>Deep Learning</i>)	30
2.5.2	Regressão e Classificação	30
2.5.3	Classificação binária X Classificação multiclasse	31
2.5.4	Redes Neurais Artificiais (RNA)	31
2.5.4.1	Redes Neurais Convolucionais (RNC)	33
2.5.5	Métricas de avaliação	38
2.5.5.1	Acurácia	38
2.5.5.2	Precisão	38
2.5.5.3	<i>Recall</i>	38
2.5.5.4	Erro médio absoluto	39
2.5.5.5	Erro quadrático médio	39
2.5.5.6	Raiz do erro médio quadrático	39
2.5.6	Matriz de confusão	39

2.6	A linguagem Python	40
3	METODOLOGIA	43
3.1	Mapa Resumo	43
3.2	Coleta de dados	43
3.3	Forma de análise dos dados	44
3.4	Forma de análise dos resultados	47
4	RESULTADOS E DISCUSSÕES	48
4.1	Resultados obtidos com a aplicação da CNN sequencial	48
4.2	Resultados obtidos com a aplicação da rede Resnet50	50
4.3	Discussão	51
4.3.1	Gráfico comparativo de acurácias de treino e validação para 3 partições do banco	56
4.3.1.1	Divisão 60% treino/40% validação	57
4.3.1.2	Divisão 70% treino/30% validação	58
4.3.1.3	Divisão 80% treino/20% validação	59
4.3.2	Tabela geral de acurácias	59
5	CONCLUSÕES	62
5.1	Trabalhos futuros	62
	REFERÊNCIAS	63

1 Introdução

Não é recente o esforço de introduzir a inteligência artificial (IA) no diagnóstico médico. O sistema especialista MYCIN foi o pioneiro e abriu um precedente sobre a aplicação da inteligência artificial na medicina como também gerou expectativas além do normal sobre o seu alcance. Conforme [Buchanan e Shortliffe \(1984\)](#), o sistema especialista utilizava-se de regras, que consistiam em perguntas feitas ao paciente cuja resposta seria sim ou não, e baseado nas respostas era elaborado um diagnóstico.

Os algoritmos de IA, por vezes são alimentados com dados oriundos de exames de imagem. Uma das técnicas empregadas com esse fim é a de imageamento por Tomografia de Coerência Óptica (OCT), onde padrões de interferência são usados para gerar imagens transversais dos tecidos com alta resolução ([AL-ASWAD, 2019](#)). Através deste exame é possível detectar patologias nos olhos, seja por decorrência de doença ou simplesmente devido à passagem dos anos, sendo alguns exemplos a degeneração macular relacionada à idade, seca e úmida, o edema macular diabético, o glaucoma, entre outros. A imagem de OCT bidimensional é gerada em tons de cinza cujas nuances revelam as diferenças fisiológicas das camadas internas da retina.

Os métodos de classificação de imagens de exames de OCT de olhos tem ganhado atenção e já há diversos estudos realizados com detecção de patologias como o glaucoma em [Burgansky-Eliash et al. \(2005\)](#), DMRI (Degeneração macular relacionada à idade) em [Moraes et al. \(2021\)](#), entre outros. Diversas técnicas são apresentadas na literatura com o intuito de classificar imagens de exames de OCT. Alguns artigos utilizam as técnicas de *machine learning* para classificar as imagens como SVM (Máquina de vetor de suporte) em [Thomas et al. \(2021b\)](#), [Burgansky-Eliash et al. \(2005\)](#) e *Random Forest* em [Asaoka et al. \(2017\)](#), [An et al. \(2019\)](#). Outros utilizam o *deep learning* e empregam CNN's nessa tarefa como [Das, Dandapat e Bora \(2019\)](#), [Thomas et al. \(2021a\)](#), [Kuwayama et al. \(2019\)](#).

O objetivo da realização de estudos nessa área é contribuir para o desenvolvimento da aplicação da inteligência artificial na medicina para servir no auxílio ao diagnóstico, inclusive facilitando a detecção de patologias em estágios iniciais. Até o momento, os métodos de análise de imagens com resultados mais confiáveis são aqueles derivados da aplicação de Inteligência Artificial, segundo [Hormel et al. \(2021\)](#).

Há, entretanto, diversas limitações nessa aliança entre IA e diagnóstico médico, por exemplo a falta da abordagem holística, do contato com o paciente, a desacreditação de profissionais em decorrência da confiança na automação, a necessidade de um largo banco de dados para treinamento e da incorporação pela IA da variabilidade e a ambiguidade naturais à clínica médica, defende [Al-Aswad \(2019\)](#).

Dado o poder das redes de aprendizado profundo, esse ramo da inteligência artificial tem sido explorado e estudado por grandes companhias de tecnologia. O *Deep mind* da Google utiliza redes de aprendizado por reforço e o método de tentativa e erro, composição do algoritmo AlphaGo que derrotou o campeão mundial de xadrez Lee Sedol em 2016, conforme Kapoor, Walters e Al-Aswad (2019).

Dentre os artigos pesquisados, há alguns cujo estudo apresentado é mais próximos do objetivo deste trabalho, que é a classificação de imagens de OCT. Neles, inclusive, é utilizado o mesmo banco de exames de imagem, são eles: P et al. (2021), Tsuji et al. (2020), Das, Dandapat e Bora (2019).

Uma classificação de quatro categorias de imagens oriundas de OCT é proposta em P et al. (2021). O banco de imagens utilizado é o apresentado em Kermany et al. (2018) e a classificação se dá entre imagens de CNV (Neovascularização Coroidal), DME (Edema Macular Diabético), DRUSEN (Drusas) e Normais. Uma CNN (Rede Neural Convolucional) foi empregada alcançando 99,69% de acurácia com apenas 3 erros de classificação dentre 968 casos de testes. Apenas 6,9% dos parâmetros apreensíveis foram utilizados, tornando o algoritmo mais simples e com potencial para aplicações que apresentem resultados em tempo real.

Em Tsuji et al. (2020) é utilizada uma rede neural "cápsula", denominação utilizada no artigo devido ao reduzido número de camadas da rede. Essa rede é constituída de apenas duas camadas convolucionais e uma completamente conectada, para classificar o banco de imagens de tomografia de olhos apresentadas em Kermany et al. (2018). As imagens são classificadas em CNV, DME, DRUSEN e NORMAL com uma acurácia de 99,6%.

Em Das, Dandapat e Bora (2019) também é utilizado o banco de dados apresentado por Kermany et al. (2018) e a classificação em CNV, DME, DRUSEN e imagens normais é executada com uma acurácia de 99,6%.

1.1 Justificativa

Desde a descoberta dos raios X, em 1858, passando pelo advento da tomografia e da ressonância magnética que os exames de imagem têm se desenvolvido e se consolidado como meios de diagnóstico médico não invasivo. O atual momento, aliás, é de desenvolvimento da telemedicina e dos aplicativos. Com isso desenvolvem-se também os algoritmos, os bancos de dados e a análise desses dados coletados. Com o advento da automação dos processos, o emprego da tecnologia na medicina é inevitável e contribui tanto para assegurar a veracidade do diagnóstico como para acelerar a análise dos exames de imagem.

1.2 Objetivo geral

Classificar a presença das doenças oculares CNV (neovascularização coroidal), DME (edema macular diabético) e DRUSEN (acúmulos lipídicos na retina) ou não (imagens normais) em um banco de imagens digitais oriundas de exames de Tomografia de Coerência Óptica (OCT - Optical Coherence Tomography).

1.2.1 Objetivos específicos

Os objetivos específicos deste trabalho são:

- Apresentar os resultados da classificação realizada com as duas redes para discutir acerca da matriz de confusão analisando os valores de verdadeiro positivo, para avaliar o acerto da classificação, e os valores de falso negativo e falso positivo, para avaliar a necessidade de um desempenho melhor do algoritmo empregado na tarefa da automatização de um diagnóstico;
- Comparar os resultados obtidos com as CNN's utilizadas na classificação para identificar aquela que apresente melhor desempenho.
- Explicar, com base nas características dos dados utilizados, a adequação do método selecionado.

1.3 Limites do trabalho

Este trabalho está voltado à tarefa de classificação. Imagens de quatro categorias cada uma com condição clínica específica da retina alimentarão dois algoritmos. Serão usadas para isto duas redes neurais convolucionais: uma rede sequencial e a rede Resnet50. Não é do escopo deste trabalho realizar segmentação de imagens.

1.4 Limitações do trabalho

A primeira dificuldade encontrada para a realização do trabalho foi a procura por um banco de imagens na internet que tivesse um tamanho considerável e que apresentasse mais categorias para classificação. Posteriormente, na etapa de execução dos códigos e testes para avaliar resultados parciais com a aplicação dos métodos empregados, a capacidade de processamento do hardware limitou a utilização do banco de dados em sua totalidade. O hardware usado foi um notebook Intel Core i5 com 8Gb de memória RAM.

1.5 Estrutura do trabalho

O Capítulo 2 esclarece os conceitos principais, explica a natureza de cada patologia e introduz a aplicação da inteligência artificial na medicina, mais especificamente no auxílio ao diagnóstico médico, fundamentando o conhecimento para o entendimento deste trabalho.

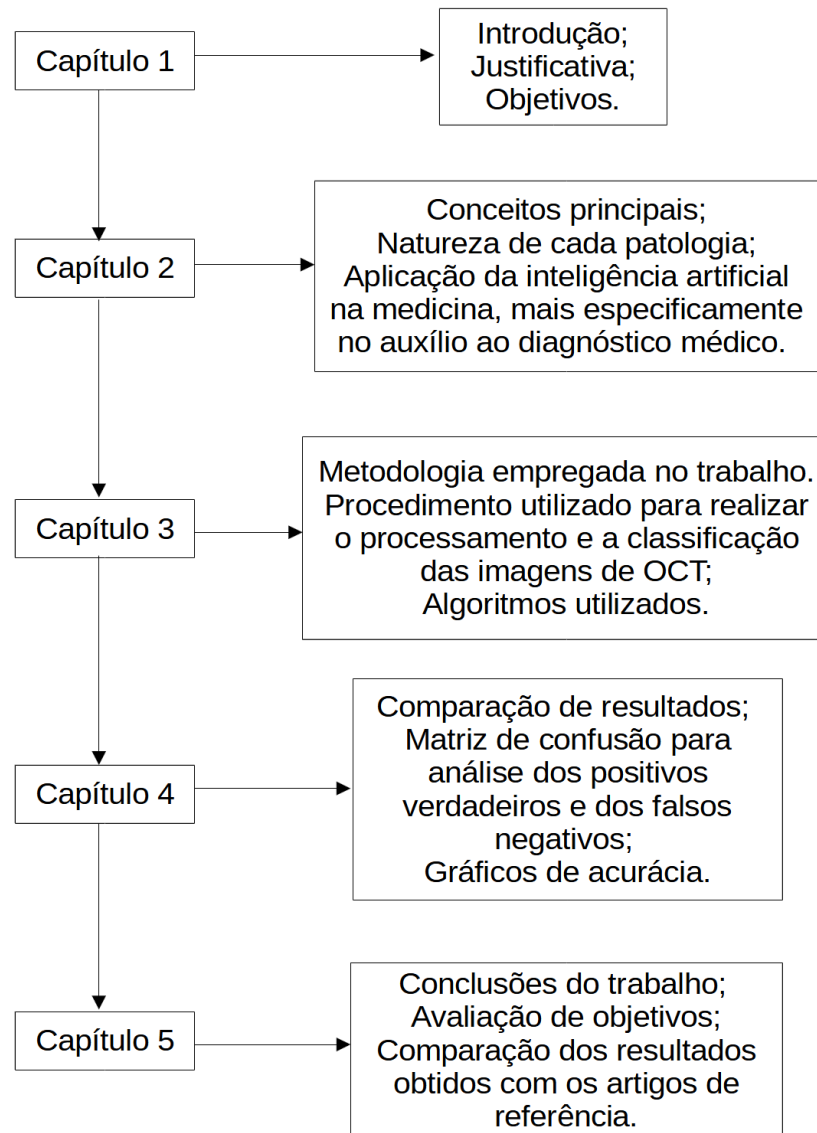
No Capítulo 3 é apresentada a metodologia empregada no trabalho. Nele é explicado o procedimento utilizado para realizar o processamento e a classificação das imagens de Tomografia de Coerência Óptica. Também são apresentados os algoritmos que foram utilizados e como ocorrerá, no capítulo seguinte, a análise dos resultados.

No Capítulo 4 são exibidos e discutidos os resultados das classificações realizadas e os resultados obtidos são comparados com os apresentados na literatura de referência. São apresentados como resultados as respectivas acurácias obtidas com cada classificação, a matriz de confusão para análise dos positivos verdadeiros e dos falsos negativos, e os gráficos de acurácia.

O Capítulo 5 apresenta as conclusões do trabalho e avalia se os objetivos listados na introdução foram alcançados por meio de comparação dos resultados obtidos com aqueles apresentados nos artigos que fundamentam este trabalho.

Na figura 1 está esquematizada a estrutura do trabalho em capítulos.

Figura 1 – Fluxograma da estrutura do trabalho capítulo a capítulo.



Fonte: Produzido pelo autor.

2 Fundamentos Teóricos

Apresentam-se neste capítulo os principais conceitos que dão suporte à metodologia desenvolvida nesta pesquisa. Este capítulo inicia-se com uma revisão bibliográfica dos artigos pesquisados sobre o assunto e que foram referência para este trabalho. O capítulo segue com esclarecimentos acerca do conceito de OCT, passando por uma breve explicação sobre as patologias oculares até os conceitos de aprendizado profundo (*Deep Learning*), de redes neurais artificiais (ANN) e convolucionais (CNN) e classificação de imagens. As patologias oculares abordadas neste trabalho são: Neovascularização coroidal, Edema Macular Diabético e Drusas.

2.1 Revisão bibliográfica

Uma classificação de quatro categorias de imagens oriundas de OCT é proposta em [P et al. \(2021\)](#). O banco de imagens utilizado é o apresentado em [Kermany et al. \(2018\)](#) e a classificação é executada entre imagens de CNV, DME, DRUSEN e Normais. Uma CNN foi empregada alcançando 99,69% de acurácia apresentando apenas 3 erros de classificação dentre 968 casos de testes. Apenas 6,9% dos parâmetros apreensíveis foram utilizados, tornando o algoritmo mais simples e com potencial para aplicações que apresentem resultados em tempo real.

[Lee, Baughman e Lee \(2017\)](#) utilizam-se de um banco de dados de 52.690 imagens de olhos saudáveis e 48.312 imagens de olhos com AMD (degeneração macular relacionada à idade), material originário de uma coletânea de aproximadamente 2,6 milhões de imagens de exames de tomografia de coerência óptica. Nessa classificação binária realizada obteve-se uma acurácia de 87,63% e área sob a curva ROC (*Receiver Operating Characteristic*) de 92,78%.

Em [Tsuji et al. \(2020\)](#) é utilizada uma rede neural "cápsula", reduzida no número de camadas, com apenas duas camadas convolucionais e uma completamente conectada, para classificar o banco de imagens de OCT de fundo de olho apresentadas em [Kermany et al. \(2018\)](#). As imagens são classificadas em CNV, DME, DRUSEN e NORMAL com uma acurácia de 99,6%.

Em [Kuwayama et al. \(2019\)](#) 1.100 imagens contendo exames de olhos normais e de 22 patologias, aumentadas para 59.400 através de operações como rotação e translação, são submetidas a uma CNN para identificar a doença. Ranqueando as doenças, aquela escolhida como a primeira opção alcança 83% de acerto no diagnóstico pela rede, enquanto a segunda doença alcança 7%.

Em [Burgansky-Eliash et al. \(2005\)](#) é usado um banco de imagens com três categorias: Olhos normais, olhos com glaucoma precoce e glaucoma avançado. Foi aplicado *machine learning* e os classificadores foram treinados com 38 parâmetros, dos quais apenas 8 foram usados no teste. Os classificadores utilizados foram análise discriminante linear, máquina de vetores de suporte, particionamento recursivo e árvore de regressão, modelo linear generalizado e modelo aditivo generalizado.

Em [Farsiu et al. \(2014\)](#) é usado um banco com imagens de olhos normais (115 imagens) e imagens de olhos com degeneração macular relacionada à idade em estágio intermediário (269 imagens). Medidas tomadas nas imagens que dizem respeito a espessuras e distâncias entre estruturas foram as entradas dos classificadores. Foi empregado o método *Leave-one-out* para melhorar a performance.

Em [Mookiah et al. \(2015\)](#) imagens de exames de OCT são classificadas para diagnóstico normal ou para a existência de degeneração macular. Nesse trabalho é feita uma extração de características de forma não linear pela técnica de EMD (Modo de decomposição empírica) para posterior classificação por meio do algoritmo de SVM. A validação cruzada é feita em 10 camadas e os resultados de acurácia, sensibilidade e especificidade alcançam 100%.

Em [Thomas et al. \(2021b\)](#) é proposta uma nova arquitetura de rede neural convolucional, multiescala e multipercurso e com seis camadas convolucionais, para detectar exames de olhos que apresentem estágios precoces de degeneração macular relacionada à idade. A avaliação de desempenho é feita por validação cruzada em 10 camadas utilizando vários classificadores como Máquina de Vetor de Suporte, *Perceptron* Multicamadas e *Random Forest*. Com este último obteve-se o melhor resultado: 96,66% com o primeiro conjunto de dados, 98,97% com o segundo, 99,74% com o terceiro e 99,78% com o quarto.

Em [Das, Dandapat e Bora \(2019\)](#), é utilizado o banco de dados de imagens de OCT apresentado por [Kermany et al. \(2018\)](#) e a classificação em CNV, DME, DRUSEN e imagens normais é executada com uma acurácia de 99,6% por meio da aplicação de Rede Neural Convolucional. Dos artigos pesquisados, a maioria, [P et al. \(2021\)](#), [Tsuji et al. \(2020\)](#), [Kuwayama et al. \(2019\)](#), [Das, Dandapat e Bora \(2019\)](#), utiliza como ferramenta de processamento e classificação do banco de dados Redes Neurais Convolucionais. [Thomas et al. \(2021b\)](#) inclusive aliam a atuação da Rede Neural a de um classificador de *Machine Learning*, sendo estes pouco usados isoladamente como em [Burgansky-Eliash et al. \(2005\)](#), [Farsiu et al. \(2014\)](#), sendo executada uma classificação apenas binária em [Lee, Baughman e Lee \(2017\)](#).

2.2 Tomografia de Coerência Óptica

Dentre os meios de realização de exames de imageamento óptico estão o Raio X Computadorizado, a Ressonância Magnética, o Imageamento Ultrassônico e a Tomografia de Coerência Óptica (OCT). Cada uma mede uma característica física do meio em questão e tem determinados alcance e penetração nos tecidos, sendo cada uma adequada para situações específicas. A técnica de OCT é o enfoque deste trabalho. A Tomografia de coerência óptica é um método de imageamento transeccional não invasiva de sistemas biológicos, imprimindo em um mapa 2D a refletividade óptica dos tecidos. É utilizada na realização de exames oculares de imagem para a detecção de anomalias, conforme [Lin \(1991\)](#). Segundo [Motta, Coblenz e Melo \(2009\)](#), essa modalidade de Tomografia utiliza luz com comprimento de onda próximo ao infravermelho para produzir imagens da retina em sentido sagital, que divide o corpo lateralmente em esquerda e direita. O exame permite obter imagem que se assemelha a um corte histológico, sendo possível mensurar a espessura da retina e visualizar suas camadas.

2.3 Patologias oculares

Ao longo da vida, o ser humano está sujeito a ter reduções de sua capacidade visual seja como consequência de alguma doença, seja por decorrência da idade. A patologia mais conhecida é a redução da capacidade de leitura, que pode ser decorrente de uma doença ou apenas do envelhecimento natural das células do olho. Esse envelhecimento natural atinge as células de uma região da retina chamada mácula, que tem a função de percepção dos detalhes na visão. Como consequência a visão pode apresentar-se embaçada, com manchas no centro, dificuldade de focalização de objetos, sensibilidade a luzes brilhantes e dificuldades de enxergar de perto ([NEOOFTALMO, 2020](#)).

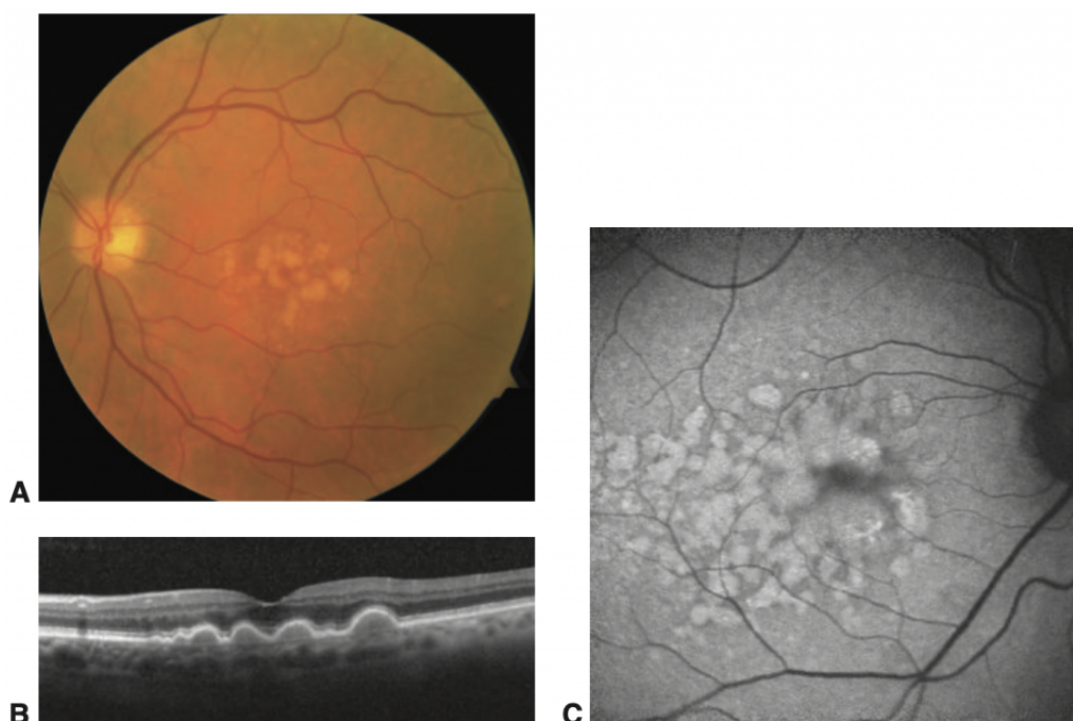
2.3.1 Degeneração Macular relacionada à idade

A Degeneração Macular relacionada à idade (DMRI) é uma doença degenerativa que atinge a parte central da retina, a mácula. Essa região é responsável pela percepção dos detalhes na visão. Uma das causas de seu aparecimento é o envelhecimento natural das células que ocorre com a idade, causando sintomas como o embaçamento da visão no centro, dificuldade de distinguir cores, dificuldade em focalizar objetos, aumento da sensibilidade a luzes brilhantes e dificuldade de enxergar de perto. O diagnóstico médico irá identificar qual o tipo de DMRI, se a DMRI é seca (atrófica) ou úmida (exsudativa) ([NEOOFTALMO, 2020](#)).

2.3.1.1 Degeneração Macular relacionada à idade (seca)

É a mais comum, ocorrendo em 90% dos casos. Configura-se como a perda progressiva e lenta das células da retina, causando perda de visão para leitura. Também diagnosticada pela presença de Drusas, do inglês *Drusen*, drusas são depósitos amarelos de lipídeos e proteínas sob a retina, conforme ilustrado na figura 2. A relação entre drusas e a degeneração macular com a idade não é clara, geralmente não apresenta sintomas, entretanto conforme [Porter \(2022\)](#), a presença de drusas em quantidade pode ser um sinal precoce de DMRI.

Figura 2 – Exemplo de exame de OCT revelando a existência de drusas no olho do paciente.

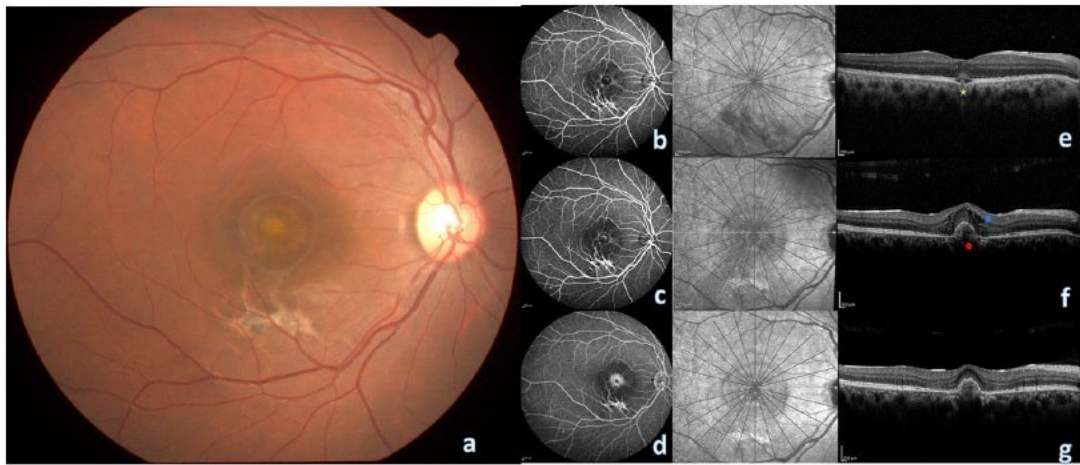


Fonte: [Rocha \(2021\)](#).

2.3.1.2 Degeneração Macular relacionada à idade (úmida)

É mais rara e mais severa. É decorrente do crescimento de novos vasos sanguíneos por baixo da retina e que costumam vaziar líquidos, danificando os fotorreceptores. Também é conhecida como Neovascularização coroidal, do inglês *Choroidal Neo Vascularization - CNV* e pode ser observada na figura 3. A maioria dos casos de cegueira é decorrente da complicação desse quadro ([PORTER, 2022](#)).

Figura 3 – Exemplo de exame de OCT revelando a existência de neovascularização coroidal no olho do paciente.



Fonte: [Venkatesh, Bavaharan e Yadav \(2019\)](#).

2.3.2 Edema Macular Diabético

O Edema Macular Diabético, do inglês *Diabetic Macular Edema* - DME, ilustrado na figura 4, é uma complicação que pode ser decorrente dos dois tipos de diabetes. Configura-se pelo acúmulo de líquido na mácula, região cuja função é focalizar as imagens e enxergar os detalhes ([HEALTHLINE, 2020](#)).

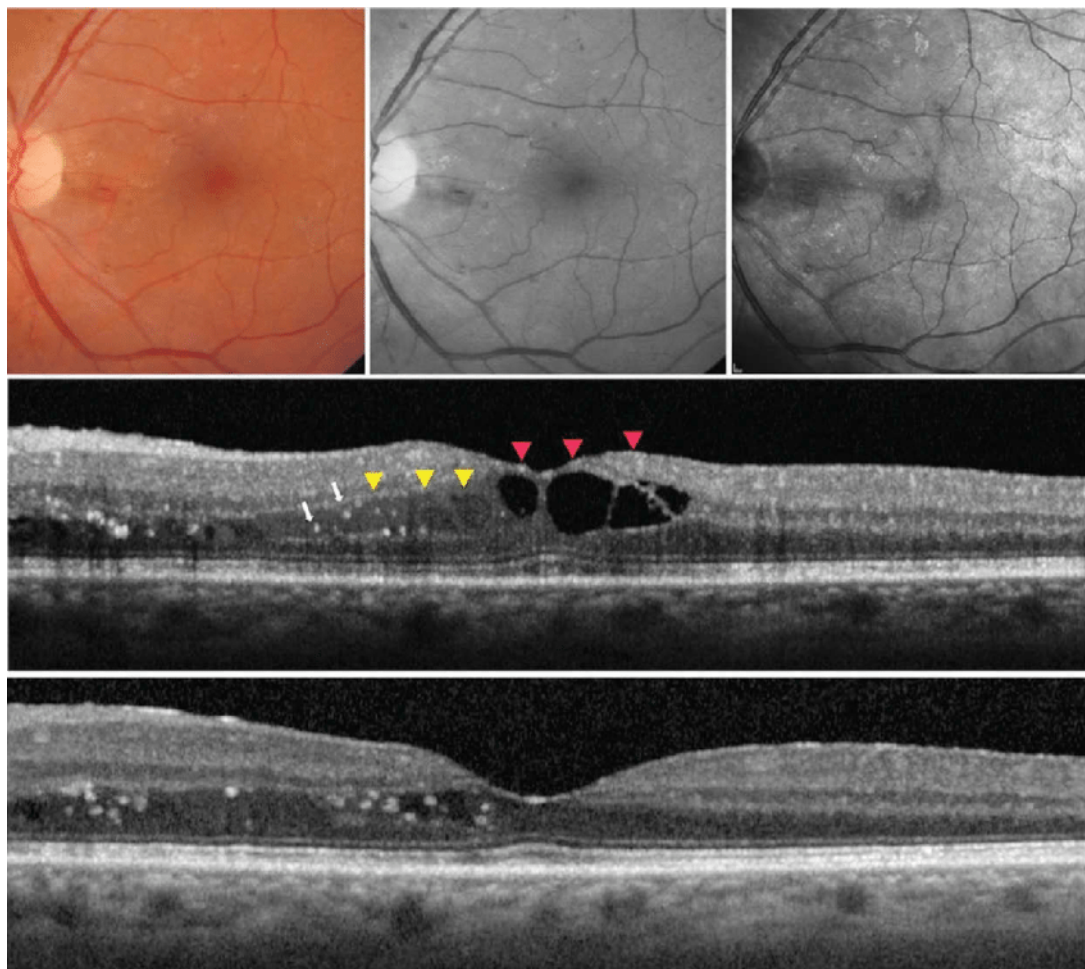
2.4 Processamento de Imagens

Necessário a todos os métodos que se utilizam de imagens, o processamento de imagem é um passo entre a geração da imagem e sua interpretação. Compreende o pré-processamento, onde a imagem sofre a ação de filtros, borramento, binarização, mudança de sistema de cores entre outros processos, e o processamento em si, onde a imagem preparada passará por detecção de objetos, linhas horizontais, verticais, geração de histograma, extração de características ([GONZALEZ RAFAEL C.; WOODS, 2010](#)).

2.4.1 Sistema de cores RGB

O sistema de cores RGB (*Red - Green - Blue*) é o mais conhecido e mais utilizado nas câmeras que encontramos nos aparelhos eletrônicos. A luz de led é baseada nesse sistema de cores. Nele cada cor possui uma saturação na imagem, ou seja, um ponto ótimo que faz a imagem traduzir as cores do momento fotografado ([ALFAMIDIA, 2019](#)).

Figura 4 – Exemplo de exame de OCT revelando a existência de edema macular diabético no olho do paciente.



Fonte: [Goyal et al. \(2017\)](#).

2.4.2 Sistema de cores HSV

O sistema de cor HSV (*Hue - Saturation - Value*), também conhecido como HSB (*Hue - Saturation - Brightness*) é composto por três variáveis: a matiz, que é a cor; a saturação, que é a intensidade da cor, variando de cinza até a cor mais pura; e o brilho, que varia de preto até a cor mais brilhante ([ALFAMIDIA, 2019](#)).

2.4.3 Extração de Características

Cada canal dos sistemas de cores citados acima e também a escala de tons de cinza podem ser representados por um histograma, onde são calculados alguns valores como: mínimo, máximo, média, variância, assimetria ou distorção e curtose.

O mínimo pode ser entendido como o menor valor em um conjunto de dados. Já o máximo será o maior valor contido no mesmo conjunto. A média é um valor calculado pela soma de todos os valores em um conjunto de dados dividido pelo somatório da quantidade

de elementos nesse conjunto. A média indicará a concentração dos valores considerados.

A variância pode ser compreendida como a medida de quão longe os valores estão em relação a um valor médio, ou seja, o quanto uma variável se afasta da média. A raiz quadrada da variância é o desvio padrão, que exprime o quanto os dados de uma amostra variam, conforme [Noletto \(2022\)](#).

Em [Vieira \(2018a\)](#) é mostrada a relação do parâmetro assimetria com a média dos valores da distribuição. Conforme a figura 5, quando a curva é mais pronunciada no início, com sua cauda ou extremidade final alongada, a assimetria é negativa e a média é menor. Quando a curva é mais pronunciada no final, com sua cauda ou extremidade inicial mais alongada, a assimetria é positiva e a média é maior. Dados reais, entretanto, salvo os casos em que se trabalha com grandes amostras, não obrigatoriamente vão ter uma distribuição em forma de sino podendo, às vezes, apresentar-se como distribuições bimodais, com dois picos distintos em sua curva.

O parâmetro curtose, observado na figura 6, é usado para quantificar o quanto de uma variável se encontra nas extremidades da distribuição. São três os tipos de curtose: mesocúrtica, leptocúrtica e platicúrtica. Uma distribuição mesocúrtica é parecida com uma distribuição normal. A curtose de uma distribuição normal funciona como um padrão para se definir os outros dois tipos. A distribuição leptocúrtica tem a curtose maior do que o padrão mesocúrtico, apresentando as extremidades direita e esquerda com uma curvatura maior, mais "pesadas". Em contrapartida têm picos mais altos, fazendo com que haja uma maior quantidade da variável nas extremidades. Já as distribuições platicúrticas têm extremidades com uma curvatura menor, mais "leves", e em geral têm picos mais baixos que as leptocúrticas, fazendo com que haja uma menor quantidade da variável nas suas extremidades e conseqüentemente uma menor curtose ([VIEIRA, 2018b](#)).

Figura 5 – Assimetrias negativa e positiva.

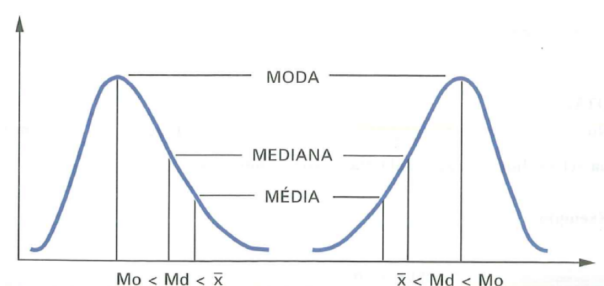
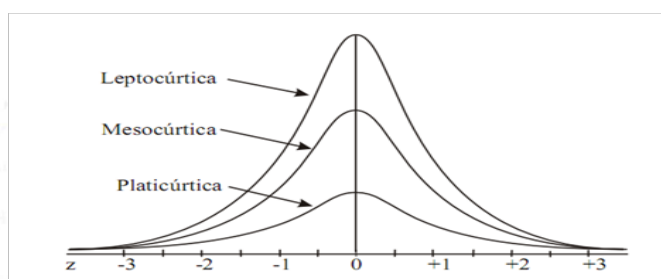


Figura 6 – Tipos de curtose.



Fonte: [Vieira \(2018b\)](#).

Esses indicadores são depois tabelados como banco de dados das fotos utilizadas no estudo para que se possa relacionar as características extraídas através da análise do histograma com a condição de presença ou não de patologia e qual o tipo de doença ocular

possível de ser identificada na imagem. Essa modalidade de extração de características é empregada quando o algoritmo recebe os dados em formato de tabela, em um arquivo com extensão ".csv", por exemplo. Há, entretanto, uma forma de executar a extração de características dentro do próprio algoritmo. Assim é feito quando se aplica redes neurais. A cascata de camadas convolucionais e de *pooling* são responsáveis pela captação dessas informações que servirão posteriormente, na camada "*fully connected*", ao procedimento de classificação.

Um método bastante conhecido e empregado na extração de características de imagens é o Padrão Binário Local. Esse é um método específico para imagens em escala de cinza, tem baixo custo de cálculo e consiste na avaliação de um pixel com base nos pixels que o circundam, forma que permite ao método descrever características locais, detectar linhas, manchas, bordas, auxiliando em segmentação de imagens por textura. Podemos descrever as etapas do algoritmo da seguinte forma:

- 1 - Conversão da imagem para escala de cinza;
- 2 - Cada pixel da imagem terá sua vizinhança selecionada e o pixel central será definido como limiar, para referência;
- 3 - Cada pixel da vizinhança será avaliado com base no pixel central. Se a intensidade do pixel vizinho for maior ou igual a do pixel central, lhe será atribuído o valor 1. Caso sua intensidade seja menor, lhe será atribuído 0.
- 4 - Os valores binários são colocados em sequência, no sentido anti-horário, iniciando pelo pixel posicionado no meio da lateral direita da vizinhança selecionada.
- 5 - O 'valor LBP' é calculado por meio de uma transformação binário-decimal. Este valor será atribuído, na matriz de valores LBP, ao pixel central à vizinhança ([HARMOUCH, 2020](#)).

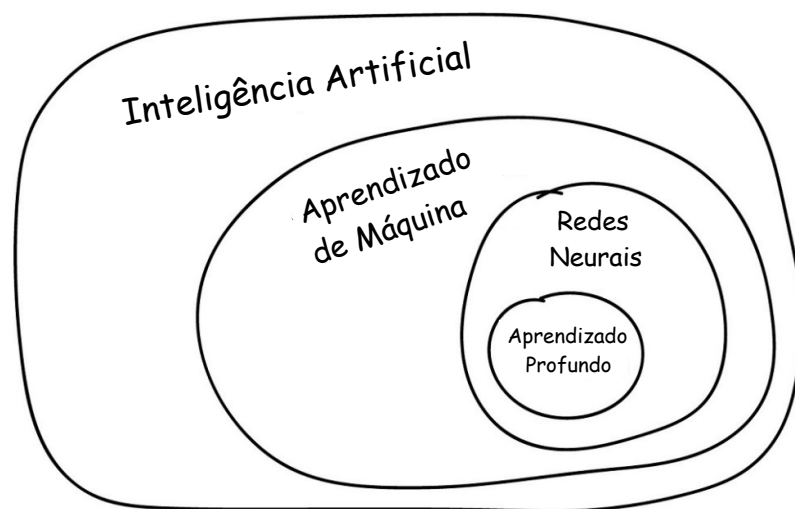
2.5 Inteligência Artificial

A inteligência artificial visa simular o processamento de informações realizado pelo cérebro humano. Esse termo engloba vários outros como o aprendizado de máquina (*Machine learning*), o aprendizado profundo (*Deep learning*) e o processamento de linguagem natural (*Natural Language processing*), que é a compilação de informações oriundas de dados não estruturados para que possam ser analisadas por técnicas de aprendizado de máquina, conforme [Al-Aswad \(2019\)](#).

2.5.1 Aprendizado de Máquina (*Machine Learning*) e Aprendizado Profundo (*Deep Learning*)

O Aprendizado de Máquina é um subcampo da Inteligência Artificial, conforme ilustrado na figura 7. O Aprendizado Profundo é um subcampo do próprio aprendizado de máquina. Neste o algoritmo muda constantemente seus parâmetros à medida que processa mais dados. Esses dados, inclusive, podem ser extraídos primeiro para que sejam submetidos ao algoritmo. A diferença que se encontra no Aprendizado Profundo é que a presença de diversas camadas de código dispensam que se alimente o algoritmo com características específicas, pois ele próprio é capaz de aprender automaticamente através dos parâmetros resultado do treinamento. Necessitam, todavia, de um banco de dados bem maior para executar essa tarefa (AL-ASWAD, 2019).

Figura 7 – Instâncias de cada método computacional no conjunto de técnicas de Inteligência Artificial.



Fonte: Adaptado de Gomes (2019).

2.5.2 Regressão e Classificação

Os algoritmos de aprendizado de máquina, que exercem aprendizado supervisionado, podem executar basicamente duas funções: regressão e classificação. A regressão consiste em algo parecido com a classificação, mas no final é predito um número e não uma categoria. É utilizada para traçar correlações numéricas no geral, um exemplo é a análise do preço de ações da bolsa. As regressões podem ser lineares, polinomiais ou de natureza mais complexa, como a regressão logística, que já não é mais uma técnica de regressão, mas sim de classificação (GOMES, 2019).

Já a classificação divide os objetos da análise em categorias baseadas em atributos conhecidos. O reconhecimento de caracteres e números manuscritos e a detecção de idioma e de fraudes são exemplos da aplicação da classificação. A classificação é um aprendizado

essencialmente supervisionado. É necessário um "professor", ou seja, os dados devem ser rotulados e é baseado nessa informação que o algoritmo trabalha para atribuir classes aos dados. Máquinas de vetor de suporte (SVM) e Árvores de decisão (DT) são alguns exemplos de algoritmos de classificação. Algoritmos de clusterização também realizam classificação, mas de um banco de dados com características desconhecidas. Um exemplo é o KNN (*K-nearest neighbours*) (GOMES, 2019).

2.5.3 Classificação binária X Classificação multiclasse

A classificação binária realiza a predição dos dados do dataset em duas classes, por exemplo, uma classe normal e outra com uma patologia. É mais simples e os algoritmos classificadores mais conhecidos são aplicáveis. Um algoritmo Máquina de Vetor de Suporte - SVM - realiza a tarefa da classificação separando os dados através de um hiper-plano ou uma reta. Esse é o kernel mais simples, o linear, havendo também o radial, o polinomial e o sigmóide. Entretanto na classificação de um banco de dados que contenha mais de duas classes, esse algoritmo já não será mais aplicação. Nessa tarefa, de acordo com a tabela 1, algoritmos como Knn, *Naïve Bayes*, *Decision tree*, *Random forest* e *K-means* são mais apropriados (AUGUSTO, 2020).

Tabela 1 – Adequação dos métodos à classificação binária e à classificação multiclasse

Classificação binária	Classificação multiclasse
SVM	KNN
LDA	RF/DT
GLM	<i>Naïve bayes</i>
<i>Adaboost</i>	<i>K-means</i>

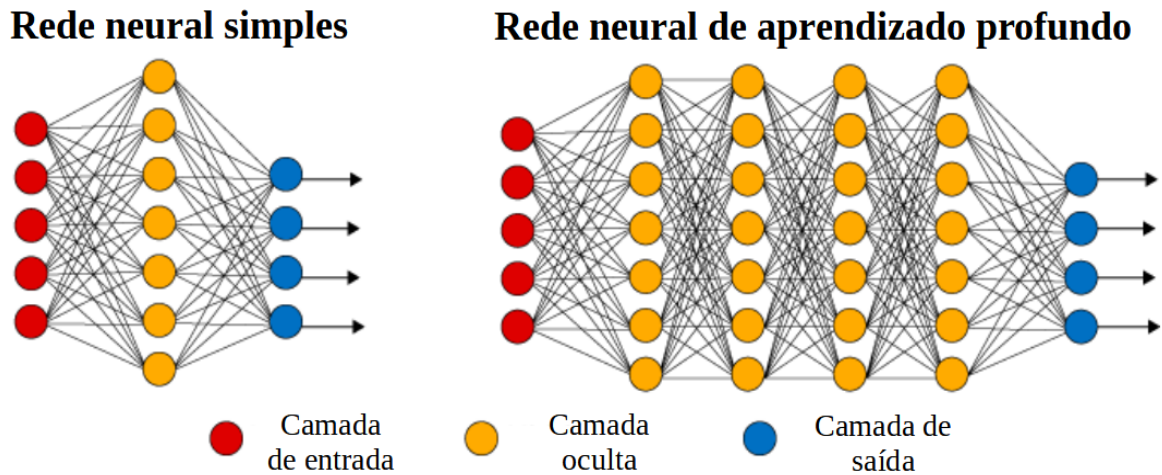
Fonte: Produzido pelo autor.

Essa adequação se deve pela forma como esses métodos realizam a tarefa da classificação. Os algoritmos Knn e *K-means*, por exemplo, trabalham com cálculos de distâncias. O *Naïve Bayes*, com cálculo de probabilidades. *Decision Tree* e *Random Forest*, com base em regras que se repetem recursivamente (AUGUSTO, 2020). Apesar de serem métodos com capacidade de classificação de datasets com mais de duas classes, não deixam de ser métodos efetivos para classificações binárias.

2.5.4 Redes Neurais Artificiais (RNA)

As redes neurais artificiais se baseiam no funcionamento dos neurônios humanos com o objetivo de reproduzir o aprendizado através de exemplos de treinamento. Essas redes são compostas basicamente por camadas de um nó, sendo uma camada de entrada, uma camada de saída e uma camada oculta, conforme a figura 8 (ALVES, 2020).

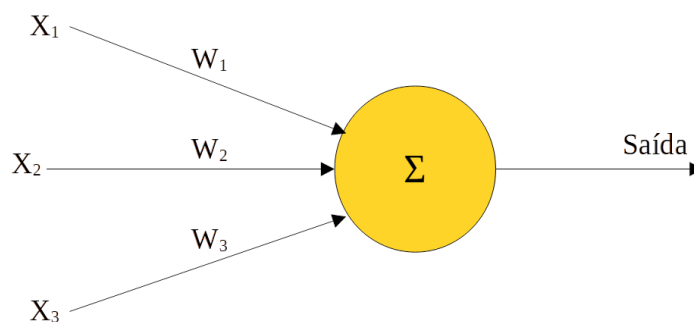
Figura 8 – Ilustração da estrutura de uma rede neural artificial. Para redes mais complexas de deep learning haverá um número bem maior de camadas de neurônios ocultas.



Fonte: Adaptado de [Academy \(2022\)](#).

Cada nó, ou neurônio, ilustrado na figura 9, tem conexão com outro e tem um peso e um valor limite, também chamado de viés. Caso a saída de um neurônio esteja acima do valor limite, esse nó será ativado e enviará o sinal adiante, para a próxima camada. Caso a saída esteja abaixo do valor limite, o nó não será ativado e não repassará o sinal ([CLOUD, 2020](#)).

Figura 9 – Ilustração da estrutura de um neurônio artificial. Na figura, as entradas são multiplicadas por seus respectivos pesos e somadas. Essa soma é avaliada com base no valor limite, ou viés, estabelecido para o neurônio para que a saída atribuída.



Fonte: Produzido pelo autor.

O primeiro tipo de neurônio foi proposto por Frank Rosenblatt em 1958 e foi chamado de *Perceptron*. A proposição desse neurônio foi possível devido a um estudo anterior, realizado por Warren S. McCulloch e Walter Pitts onde foi pesquisada a geração

de padrões por células neurais conectadas. Como resultado da pesquisa concluiu-se que era possível utilizar a lógica booleana para construir um neurônio artificial.

De arquitetura simples, o *Perceptron* logo deixou de atender a problemas mais complexos. Hoje são usados o *Perceptron* multicamadas e o sigmóide em redes de *deep learning* e na tarefa de classificação (ALVES, 2020).

Para entender o funcionamento de uma rede neural artificial faz-se necessário compreender o funcionamento de cada neurônio. Cada nó terá dados de entrada, um conjunto de pesos, um viés e os dados de saída. Os pesos são valores que multiplicam os sinais de entrada de forma a dar maior ou menor importância a esses valores no sinal de saída.

$$\sum_{i=0}^m (w_i x_i + \text{viés}) = w_1 x_1 + w_2 x_2 + w_3 x_3 + \text{viés} \quad (2.1)$$

Após essa multiplicação dos pesos por suas respectivas entradas, os valores são somados e uma comparação será feita com base no limite estabelecido. Caso o valor da soma ultrapasse o viés, o sinal é transmitido para a próxima camada. Caso seja inferior, o sinal não será repassado (CLOUD, 2020).

$$Saída = f(x) = \begin{cases} 1, & \text{se } \sum_{i=0}^m (w_i x_i + \text{viés}) \geq 0 \\ 0, & \text{se } \sum_{i=0}^m (w_i x_i + \text{viés}) < 0 \end{cases} \quad (2.2)$$

Redes cuja transmissão de dados se dá de uma camada para a seguinte são chamadas de *feedforward* e o seu treinamento pode ser realizado através de *backpropagation*, que consiste em retroagir na rede, calculando o erro associado a cada neurônio para que esses valores contribuam para um melhor ajuste dos parâmetros do modelo (CLOUD, 2020).

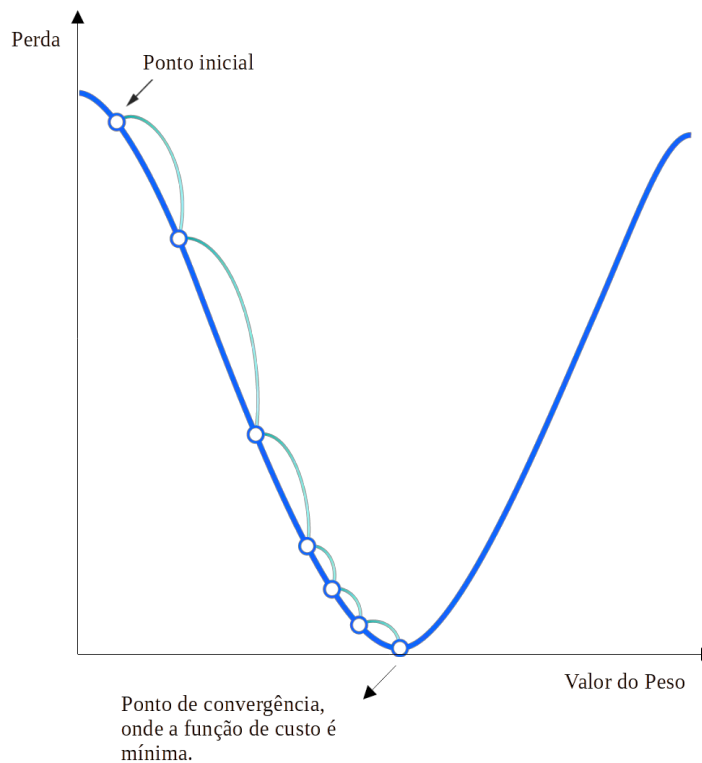
À medida que ocorre o treinamento do modelo, faz-se necessário calcular a sua precisão através de uma função de custo ou de perda, o que conhecemos por Erro Quadrático Médio (MSE). Buscando minimizar a função de custos, o modelo vai reajustando seus pesos e vieses por meio da descida de gradiente, conforme a figura 10.

O que define a arquitetura da rede neural a ser empregada em uma aplicação é o tipo de aprendizado que ela desenvolverá. Dentre os principais tipos de redes tem-se: redes *multilayer perceptron* (MLP), redes neurais recorrentes (RNR) e redes neurais convolucionais (RNC) (ALVES, 2020).

2.5.4.1 Redes Neurais Convolucionais (RNC)

As redes neurais convolucionais (do inglês *Convolutional Neural Networks* - CNN) têm o poder de executar o que os algoritmos supracitados já realizam e um pouco mais, como

Figura 10 – Minimização da função de custo.



Fonte: Adaptado de [Academy \(2022\)](#).

reconhecimento de objetos em fotos, tradução de textos ou de áudios, reconhecimento de fala e tantos outros. Suas unidades básicas de processamento, os neurônios, são organizados em camadas, através das quais a informação é repassada, os pesos são reajustados e por fim se obtém um resultado que é compatível com os algoritmos convencionais, muitas vezes com melhor desempenho ([GOMES, 2019](#)).

Uma rede neural convolucional é composta basicamente por cinco peças: a camada de Convolução, a camada de *Pooling*, a Totalmente Conectada, a camada de *Dropout* e a Função de Ativação, como ilustrado na figura 11. As camadas de Convolução são responsáveis por extrair informações das imagens por meio de filtros que atuam em sub-blocos dessas imagens e que conseguem identificar bordas, linhas, texturas, cores, cantos, etc, podendo até identificar algo mais específico como um objeto ou uma pessoa, em redes com camadas mais profundas ([PASSOS, 2021](#)). A linha de código a seguir exerce a função da camada convolucional, recebendo como argumentos numéricos a quantidade de filtros e o tamanho da janela de convolução e por fim a função de ativação.

$$layers.Conv2D(2, 3, activation = 'relu') \quad (2.3)$$

A camada de *Pooling* recebe os blocos de extração de características das camadas

de convolução e simplifica a informação recebida resumindo-a em um só valor. Em outras palavras, essa camada reduz a resolução de entrada em termos de dimensões espaciais, ou seja, reduz a altura e a largura da matriz, tomando o valor máximo em uma janela de entrada para cada canal de entrada. O argumento '*pool size*' determina o tamanho da janela para a qual será tomado esse valor máximo, por exemplo, (2,2) representará uma janela 2x2.

$$layers.MaxPooling2D(pool_size = (2, 2)) \quad (2.4)$$

A camada Totalmente Conectada recebe a informação e a transforma em uma só linha e inicia o processo de classificação.

$$layers.Dense(128, activation = 'relu') \quad (2.5)$$

A camada de *Dropout* é responsável por reduzir o *overfitting*, que acontece quando o modelo se adequa bem aos dados, mas não apresenta capacidade de generalização, ou seja, não consegue ser aplicado a novos dados.

$$tf.keras.layers.Dropout(0.5) \quad (2.6)$$

Por fim, a Função de Ativação comanda o aprendizado em si, decidindo quais neurônios serão ativados. Na biblioteca *Keras* existem alguns tipos de função de ativação disponíveis. ReLu, Sigmóide e *Softmax* são alguns exemplos. A ativação ReLU aplica a função de ativação da unidade linear retificada.

$$tf.keras.activations.relu(x, alpha = 0.0, max_value = None, threshold = 0.0) \quad (2.7)$$

A segunda ativação disponível é a sigmóide, que retorna sempre um valor entre 0 e 1. Para valores pequenos (<-5) retorna valores próximos de zero. Para valores grandes retorna valores próximos de 1.

$$tf.keras.activations.sigmoid(x) \quad (2.8)$$

Onde:

$$sigmoid(x) = 1/(1 + \exp(-x)) \quad (2.9)$$

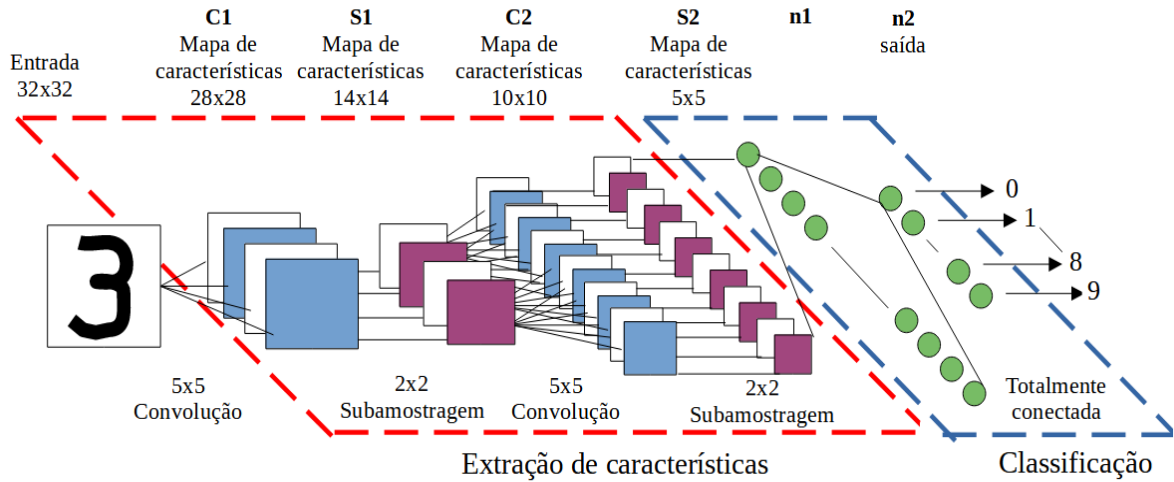
Por fim temos a ativação *Softmax*, que transforma um vetor de valores em uma distribuição de probabilidades, onde os valores estão entre 0 e 1 e somam 1.

$$tf.keras.activations.softmax(x, axis = -1) \quad (2.10)$$

Onde:

$$\text{softmax}(x) = \exp(x) / \text{tf.reduce_sum}(\exp(x)) \quad (2.11)$$

Figura 11 – Esquematisação das camadas que compõem uma Rede Neural Convolucional. Há na ilustração duas regiões distintas: A região de extração de características, delimitada em vermelho, e a região de classificação, destacada em azul. Na camada de classificação estão apenas as camadas totalmente conectadas, enquanto na camada de extração de características pode-se identificar as camadas de convolução e de *pooling*.



Fonte: Adaptado de [Ravindra \(2017\)](#).

O modelo é escolhido, caso já esteja implementado na biblioteca, ou montado, a partir das orientações encontradas na documentação. Em seguida é compilado pela função:

$$\text{model.compile(optimizer = 'adam',} \quad (2.12)$$

$$\text{loss = 'categorical_crossentropy',} \quad (2.13)$$

$$\text{metrics = 'accuracy'}) \quad (2.14)$$

Os três parâmetros citados na função são sempre exigidos e podem ser adaptados de acordo com os dados e o resultado esperado. Para otimizador há as opções 'Adam' e 'SGD' (gradiente descendente), para a função de perda há as definições 'binary', para classificações binárias, 'sparse', para classificações onde os rótulos estão na forma numérica e 'categorical' para dados com rótulos escritos. Na definição das métricas, acurácia, erro médio absoluto são nativas e podem apenas ser citadas como parâmetro. Para calcular métricas diferenciadas como precisão e *recall*, por exemplo, é necessário, para as novas versões da biblioteca, definir uma função para cada uma para que possam ser listadas como métricas na função de compilação do modelo.

Por fim, o modelo é treinado pela função:

$$history = model.fit(training_data = treino, \quad (2.15)$$

$$callbacks = [checkpoint, early_stop], \quad (2.16)$$

$$validation_data = valid, \quad (2.17)$$

$$epochs = 20) \quad (2.18)$$

Nessa função definimos os dados de treinamento, de validação, a quantidade de épocas e os retornos de chamada ou *callbacks*. Dois exemplos são o *Checkpoint* e o *Early Stopping*.

O primeiro, *Checkpoint*, é definido pela seguinte função e possibilita que o melhor conjunto de pesos, ou mesmo o modelo inteiro, seja salvo naquela época que apresentou melhor resultado para a métrica monitorada. Pode-se monitorar acurácia (do treinamento), acurácia de validação, perda de validação, erro médio absoluto ou aquela que melhor se adequar à informação desejada. Também é definido o modo como essa métrica sofrerá monitoramento, por exemplo, se pelo valor mínimo apresentado, ou se pelo valor máximo. Para monitoramento de acurácia de validação, espera-se que o valor seja máximo. Já para perdas, espera-se que o valor seja mínimo.

$$checkpoint = ModelCheckpoint('class.hdf5', \quad (2.19)$$

$$save_weights_only = True, \quad (2.20)$$

$$monitor = 'val_accuracy', \quad (2.21)$$

$$mode = 'max') \quad (2.22)$$

O segundo, *Early Stopping*, realiza uma parada quando for detectado que a métrica monitorada não apresentou melhoria nos resultados. Possibilita também definir após quantas épocas ficará decidido que não houve melhora na métrica monitorada, para que os pesos sejam salvos e o treinamento encerrado. Esse dado é denominado '*patience*'. É possível também definir um valor numérico que, ao ser detectado como acréscimo na métrica monitorada, representará uma melhora nos resultados (??).

$$early_stop = EarlyStopping(monitor = 'val_accuracy', \quad (2.23)$$

$$min_delta = 0.01, patience = 2, mode = 'max', \quad (2.24)$$

$$verbose = 1, baseline = None, \quad (2.25)$$

$$restore_best_weights = True) \quad (2.26)$$

2.5.5 Métricas de avaliação

2.5.5.1 Acurácia

A acurácia é a métrica mais conhecida para avaliar uma classificação. Representa a porcentagem de acerto na classificação realizada, calculada como a razão entre a soma dos Verdadeiros positivos e Verdadeiros negativos pela soma do total de classificações, em outras palavras, é a razão entre as previsão corretas e o total de amostras. Pode ser representada pela seguinte equação, onde os índices subscritos correspondem a: v(verdadeiro), p(previsto) e a(amostras) ([LEARN, 2022](#)).

$$Acurácia(y_v, y_p) = \frac{1}{n_a} \left(\sum_{i=0}^{n_a-1} (y_p = y_v) \right) \quad (2.27)$$

Ou

$$Acurácia = \frac{VP + VN}{VP + VN + FP + FN} \quad (2.28)$$

Na equação anterior, entende-se como numerador o número de classificações corretas, os Verdadeiros Positivos (VP), quantidade identificada como a diagonal da matriz de confusão. Já o denominador é representado como uma soma dos Verdadeiros positivos (VP), Verdadeiros negativos (VN), Falsos positivos (FP) e Falsos negativos (FN). Essa soma representa todo o conjunto utilizado, é o total de amostras.

2.5.5.2 Precisão

A precisão, também denominada valor preditivo positivo (VPP), é uma métrica empregada quando o número de resultados Falsos positivos são mais relevantes/prejudiciais do que o de Falsos negativos. Em um caso de avaliação de viabilidade de investimentos financeiros, o Falso positivo indicará que um bom investimento é ruim, enquanto o Falso negativo indicará que um mau investimento é bom. Nessa situação, bons investimentos deixarão de ser feitos. Um valor de precisão alto, indicará que esse quantitativo é baixo ([RODRIGUES, 2019](#)).

$$Precisão = \frac{VP}{VP + FP} \quad (2.29)$$

2.5.5.3 Recall

O *recall*, também conhecido como sensibilidade ou taxa de acerto, é uma métrica empregada quando o número de resultados Falsos negativos são mais relevantes/prejudiciais

do que o de Falsos positivos. Neste caso, considerando a situação de classificação de exames para detectar doenças, os Falsos negativos indicarão que a pessoa doente não tem a doença, e os Falsos positivos indicarão que a pessoa não tem a doença. O primeiro indicador é mais preocupante pois tira a preocupação do paciente sobre ter ou não uma enfermidade, sendo que ele a tem. Então, para essa situação, um valor de *recall* alto será desejado, indicando que a quantidade de Falsos negativos é baixa (RODRIGUES, 2019).

$$Recall = \frac{VP}{VP + FN} \quad (2.30)$$

2.5.5.4 Erro médio absoluto

O MAE é o erro médio absoluto e pode ser calculado pela seguinte equação:

$$MAE(y_v, y_p) = \frac{1}{n_a} \left(\sum_{i=0}^{n_a-1} |y_v - y_p| \right) \quad (2.31)$$

onde os índices subscritos correspondem a: v(verdadeiro), p(previsto) e a(amostras).

2.5.5.5 Erro quadrático médio

O MSE é o erro médio quadrático e pode ser calculado pela seguinte equação:

$$MSE(y_v, y_p) = \frac{1}{n_a} \left(\sum_{i=0}^{n_a-1} |y_v - y_p|^2 \right) \quad (2.32)$$

2.5.5.6 Raiz do erro médio quadrático

O RMSE é a raiz do erro médio quadrático e pode ser calculado pela seguinte equação:

$$RMSE(y_v, y_p) = \sqrt{\frac{1}{n_a} \left(\sum_{i=0}^{n_a-1} |y_v - y_p|^2 \right)} \quad (2.33)$$

2.5.6 Matriz de confusão

A matriz de confusão é uma forma de avaliação que permite comparar os quantitativos referentes à classificação em cada categoria. Em outras palavras, a matriz de

confusão guarda os valores referentes aos Verdadeiros positivos e negativos e aos Falsos positivos e negativos.

O Verdadeiro positivo é a classificação exata. Significa dizer que a pessoa possui a doença em questão e o algoritmo acertou em prever que a pessoa tem a doença. Esse é o quantitativo que chama a atenção em um primeiro momento, afinal ele traduz a capacidade da inteligência artificial utilizada em prever o diagnóstico correto. O Verdadeiro negativo ocorre quando a pessoa não tem a doença e o algoritmo foi assertivo em prever o mesmo. Ambos os verdadeiros confirmam o status do paciente acerca de ter ou não uma patologia.

Já os Falsos são valores discordantes. O Falso positivo significa que a pessoa não tem a doença, mas o algoritmo concluiu que ela tem a doença, configurando uma superestimação no resultado. Apesar de apresentar um resultado diverso da realidade, a realização de outros exames ou a análise de um especialista podem sanar as dúvidas.

Por fim temos os Falsos negativos. Esse quantitativo significa que a pessoa examinada tem a doença, mas o algoritmo previu que ela não tem a doença, havendo assim uma subestimação no diagnóstico. Esse valor também chama a atenção e gera preocupação pois representa um alívio para o paciente, sendo que ele está doente. A redução dos valores de Falsos negativos representa um melhor desempenho do algoritmo empregado e uma maior confiabilidade no diagnóstico automatizado (BHANDARI, 2022).

A matriz de confusão 2x2, conforme apresentado na figura 12, é a mais simples e permite o entendimento mais facilitado do significado dos valores distribuídos nela. Classificações, entretanto, nem sempre serão binárias. Para classificações de bancos de dados em mais de duas classes, as variáveis da matriz de confusão 2x2 apresentadas acima mudarão. A matriz exemplificada na figura 13 tem dimensões 4x4 e abaixo dela há uma lista com os valores que correspondem aos Verdadeiros Positivos, Verdadeiros Negativos, Falsos Positivos e Falsos Negativos. Nessa matriz, algumas das variáveis terão seu valor final resultado de uma soma (BHANDARI, 2022).

2.6 A linguagem Python

A linguagem *python* foi criada na década de 90 por Guido Van Rossum com o objetivo de ser intuitiva, de fácil aprendizado, de código aberto e utilizável para projetos dos mais variados tipos de complexidade (INFNET, 2021). *Python* é uma das linguagens de programação mais populares do mundo, logo atrás de Java e C. Para trabalhar com ciência de dados, *Python* possui as bibliotecas *Numpy*, *Pandas* e *Matplotlib*, para trabalhar respectivamente com matrizes grandes, análise de dados e visualização de dados. Para a implementação de Aprendizado de Máquina em grande escala, *python* é a linguagem adequada, possuindo bibliotecas especializadas em Aprendizado de máquina e Aprendizado profundo como *Scikit-learn*, *Keras* e *TensorFlow*. Para executar essas rotinas há, por

Figura 12 – Formatação de uma matriz de confusão 2x2. Na figura está exemplificada uma matriz de confusão para uma classificação em duas categorias: uma denominada "Sim" e outra denominada "Não". Esses exemplos de classes podem ser interpretados como ter ou não ter determinada doença. No eixo vertical temos os valores reais, que são os rótulos dos dados de entrada no algoritmo. No eixo horizontal temos os valores preditos, que serão calculados a partir dos pesos obtidos com o treinamento. Comparando os valores desses dois eixos tem-se: os Verdadeiros Positivos, que são os valores cuja predição concorda com o valor real, ou seja, a pessoa tem a doença assim como é afirmado pelo rótulo; os Verdadeiros negativos, que também significam uma concordância dos valores reais e preditos, mas dessa vez em favor da segunda classe considerada, a pessoa não tem patologia e o rótulo é confirmado pela inteligência artificial.

		Valores preditos	
		Classe 1-Sim	Classe 2-Não
Valores reais	Classe 1-Sim	Verdadeiro Positivo	Falso Negativo
	Classe 2-Não	Falso Positivo	Verdadeiro Negativo

Fonte: Produzido pelo autor.

exemplo, as plataformas *Jupyter Notebook* e *Google Colaboratory* ([CLOUD, 2021](#)).

Python é uma linguagem de uso mais geral, multifuncional, oferecendo ao usuário bibliotecas como *Numpy* para modelagem numérica de dados, *Pandas* para análise de dados, *Scipy* para computação e cálculos científicos e *Scikit-learn* para algoritmos de aprendizado de máquina.

Figura 13 – Formatação de uma matriz de confusão 4X4. Abaixo da matriz é possível observar as combinações que compõem os Verdadeiros Positivos, Verdadeiros Negativos, Falsos Positivos e Falsos Negativos para cada uma das 4 classes.

		Valores Preditos			
		Classe 1	Classe 2	Classe 3	Classe 4
Valores Reais	Classe 1	A	B	C	D
	Classe 2	E	F	G	H
	Classe 3	I	J	K	L
	Classe 4	M	N	O	P

Classe 1

VP = A
VN = F + G + H + J + K + L + N + O + P
FP = E + I + M
FN = B + C + D

Classe 2

VP = F
VN = A + C + D + I + K + L + M + O + P
FP = B + J + N
FN = E + G + H

Classe 3

VP = K
VN = A + B + D + E + F + H + M + N + P
FP = C + G + O
FN = I + J + L

Classe 4

VP = P
VN = A + B + C + E + F + G + I + J + K
FP = D + H + L
FN = M + N + O

VP = Verdadeiros Positivos
VN = Verdadeiros Negativos
FP = Falsos Positivos
FN = Falsos Negativos

3 Metodologia

As imagens utilizadas são oriundas do banco de dados apresentado por [Kermany et al. \(2018\)](#) e disponibilizada para download na plataforma *Mendeley*. O banco possui quatro categorias, a saber: NORMAL (exemplificada na figura 14), DME (exemplificada na figura 15), CNV (exemplificada na figura 16) e DRUSEN (exemplificada na figura 17), com uma alíquota maior em uma pasta reservada para o treinamento e outra menor em uma pasta dedicada ao teste. Neste trabalho foi utilizada apenas a pasta original dedicada ao treinamento e o algoritmo utilizado fez a subdivisão desse material para os processos de treino, validação e teste.

Figura 14 – Normal.

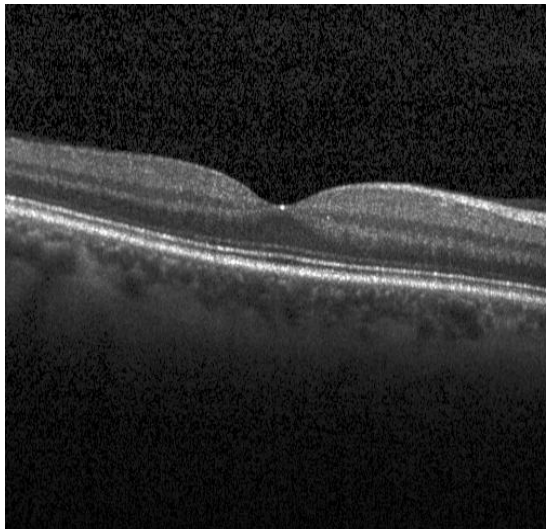
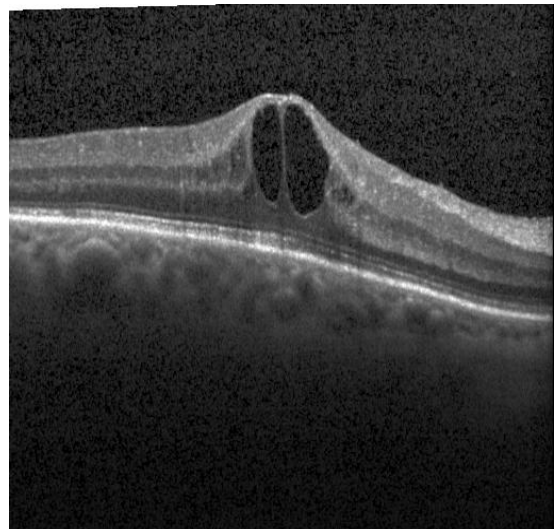


Figura 15 – Edema macular diabético.



Fonte: [Kermany et al. \(2018\)](#).

3.1 Mapa Resumo

No mapa de ideias ilustrado na figura 18 é apresentado um resumo da metodologia utilizada com a indicação da técnica escolhida para classificação e do tamanho do volume de dados (imagens) utilizado.

3.2 Coleta de dados

As imagens de exames de OCT submetidas à classificação pertencem ao banco de dados apresentado em [Kermany et al. \(2018\)](#) e disponibilizadas na internet. São

Figura 16 – Neo-vascularização coroidal.

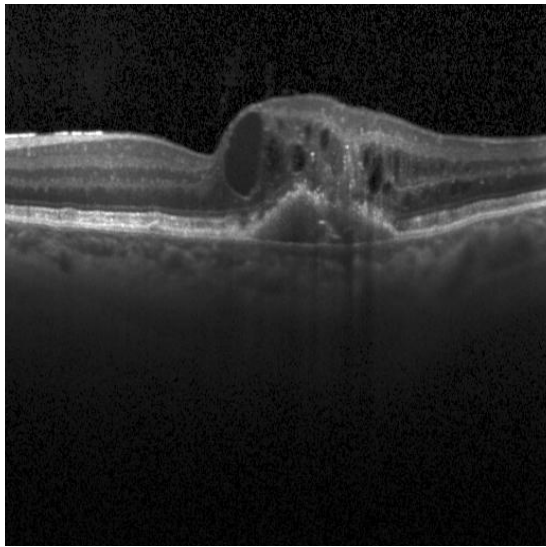
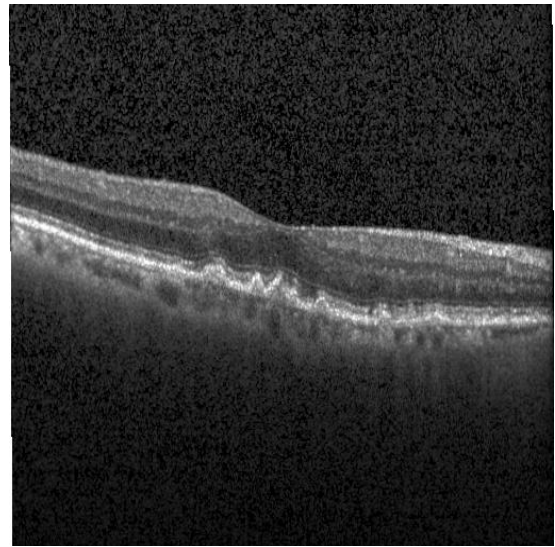


Figura 17 – Drusas.



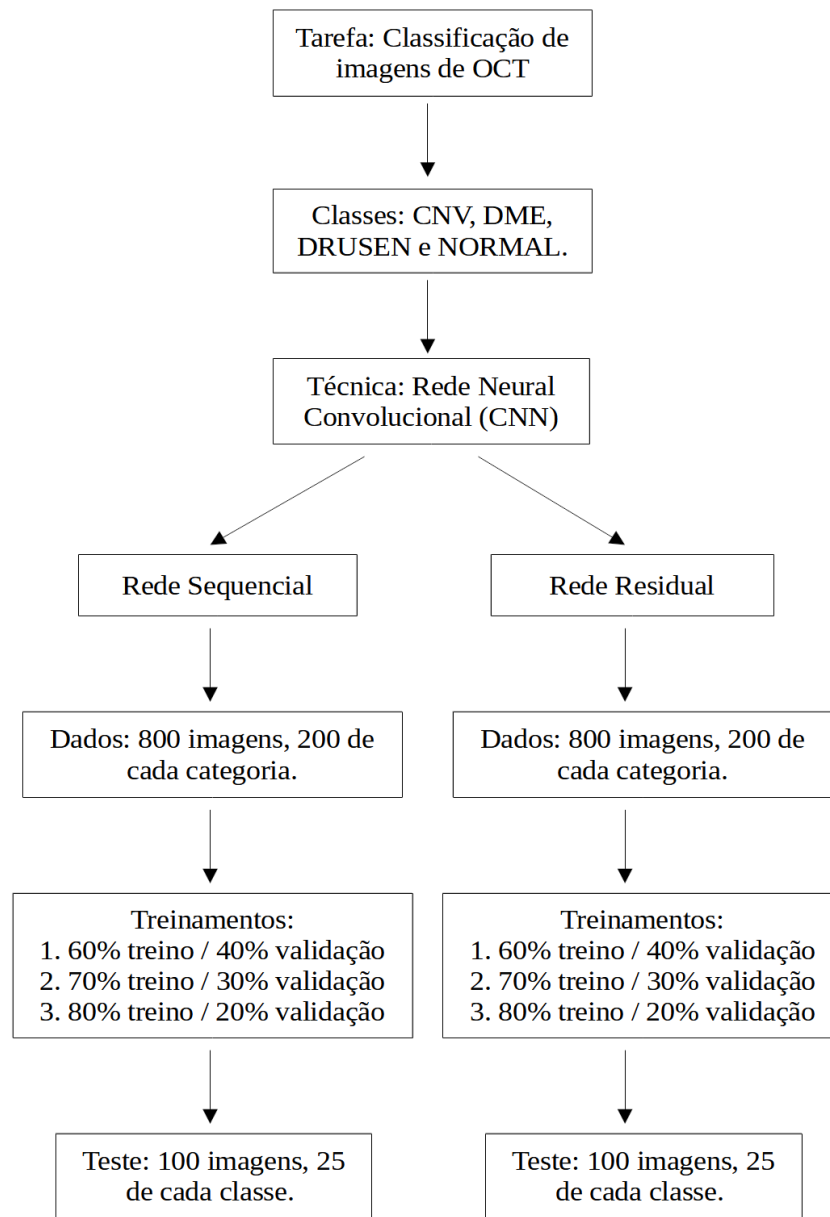
Fonte: [Kermany et al. \(2018\)](#).

84.484 imagens alocadas em duas pastas: a de treinamento com 83.484 e a de teste com 1.000 imagens, 250 para cada categoria. Foi utilizada apenas a pasta dedicada ao treinamento. Esta pasta está dividida em 4 subpastas, a de imagens normais, que conta com 26.315 arquivos e as de três patologias: a neovascularização coroidal (CNV) ou degeneração macular relacionada à idade do tipo úmida, com 37.205 arquivos; o edema macular diabético (DME), com 11.348 arquivos; e a presença de drusas (DRUSEN), que revelam uma precoce degeneração macular relacionada à idade do tipo seca, com 8.616 arquivos. Conforme [Das, Dandapat e Bora \(2019\)](#), a presença de drusas pode ser considerada um estágio inicial de degeneração macular como resultado da idade, enquanto a neovascularização coroidal configuraria um estágio mais avançado, tardio.

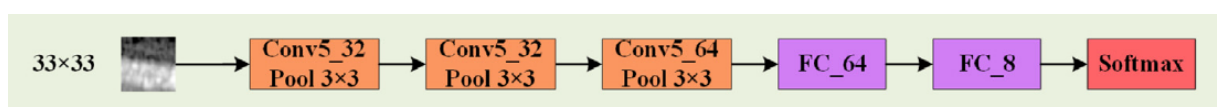
3.3 Forma de análise dos dados

Primeiramente foi feita uma subamostragem do banco de dados. Foi retirada de forma aleatória uma parcela de 200 imagens das pastas originais de cada categoria, formando uma pasta de treinamento com 800 imagens. Para o teste foram salvas em outra pasta 25 imagens inéditas de cada categoria, também escolhidas aleatoriamente, totalizando 100 arquivos. Em um segundo momento, foram aplicadas diretamente ao grupo de 800 imagens duas redes neurais convolucionais, a primeira baseada em uma rede sequencial exibida em [Hu et al. \(2019\)](#), ilustrada na figura 16. A segunda é a rede Resnet50. As duas redes selecionadas foram escolhidas dentre aquelas utilizadas nos artigos da revisão bibliográfica.

Figura 18 – Mapa de Ideias da metodologia utilizada neste trabalho.



Fonte: Produzido pelo autor.

Figura 19 – Esquema da rede sequencial apresentada no artigo [Hu et al. \(2019\)](#). Neste diagrama tem-se um patch de 33x33 de uma imagem sendo submetido a três ciclos de camadas de Convolução seguidas de *Pooling*. Logo em seguida há duas camadas sequenciais totalmente conectadas e por último o classificador *softmax*.Fonte: [Hu et al. \(2019\)](#).

Nessa primeira rede, que é sequencial, cada imagem é recortada em uma amostra de 33x33. Em seguida é submetida por três vezes a uma camada convolucional de kernel 5x5 e *pooling* de filtro 3x3 para reduzir a resolução. Logo após, duas camadas '*Fully connected*' e o classificador *softmax*. Essa rede foi baseada na arquitetura CIFAR, criada para classificar 10 categorias de imagens coloridas de pequeno tamanho e baixa resolução (HU et al., 2019).

A alíquota do banco usada foi de 800 imagens em 20 épocas, batch size de 50 e 10 treinamentos. Foram executados três treinamentos, cada um com porcentagens diferentes do banco para treino e validação. No primeiro, 60% do volume de dados foi reservado para treino e 40% para validação. No segundo, 70% para treino e 30% para validação. No terceiro, 80% para treino e 20% para validação.

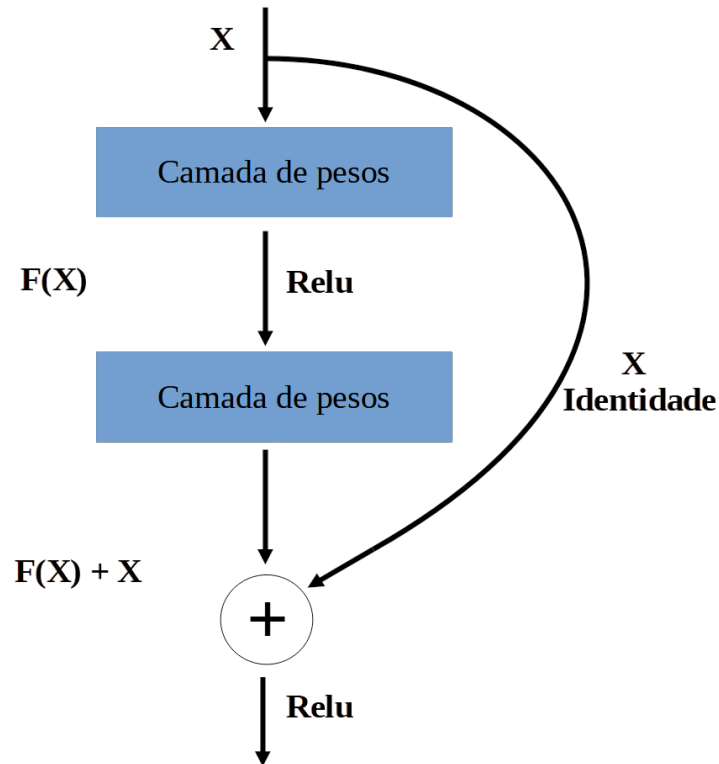
A segunda rede neural convolucional utilizada foi a Resnet50. Essa é uma rede residual, com 50 camadas. A característica residual é relativa a um tipo de conexão específica presente em sua arquitetura, criada para contornar o problema do gradiente de fuga/explosão, que ocorria quando o erro tendia a zero ou a um valor muito grande à medida que se aumentava o número de camadas. Essas ligações específicas são chamadas de conexões de salto, conforme ilustrado na figura 20. Seu funcionamento se concentra no fato de permitir que no caso de alguma camada prejudicar a rede, haverá a possibilidade de uma regularização, onde a rede tenta se adequar ao mapeamento residual (LIMA, 2022). A alíquota tomada para a classificação foi a mesma da rede anteriormente citada, 800 imagens, 200 de cada categoria, e os treinamentos foram feitos com as mesmas três divisões para treino e validação: 60%/40%, 70%/30% e 80%/20% , em 20 épocas, batch size de 50 e 10 treinamentos.

Ambas foram aplicadas ao banco de dados em linguagem *Python* na plataforma *Google Colaboratory*, online. As bibliotecas usadas foram *Scikit-learn*, *Keras* e *TensorFlow*.

Ambas as redes foram executadas com os callbacks (retorno de chamada) *Checkpoint* e *Early Stopping*, nativos da biblioteca *Keras*. A função *Checkpoint* permite que seja realizada a atualização dos pesos e o salvamento do melhor conjunto. A função *Early Stopping* permite monitorar uma métrica e parar a execução quando essa variável não apresentar mais melhora.

A característica intrínseca dos dados, a forma exibida em um histograma de frequências, por exemplo, orienta a escolha do método de classificação. Os classificadores do tipo regressão são mais adequados a dados padronizados e recebem dados de extração de características tabelados, salvos em arquivos '.csv', enquanto as CNN's têm melhor performance quando aplicadas a dados com maior variedade, sendo que utilizam as próprias imagens durante todo o processo.

Figura 20 – Formatação genérica do bloco residual, também chamado conexão de salto, característica da rede Resnet50.



Fonte: Produzido pelo autor.

3.4 Forma de análise dos resultados

Os resultados são analisados por meio de métricas como a acurácia, mae, rmse, precisão e recall. Também há uma comparação com as classificações já realizadas e citadas na literatura, os classificadores utilizados e as acurácias alcançadas. Para a análise dos dados utilizados são plotados gráficos com os histogramas de cada atributo para verificar sua forma de distribuição e a adequação ao método utilizado.

4 Resultados e Discussões

O desempenho das redes neurais empregadas é avaliado através dos valores de acurácia e erro correspondentes aos treinamentos realizados e dos valores calculados a partir da matriz de confusão resultante da aplicação do melhor conjunto de pesos a um grupo de imagens inéditas para teste.

Será analisado o melhor resultado obtido, que foi aquele resultado da divisão do banco em 80% para treino e 20% para validação. Na seção 4.3.1 da discussão serão apresentados os resultados obtidos para as outras divisões do banco, 70%/30% e 60%/40%, com seus respectivos gráficos e valores de acurácia e erros.

Abaixo os resultados são apresentados e discutidos.

4.1 Resultados obtidos com a aplicação da CNN sequencial

Os dados apresentados aqui foram obtidos com a execução de dois métodos, uma CNN sequencial e a CNN Resnet50, em 10 treinamentos, com 20 épocas cada, e para uma divisão do banco de 80% para treino e 20% para teste. As matrizes de confusão apresentadas são referentes à aplicação da rede à pasta de teste, que possui ao todo 100 arquivos de imagem '.jpeg'. O conjunto de pesos utilizado para isso foi o último melhor conjunto salvo pelo retorno de chamada '*Early Stopping*'.

Tabela 2 – Matriz de Confusão da CNN sequencial

real \ predito	CNV	DME	DRUSEN	NORMAL
CNV	5	15	0	5
DME	4	6	5	10
DRUSEN	2	2	6	15
NORMAL	0	0	5	20

Fonte: Produzido pelo autor.

A partir da matriz de confusão apresentada na tabela 2 e com base na matriz 4x4 mostrada na fundamentação teórica deste trabalho, a tabela 3 traz os valores de Verdadeiros Positivos, Verdadeiros Negativos, Falsos Positivos e Falsos Negativos referentes ao resultado. Com esses valores são calculados em seguida, na tabela 4, para cada uma das categorias classificadas (CNV, DME, DRUSEN e NORMAL), acurácia, *recall* e precisão.

Tabela 3 – Da matriz de confusão anterior, tem-se:

Classes	VP	VN	FP	FN
CNV	5	69	6	20
DME	6	58	17	19
DRUSEN	6	65	10	19
NORMAL	20	45	30	5

Fonte: Produzido pelo autor.

Tabela 4 – Acurácia, precisão e *recall*:

Classes	Acurácia	Precisão	<i>Recall</i>
CNV	74%	45,45%	20%
DME	64%	26,09%	24%
DRUSEN	71%	37,5%	24%
NORMAL	85%	40%	80%

Fonte: Produzido pelo autor.

Observando a matriz de confusão, a classe que apresenta o maior número de classificações corretas é a NORMAL, apesar de a predição para essa classe apresentar muitas classificações erradas (Falsos positivos). Para todas as outras classes o número de Verdadeiros Positivos é baixo. Para uma análise inicial do acerto da classificação, ao observar a coluna de Verdadeiros positivos (VP) da tabela 3 e sabendo que para o teste são usados 25 arquivos de cada classe, então a acurácia apenas para os verdadeiros positivos é de 20% para a classe CNV, 24% para DME e DRUSEN, e 80% para a classe NORMAL.

Apesar dos Falsos Positivos apresentarem um número alto, esses valores não são tão prejudiciais, pois diagnosticar que alguém saudável apresenta uma doença levará a pessoa a realizar novos exames e a confusão será sanada. Os valores de Falsos Negativos é que são perigosos, caso altos, pois indicam que a pessoa não tem qualquer problema de saúde.

A matriz em análise, entretanto, apresenta como categoria com melhor classificação aquela que representa o grupo de pacientes normais, sem doença. Nesse caso, o valor de Falsos Positivos é que é mais importante e deve apresentar um número baixo, o que não acontece. Das predições para NORMAL, 5 são CNV verdadeiro, 10 são DME verdadeiro e 15 são DRUSEN verdadeiro, ou seja, em 30 casos de 100, patologias oculares recebem o diagnóstico NORMAL. Das 25 amostras verdadeiras para NORMAL, 20 são corretamente classificadas e apenas 5 recebem diagnóstico de patologia.

A classe NORMAL apresenta acurácia de 85% na classificação, *recall* de 80% e precisão de 40%. A métrica *recall* diz respeito a uma análise feita com a referência na linha de valores reais e leva em consideração o número de Falsos Negativos. Já a precisão analisa a coluna de valores preditos para a classe, levando em consideração o número de

Falsos positivos. Apesar de apresentar o maior número de acertos na classificação, a classe NORMAL apresenta baixa precisão.

4.2 Resultados obtidos com a aplicação da rede Resnet50

O segundo algoritmo a ser utilizado também foi aplicado a mesma quantidade de imagens, 800, em uma parcela de 80% para treino e 20% para validação, para a execução de 10 treinamentos. Uma segunda pasta, com 100 imagens, foi usada para o teste e para a matriz de confusão.

Tabela 5 – Matriz de Confusão da rede Resnet50

real \ predito	CNV	DME	DRUSEN	NORMAL
CNV	25	0	0	0
DME	21	1	0	3
DRUSEN	4	0	21	0
NORMAL	0	1	0	24

Fonte: Produzido pelo autor.

A partir da matriz de confusão apresentada na tabela 5 e com base na matriz 4x4 mostrada na fundamentação teórica deste trabalho, a tabela 6 traz os valores de Verdadeiros Positivos, Verdadeiros Negativos, Falsos Positivos e Falsos Negativos referentes ao resultado. Com esses valores são calculados em seguida, na tabela 7, para cada uma das categorias classificadas (CNV, DME, DRUSEN e NORMAL), acurácia, *recall* e precisão.

Tabela 6 – Da matriz de confusão anterior, tem-se:

Classes	VP	VN	FP	FN
CNV	25	50	25	0
DME	1	73	1	24
DRUSEN	21	75	0	4
NORMAL	24	72	3	1

Fonte: Produzido pelo autor.

Tabela 7 – Acurácia, precisão e *recall*:

Classes	Acurácia	Precisão	<i>Recall</i>
CNV	75%	50%	100%
DME	74%	50%	4%
DRUSEN	96%	100%	84%
NORMAL	96%	88,88%	96%

Fonte: Produzido pelo autor.

Analisando a matriz de confusão, observa-se que três categorias alcançaram valores altos para os Verdadeiros Positivos. A classe CNV teve as 25 amostras classificadas corretamente. A classe DRUSEN teve 21 amostras classificadas corretas dentre 25 casos e a classe NORMAL teve 24 classificações corretas dentre 25. Analisando apenas os Verdadeiros positivos e sabendo que cada classe tem 25 arquivos na pasta de teste, tem-se as acurácias: 100% para a classe CNV, 4% para DME, 84% para DRUSEN, e 96% para a classe NORMAL.

Em relação aos valores de Falsos Negativos, que se mostram importantes para as três classes de patologias, a classe CNV não apresentou classificações para esse indicador, levando a um *recall* de 100%. A classe DME foi a que mais apresentou erros, exibindo o mais baixo *recall*, 4%. A classe DRUSEN apresenta, juntamente com a classe NORMAL, uma acurácia de 96% nas suas respectivas classificações. Apresentam também altos *recall* e precisão.

Entretanto, a interpretação desses dois indicadores é diferente para as duas classes. Conforme já comentado, na classe com doença o indicador mais importante é o número de Falsos Negativos, que é a soma dos outros valores da mesma linha do Verdadeiro Positivo em questão.

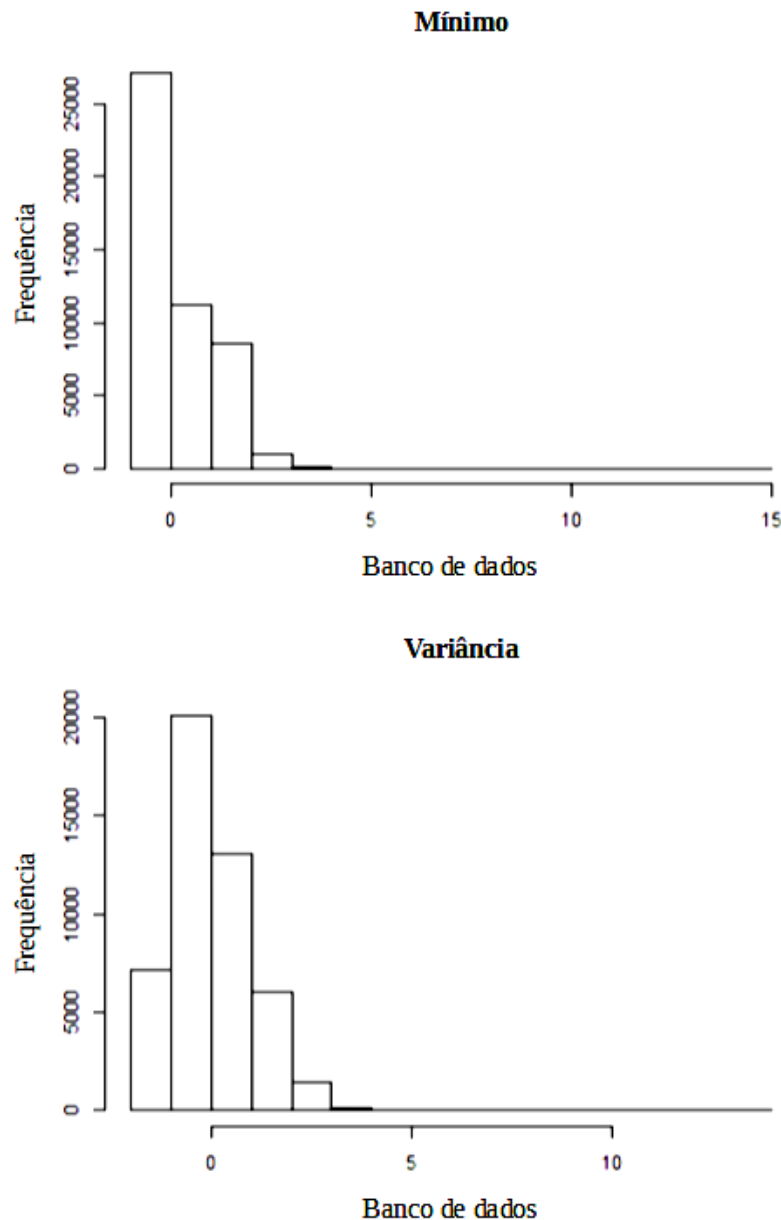
As classes CNV, DRUSEN e NORMAL apresentam baixo número de Falsos negativos, o que justifica o alto *recall*. Mas para a interpretação dos resultados da classe NORMAL, o indicador mais importante é o número de Falsos Positivos, que indicam que a pessoa recebeu um diagnóstico negativo apesar de apresentar patologia. Esse valor influencia a precisão, que para essa classe, foi de 88,88%.

4.3 Discussão

A partir da revisão bibliográfica realizada é possível observar que a classificação realizada nos artigos pesquisados, principalmente naqueles que usam classificadores de *machine learning*, é binária, ou seja, apenas duas categorias passam pela classificação. Por exemplo, um banco é classificado entre imagens com olhos normais e imagens com olhos com alguma patologia [Lee, Baughman e Lee \(2017\)](#), [Farsiu et al. \(2014\)](#), [Thomas et al. \(2021b\)](#), ou então a classificação ocorre entre imagens com uma patologia precoce e imagens com a mesma patologia, mas em estágio mais avançado [Burgansky-Eliash et al. \(2005\)](#). Neste trabalho houve uma tentativa de classificar todas as categorias do banco utilizado juntas, método adotado em [Das, Dandapat e Bora \(2019\)](#), [Kuwayama et al. \(2019\)](#), [Tsuji et al. \(2020\)](#). Em [Das, Dandapat e Bora \(2019\)](#), [Tsuji et al. \(2020\)](#), a classificação envolve as mesmas quatro categorias classificadas neste trabalho e o mesmo banco de dados. Particularmente neste trabalho, a classificação foi realizada com a aplicação de redes neurais convolucionais, alcançando resultados compatíveis com aqueles apresentados nos

artigos de referência.

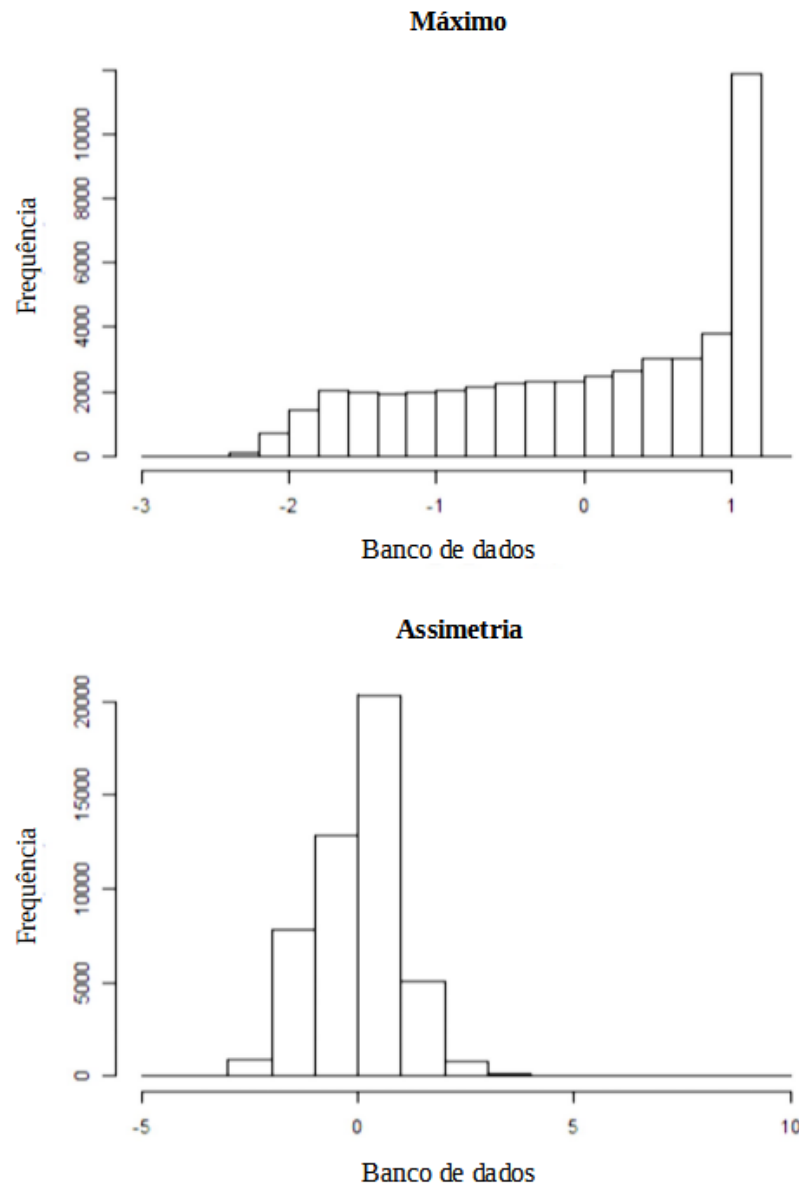
Figura 21 – Conforme os dados sejam discretos, pode-se plotar os histogramas de frequências para os atributos máximo, mínimo, média, variância, assimetria e curtose oriundos da extração de características das imagens. Nesta figura é possível observar os histogramas para mínimo e variância.



Fonte: Produzido pelo autor.

Conforme as figuras 21 a 26, há atributos que apresentam uma distribuição exponencial, enquanto outro apresenta uma distribuição mais próxima da normal. Assim, com escalas diferentes, os dados não são tão bem classificados por métodos de instância, como o KNN, enquanto métodos de regressão são mais eficazes com dados de entrada padronizados. Essa não é, entretanto, uma verdade absoluta, o que deve ser avaliado com testes de variados métodos, por exemplo: Regressão logística (LG) e Análise discriminante

Figura 22 – Conforme os dados sejam discretos, pode-se plotar os histogramas de frequências para os atributos máximo, mínimo, média, variância, assimetria e curtose oriundos da extração de características das imagens. Nesta figura é possível observar os histogramas para máximo e assimetria.

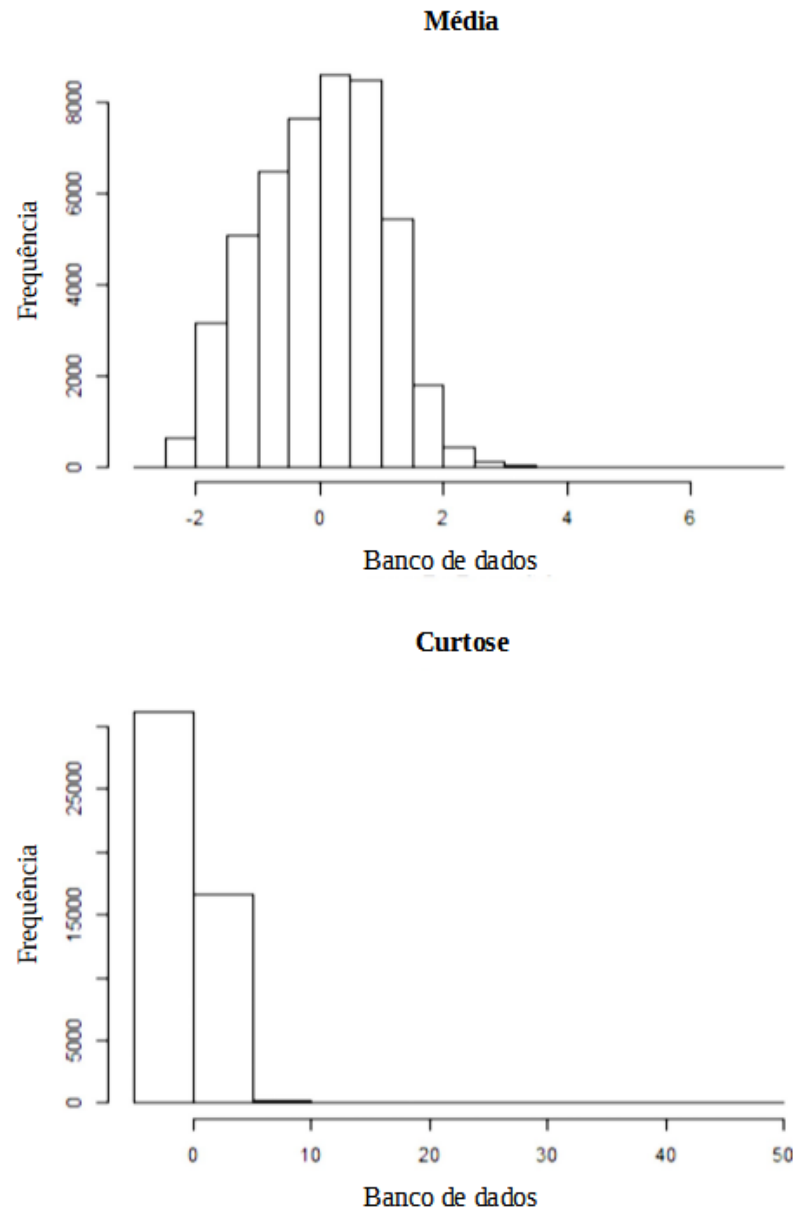


Fonte: Produzido pelo autor.

linear (LDA), para métodos lineares, e KNN e SVM, para métodos não-lineares (MONIZ, 2019).

Um algoritmo de *Machine Learning*, pela natureza do aprendizado a partir dos dados, necessita de um grande volume de informações para ter melhor performance, o que é proposto em Kuwayama et al. (2019) a partir do aumento considerável do número de imagens a partir de operações de rotação e translação sobre as originais, procedimento

Figura 23 – Conforme os dados sejam discretos, pode-se plotar os histogramas de frequências para os atributos máximo, mínimo, média, variância, assimetria e curtose oriundos da extração de características das imagens. Nesta figura é possível observar os histogramas para média e curtose.



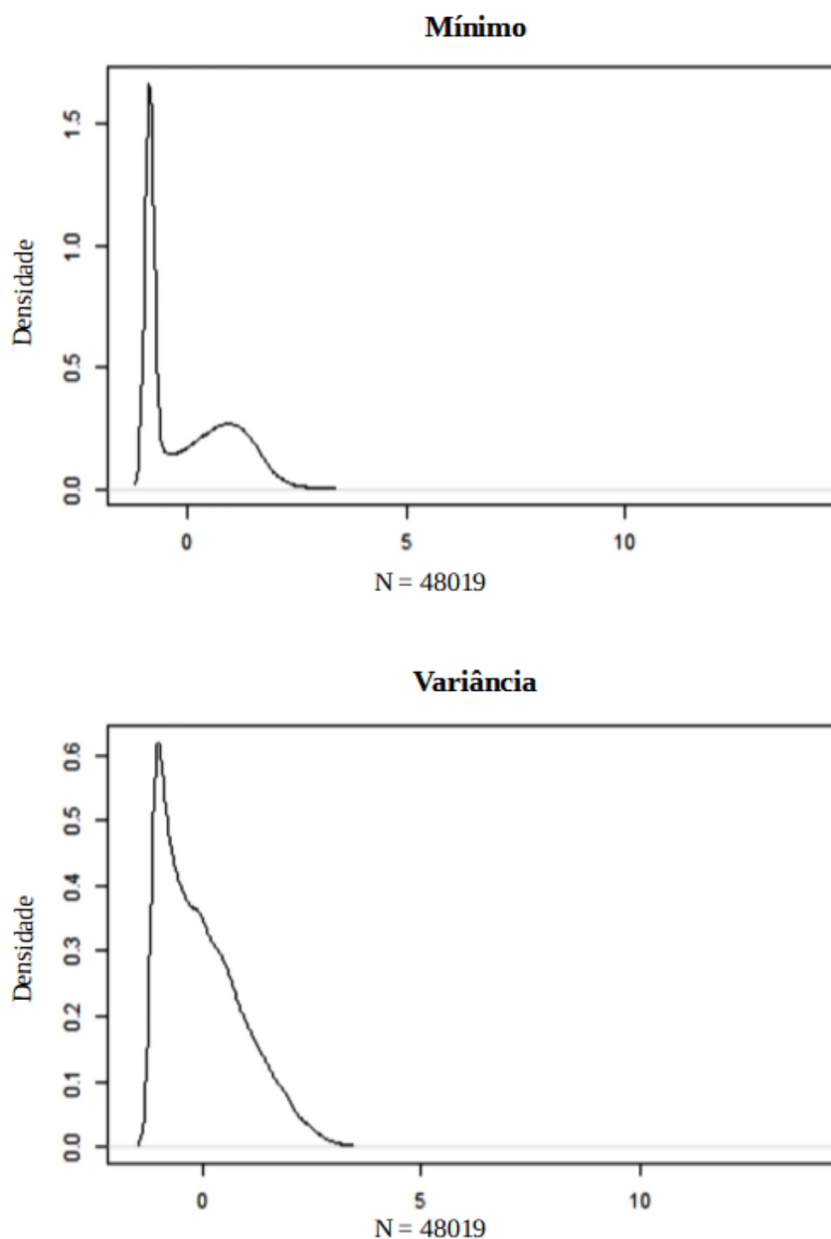
Fonte: Produzido pelo autor.

também conhecido por "aumento de dados".

Com base no desbalanceamento dos dados apresentados nas figuras 21 a 26, o emprego de redes neurais convolucionais para a tarefa de classificação é uma alternativa.

Há uma diferença peculiar entre a aplicação de algoritmos de classificação de *machine learning* e a aplicação de redes neurais convolucionais. A entrada dos dados para os classificadores transforma cada matriz de imagem em uma linha, não considerando

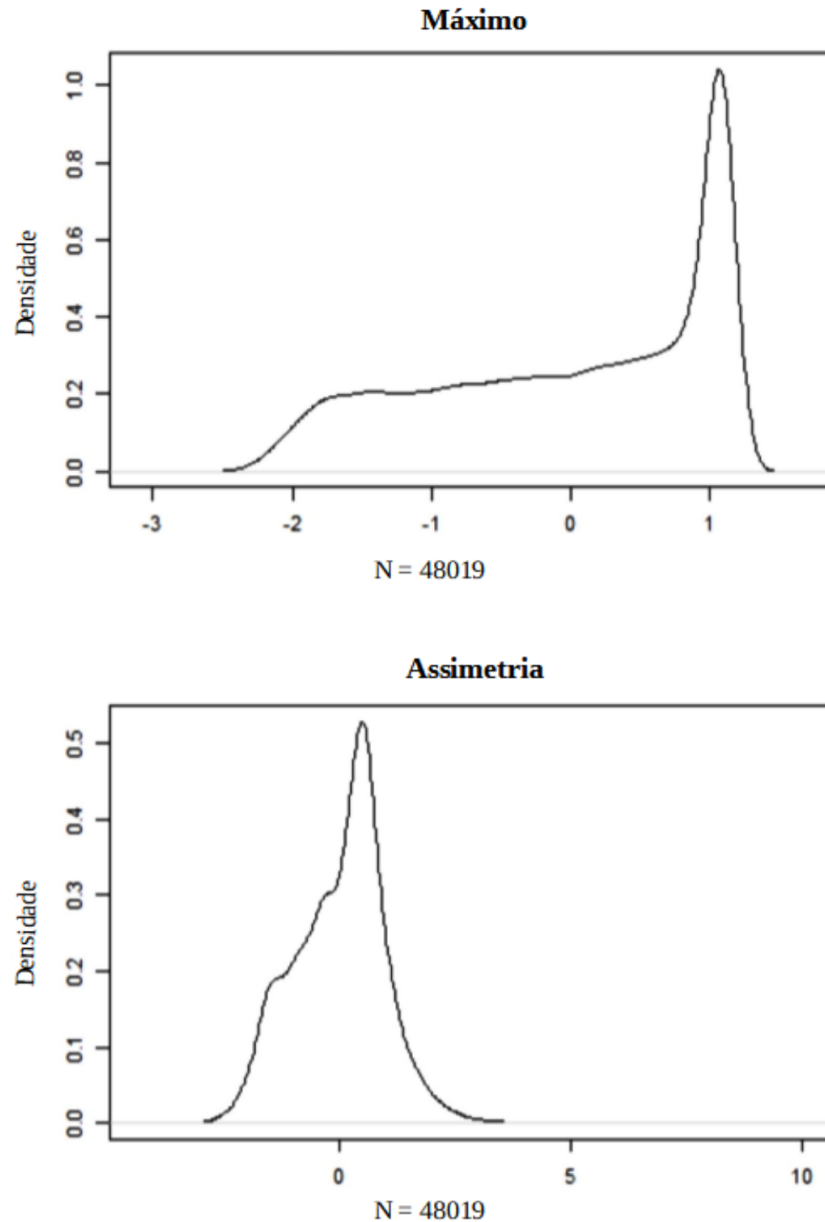
Figura 24 – Gráficos de densidade para os atributos máximo, mínimo, média, variância, assimetria e curtose calculados sobre os histogramas de extração de características das imagens. Nesta figura é possível observar os gráficos de densidade para mínimo e variância.



Fonte: Produzido pelo autor.

a posição de cada pixel. Já na aplicação de redes neurais, os dados são recebidos como 'tensor', formato que carrega as informações de cada imagem na forma de matriz e a quantidade de *labels* presentes no *dataset*. No início do processamento, a camada de convolução realiza a extração de características e leva em consideração o posicionamento de cada informação na imagem, pois trabalha mantendo o formato dos dados de entrada.

Figura 25 – Gráficos de densidade para os atributos máximo, mínimo, média, variância, assimetria e curtose calculados sobre os histogramas de extração de características das imagens. Nesta figura é possível observar os gráficos de densidade para máximo e assimetria.

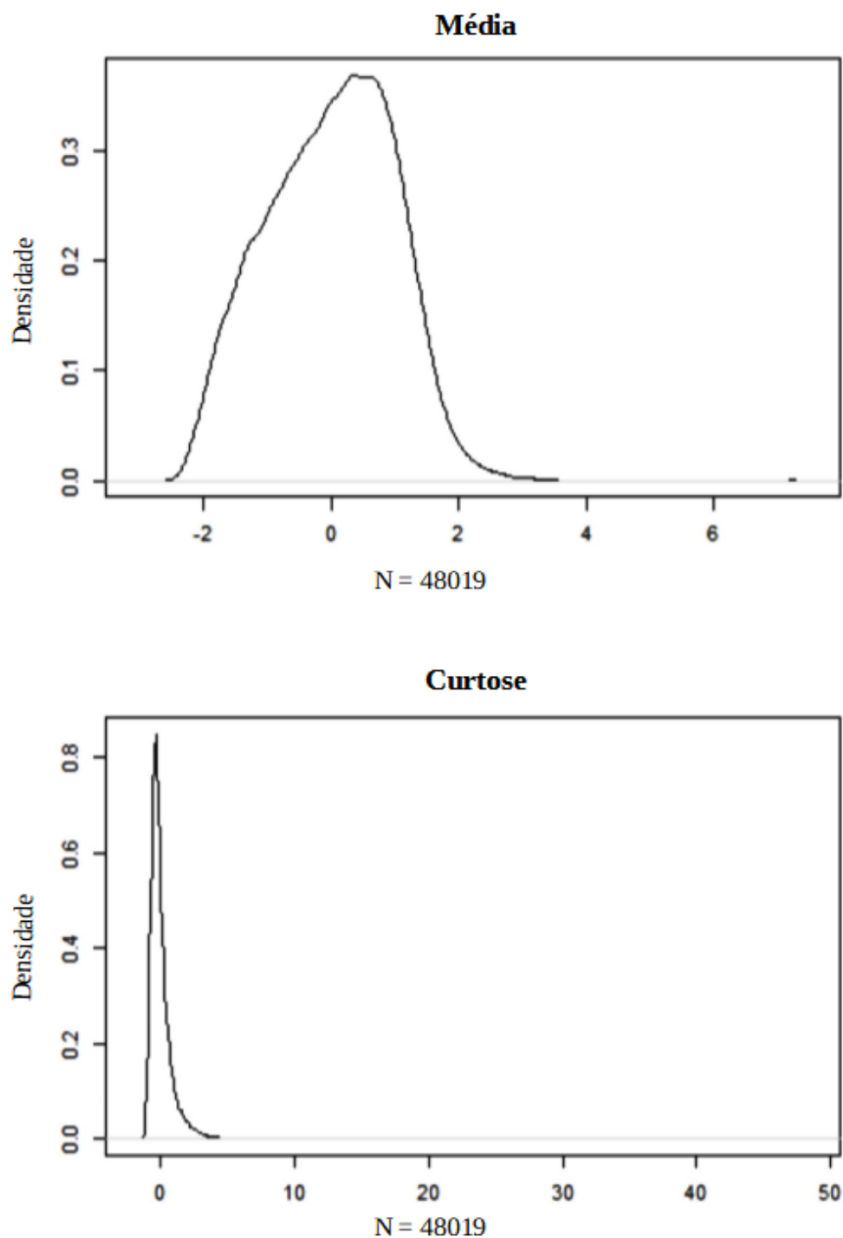


Fonte: Produzido pelo autor.

4.3.1 Gráfico comparativo de acurácias de treino e validação para 3 partições do banco

Pode-se observar nas figuras 27, 28 e 29 a evolução das acurácias de treino e validação, em 10 treinamentos, para cada um dos dois modelos de CNN empregados.

Figura 26 – Gráficos de densidade para os atributos máximo, mínimo, média, variância, assimetria e curtose calculados sobre os histogramas de extração de características das imagens. Nesta figura é possível observar os gráficos de densidade para média e curtose.



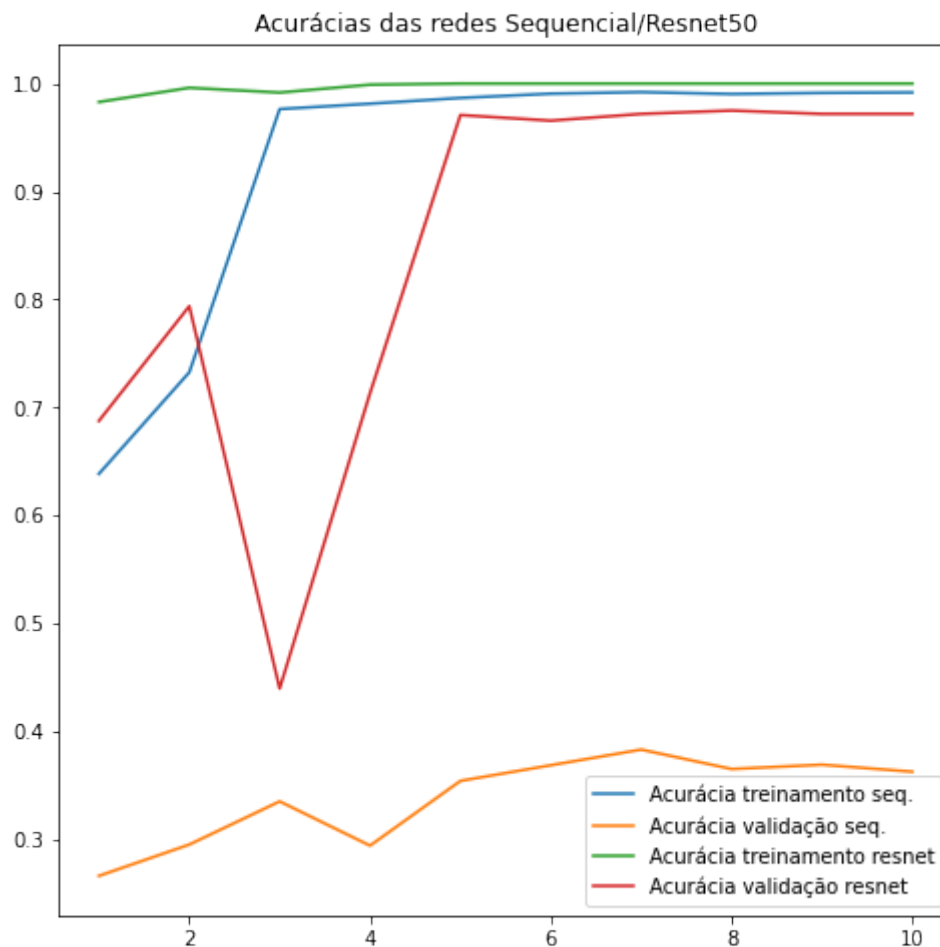
Fonte: Produzido pelo autor.

4.3.1.1 Divisão 60% treino/40% validação

A CNN sequencial apresenta, no último ciclo de treinamento, acurácia média de validação de 36,25%, contra 97,18% de acurácia média da CNN residual (Resnet50), como é possível observar na figura .

Comparando os valores de acurácia de treino e validação no décimo treinamento, tem-se um erro médio absoluto de 62,93%, um erro quadrático médio de 39,63% e um

Figura 27 – Evolução de acurácias de treinamento e validação, em 10 treinamentos, para a CNN sequencial e a CNN residual numa divisão do banco de 60%/40%.



Fonte: Produzido pelo autor.

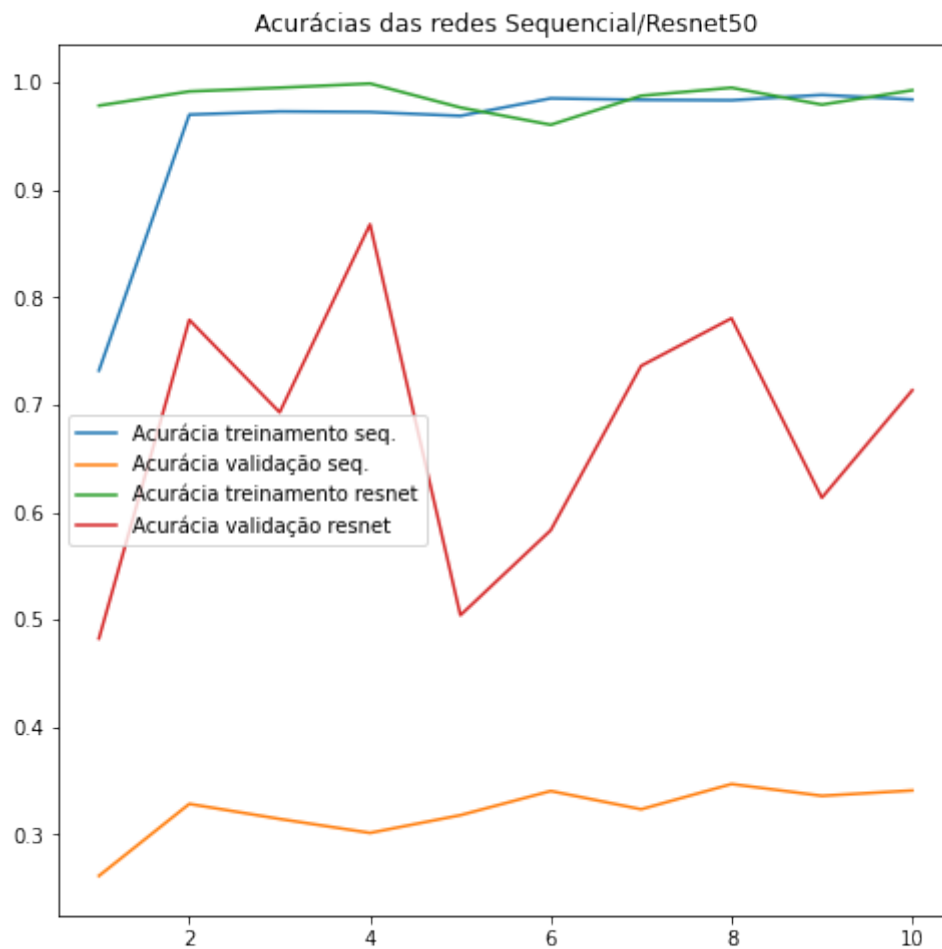
RMSE de 62,96% para a primeira rede. Para a segunda rede, tem-se um erro médio absoluto de 2,81%, um erro quadrático médio de 0,079% e um RMSE de 2,81%.

4.3.1.2 Divisão 70% treino/30% validação

A CNN sequencial apresenta, no último ciclo de treinamento, acurácia média de validação de 34,10%, contra 71,35% de acurácia média da CNN residual (Resnet50), como é possível observar na figura 28.

Comparando os valores de acurácia de treino e validação no décimo treinamento, tem-se um erro médio absoluto de 64,33%, um erro quadrático médio de 41,47% e um RMSE de 64,40% para a primeira rede. Para a segunda rede, tem-se um erro médio absoluto de 27,93%, um erro quadrático médio de 11,99% e um RMSE de 34,62%.

Figura 28 – Evolução de acurácias de treinamento e validação, em 10 treinamentos, para a CNN sequencial e a CNN residual numa divisão do banco de 70%/30%.



Fonte: Produzido pelo autor.

4.3.1.3 Divisão 80% treino/20% validação

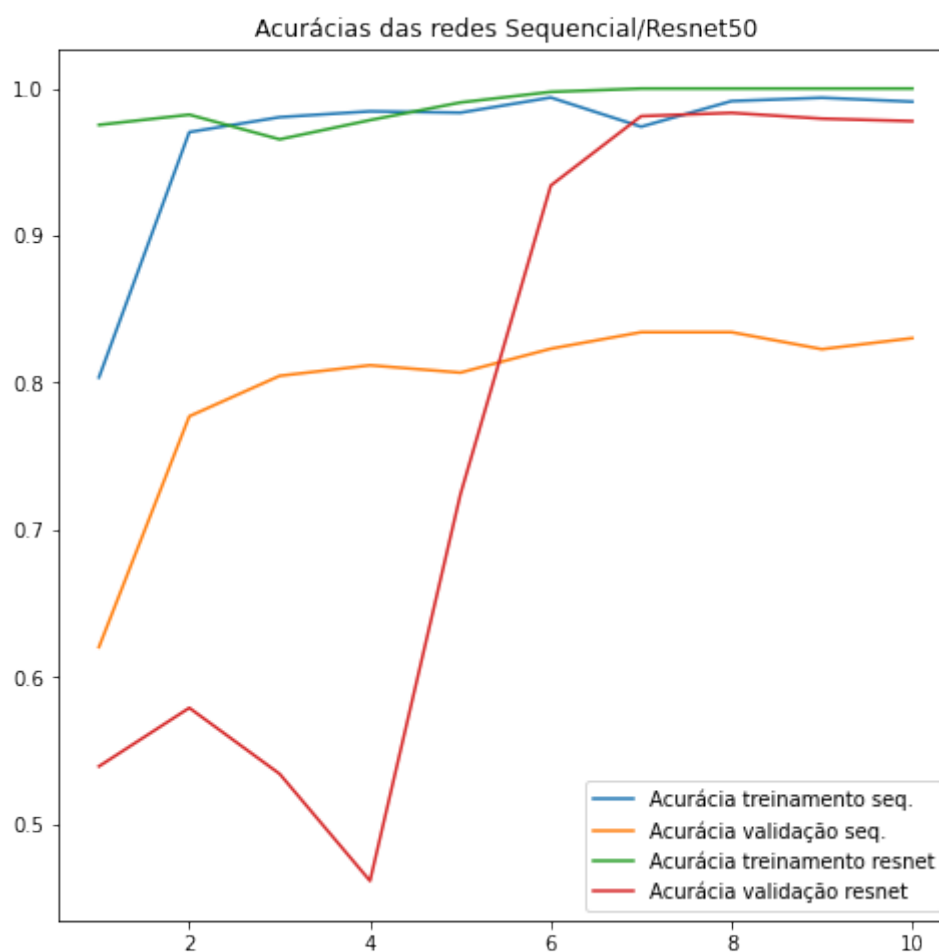
A CNN sequencial apresenta, no último ciclo de treinamento, acurácia média de validação de 83,03%, contra 97,78% de acurácia média da CNN residual (Resnet50), como é possível observar na figura 29.

Comparando os valores de acurácia de treino e validação no décimo treinamento, tem-se um erro médio absoluto de 16,08%, um erro quadrático médio de 2,61% e um RMSE de 16,17% para a primeira rede. Para a segunda rede, tem-se um erro médio absoluto de 2,21%, um erro quadrático médio de 0,06% e um RMSE de 2,46%.

4.3.2 Tabela geral de acurácias

São apresentadas nesta seção duas tabelas. A primeira, tabela 8, contém um resumo das acurácias obtidas para cada uma das 3 divisões de banco consideradas submetidas a cada uma das duas redes neurais empregadas. As partições utilizadas foram 60%/40%,

Figura 29 – Evolução de acurácias de treinamento e validação, em 10 treinamentos, para a CNN sequencial e a CNN residual numa divisão do banco de 80%/20%.



Fonte: Produzido pelo autor.

70%/30% e 80%/20%. As redes neurais utilizadas foram a sequencial e a resnet50. Em seguida, a tabela 9 contém um resumo com as melhores acurácias obtidas neste trabalho e aquelas obtidas nos artigos de referência, para comparação dos valores.

Tabela 8 – Tabela resumo de valores de acurácia para as três partições do banco de dados

Partições	Método	Acurácia
60%/40%	CNN Sequencial	36,25%
60%/40%	CNN Resnet50	97,18%
70%/30%	CNN Sequencial	34,10%
70%/30%	CNN Resnet50	71,35%
80%/20%	CNN Sequencial	83,09%
80%/20%	CNN Resnet50	97,78%

Fonte: Produzido pelo autor.

Tabela 9 – Tabela geral de valores de acurácia de classificação para os artigos pesquisados e para este trabalho

Trabalhos	Métodos	Acurácia
Este trabalho	CNN Sequencial	83,09%
Este trabalho	CNN Resnet50	97,78%
Das, Dandapat e Bora (2019)	CNN	99,6%
Farsiu et al. (2014)	GLM	>99%
Kuwayama et al. (2019)	CNN	83%
Lee, Baughman e Lee (2017)	CNN	87%
P et al. (2021)	CNN	99,69%
Thomas et al. (2021b)	CNN + SVM, RF, MLP	99,78%
Tsuji et al. (2020)	CNN	99,6%

Fonte: Produzido pelo autor.

5 Conclusões

Em conformidade com os trabalhos pesquisados que foram referência para este estudo, a aplicação de redes neurais convolucionais são possíveis para a realização de classificação de imagens de Tomografia de Coerência Óptica. As classificações realizadas pelos dois modelos empregados apresentaram valores compatíveis com a literatura. Foi utilizada uma rede sequencial, montada camada a camada e uma rede já implementada pela biblioteca *Keras*, *Resnet50*, que apresentou o melhor resultado, 97,78% de acurácia média em 10 treinamentos.

Este trabalho contribui para o desenvolvimento da aplicação de algoritmos computacionais em diagnóstico médico ao apresentar uma diferença de abordagem em comparação com grande parte dos artigos pesquisados, pois foi classificada uma quantidade maior de categorias, mais especificamente as quatro classes contidas no banco de dados. Na maioria dos trabalhos, a classificação é geralmente binária, apenas duas situações são consideradas, por exemplo, uma categoria com patologia e outra normal, ou uma categoria com patologia em estágio inicial e outra em estágio mais avançado.

No método empregado neste trabalho foi realizada a classificação de quatro quadros clínicos. Metodologia que amplia o espectro de atuação de um diagnóstico automatizado, apesar dessa abordagem ainda ser uma ferramenta auxiliar, que necessita de estudos ainda mais aprofundados para que possa no futuro ser utilizada junto a um(a) especialista.

5.1 Trabalhos futuros

Considerando a nova realidade da inteligência artificial e de suas possibilidades na área médica, novas patologias precisam ser encaixadas em estudos de classificação, bancos de dados mais abrangentes precisam ser gerados e métodos de otimização devem ter sua atuação avaliada para que se reduza ao máximo o quantitativo de falsos negativos, métrica que desfavorece os algoritmos utilizados pelo fato de representar um grupo de pacientes que teria alguma patologia, mas cujo diagnóstico é negativo.

Referências

ACADEMY, D. S. **Capítulo 3 – O Que São Redes Neurais Artificiais Profundas ou Deep Learning?** 2022. <https://www.deeplearningbook.com.br/o-que-sao-redes-neurais-artificiais-profundas/>.

AL-ASWAD, R. K. B. T. W. L. A. Artificial intelligence and optical coherence tomography imaging. **Asia-Pacific Journal of Ophthalmology**, Asia Pacific Academy of Ophthalmology, 2019.

ALFAMIDIA. **Modelos de cor: HSB ou HSV**. 2019. <http://www.alfamidia.com.br/15-01-2019-modelos-de-cor-hsb-ou-hsv/>. Accessed: 03.04.2022.

ALVES, P. M. **Inteligência Artificial e Redes Neurais**. 2020. <https://www.ipea.gov.br/cts/pt/central-de-conteudo/artigos/artigos/106-inteligencia-artificial-e-redes-neurais>.

AN, G. et al. Glaucoma diagnosis with machine learning based on optical coherence tomography and color fundus images. **Journal of Healthcare Engineering**, Hindawi Limited, v. 2019, p. 1–9, feb 2019.

ASAOKA, R. et al. Validating the usefulness of the random forests classifier to diagnose early glaucoma with optical coherence tomography. **American Journal of Ophthalmology**, Elsevier BV, v. 174, p. 95–103, feb 2017.

AUGUSTO, F. **Comitê de classificadores, Classificação multiclasse e multirrótulo**. 2020. <https://medium.com/@f2acode/comit%C3%AA-de-classificadores-classifica%C3%A7%C3%A3o-multiclasse-e-multirr%C3%B3tulo-7a01d87ee9b8>.

BHANDARI, A. **Everything you Should Know about Confusion Matrix for Machine Learning**. 2022. <https://www.analyticsvidhya.com/blog/2020/04/confusion-matrix-machine-learning/>.

BUCHANAN, B.; SHORTLIFFE, E. **Rule-Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project**. [S.l.]: Addison-Wesley Reading, 1984.

BURGANSKY-ELIASH, Z. et al. Optical coherence tomography machine learning classifiers for glaucoma detection: A preliminary study. **Investigative Ophthalmology & Visual Science**, Association for Research in Vision and Ophthalmology (ARVO), v. 46, n. 11, p. 4147, nov 2005.

CLOUD, I. **Redes neurais**. 2020. <https://www.ibm.com/br-pt/cloud/learn/neural-networks>.

CLOUD, I. **Python vs. R. Qual é a diferença?** 2021. <https://www.ibm.com/cloud/blog/python-vs-r>. Accessed: 26.06.2022.

DAS, V.; DANDAPAT, S.; BORA, P. K. Multi-scale deep feature fusion for automated classification of macular pathologies from OCT images. **Biomedical Signal Processing and Control**, Elsevier BV, v. 54, p. 101605, sep 2019.

FARSIU, S. et al. Quantitative classification of eyes with and without intermediate age-related macular degeneration using optical coherence tomography. **Ophthalmology**, Elsevier BV, v. 121, n. 1, p. 162–172, jan 2014.

GOMES, P. C. T. **Machine Learning para todos, de forma simples e com exemplos!** 2019. <https://www.datageeks.com.br/machine-learning/>. Accessed: 28.03.22.

GONZALEZ RAFAEL C.; WOODS, R. C. **Processamento Digital de Imagens**. 3^a. ed. São Paulo: Pearson Prentice Hall, 2010. ISBN 9788581435862.

GOYAL, S. et al. Review and update of cataract surgery in the diabetic eye. **Expert Review of Ophthalmology**, Informa UK Limited, v. 12, n. 5, p. 359–371, jul 2017.

HARMOUCH, M. **Algoritmo de padrão binário local: a matemática por trás disso!** 2020. <https://medium.com/swlh/local-binary-pattern-algorithm-the-math-behind-it-%EF%B8%8F-edf7b0e1c8b3>.

HEALTHLINE. **Diabetic Macular Edema: What You Need to Know**. 2020. <https://www.healthline.com/health/eye-health/diabetic-macular-edema-symptoms-treatment>. Accessed: 03.03.2022.

HORMEL, T. T. et al. Artificial intelligence in OCT angiography. **Progress in Retinal and Eye Research**, Elsevier BV, p. 100965, mar 2021.

HU, K. et al. Automatic segmentation of retinal layer boundaries in OCT images using multiscale convolutional neural network and graph search. **Neurocomputing**, Elsevier BV, v. 365, p. 302–313, nov 2019.

INFNET. **Vamos falar de Python**. 2021. <https://www.infnet.edu.br/esti/vamos-falar-de-python/:text=O%20Python%20surtiu%20na%20d%C3%A9cada,com%C3%A9dia%20brit%C3%A2ni>. Accessed: 28.03.22.

KAPOOR, R.; WALTERS, S. P.; AL-ASWAD, L. A. The current state of artificial intelligence in ophthalmology. **Survey of Ophthalmology**, Elsevier BV, v. 64, n. 2, p. 233–240, mar 2019.

KERMANY, D. S. et al. Identifying medical diagnoses and treatable diseases by image-based deep learning. **Cell**, Elsevier BV, v. 172, n. 5, p. 1122–1131.e9, feb 2018.

KUWAYAMA, S. et al. Automated detection of macular diseases by optical coherence tomography and artificial intelligence machine learning of optical coherence tomography images. **Journal of Ophthalmology**, Hindawi Limited, v. 2019, p. 1–7, apr 2019.

LEARN, S. **Model Evaluation**. 2022. https://scikit-learn.org/stable/modules/model_evaluation.html.

LEE, C. S.; BAUGHMAN, D. M.; LEE, A. Y. Deep learning is effective for classifying normal versus age-related macular degeneration OCT images. **Ophthalmology Retina**, Elsevier BV, v. 1, n. 4, p. 322–327, jul 2017.

LIMA, A. **REDES RESIDUAIS (RESNET) - APRENDIZADO PROFUNDO**. 2022. <https://acervolima.com/redes-residuais-resnet-aprendizado-profundo/:text=Bloco%20residual%3Atext=Nessa%20rede%2C%20usamos%20uma%20t%C3%A9cnica>. Accessed: 08.06.2022.

- LIN, D. H. E. A. S. C. P. Optical coherence tomography. **Science**, 1991.
- MONIZ, P. **Uma Abordagem Prática para avaliar o desempenho de algoritmos para um problema de Classificação Binária em um Projeto de Machine Learning**. 2019. [Http://rstudio-pubs-static.s3.amazonaws.com/506010_58a891968af2431486bf3dedc4949ef9.html](http://rstudio-pubs-static.s3.amazonaws.com/506010_58a891968af2431486bf3dedc4949ef9.html). Accessed: 22.05.2022.
- MOOKIAH, M. R. K. et al. Automated detection of age-related macular degeneration using empirical mode decomposition. **Knowledge-Based Systems**, Elsevier BV, v. 89, p. 654–668, nov 2015.
- MORAES, G. et al. Quantitative analysis of OCT for neovascular age-related macular degeneration using deep learning. **Ophthalmology**, Elsevier BV, v. 128, n. 5, p. 693–705, may 2021.
- MOTTA, M. M. dos S.; COBLENTZ, J.; MELO, L. G. N. de. Aspectos atuais no diagnóstico do edema macular diabético. **Revista Brasileira de Oftalmologia**, Revista Brasileira de Oftalmologia, v. 68, n. 1, p. 59–63, feb 2009.
- NEOOFTALMO. **Degeneração Macular Relacionada À Idade**. 2020. [Https://neooftalmo.com.br/doencas-oculares/degeneracao-macular-relacionada-idade](https://neooftalmo.com.br/doencas-oculares/degeneracao-macular-relacionada-idade). Accessed: 03.03.2022.
- NOLETO, C. **Variância: o que é, para que serve e como calcular?** 2022. [Https://blog.betrybe.com/variancia/](https://blog.betrybe.com/variancia/). Accessed: 10.06.2022.
- P, S. A. et al. OctNET: A lightweight CNN for retinal disease classification from optical coherence tomography images. **Computer Methods and Programs in Biomedicine**, Elsevier BV, v. 200, p. 105877, mar 2021.
- PASSOS, B. T. **O mundo do ponto de vista das Redes Neurais Convolucionais (RNCs)**. 2021. [Https://ateliware.com/blog/redes-neurais-convolucionais](https://ateliware.com/blog/redes-neurais-convolucionais). Accessed: 09.06.2022.
- PORTER, D. **What Are Drusen?** 2022. [Https://www.aao.org/eye-health/diseases/what-are-drusen](https://www.aao.org/eye-health/diseases/what-are-drusen). Accessed: 03.03.2022.
- RAVINDRA, S. **Como as Redes Neurais Convolucionais realizam o reconhecimento de imagem**. 2017. [Https://www.infoq.com/br/articles/redes-neurais-convolucionais/](https://www.infoq.com/br/articles/redes-neurais-convolucionais/). Accessed: 24.04.2022.
- ROCHA, M. **Degeneração Macular Relacionada à Idade (DMRI)**. 2021. [Https://oftalmomaster.com.br/blog/degeneracao-macular-relacionada-a-idade-dmri/](https://oftalmomaster.com.br/blog/degeneracao-macular-relacionada-a-idade-dmri/). Accessed: 03.03.2022.
- RODRIGUES, V. **Métricas de Avaliação: acurácia, precisão, recall... quais as diferenças?** 2019. [Https://vitorborbarodrigues.medium.com/metricas-de-avaliacao-acuracia-precisao-recall-quais-as-diferencas](https://vitorborbarodrigues.medium.com/metricas-de-avaliacao-acuracia-precisao-recall-quais-as-diferencas).
- THOMAS, A. et al. A novel multiscale and multipath convolutional neural network based age-related macular degeneration detection using OCT images. **Computer Methods and Programs in Biomedicine**, Elsevier BV, v. 209, p. 106294, sep 2021.

THOMAS, A. et al. RPE layer detection and baseline estimation using statistical methods and randomization for classification of AMD from retinal OCT. **Computer Methods and Programs in Biomedicine**, Elsevier BV, v. 200, p. 105822, mar 2021.

TSUJI, T. et al. Classification of optical coherence tomography images using a capsule network. **BMC Ophthalmology**, Springer Science and Business Media LLC, v. 20, n. 1, mar 2020.

VENKATESH, R.; BAVAHARAN, B.; YADAV, N. K. Predictors for choroidal neovascular membrane formation and visual outcome following blunt ocular trauma. **Therapeutic Advances in Ophthalmology**, SAGE Publications, v. 11, p. 251584141985201, jan 2019.

VIEIRA, S. **O que é assimetria (ou distorção) e como se mede?** 2018. [Http://soniavieira.blogspot.com/2018/05/por-conta-de-umapergunta-sobre-questao3.html](http://soniavieira.blogspot.com/2018/05/por-conta-de-umapergunta-sobre-questao3.html). *Acessed* : 22.05.2022.

VIEIRA, S. **O que é curtose e como se mede?** 2018. [Http://soniavieira.blogspot.com/2018/11/curtose.html](http://soniavieira.blogspot.com/2018/11/curtose.html). *Acessed*: 22.05.2022.