



Universidade Federal de Uberlândia
FACULDADE DE ENGENHARIA ELÉTRICA

JUCELINO CARDOSO MARCIANO DOS SANTOS

**TÉCNICAS DE APRENDIZADO DE
MÁQUINAS NA SEGMENTAÇÃO
AUTOMÁTICA DE VASOS SANGUÍNEOS DA
RETINA**

Uberlândia, Fevereiro de 2022

Universidade Federal de Uberlândia
FACULDADE DE ENGENHARIA ELÉTRICA

JUCELINO CARDOSO MARCIANO DOS SANTOS

**TÉCNICAS DE APRENDIZADO DE
MÁQUINAS NA SEGMENTAÇÃO
AUTOMÁTICA DE VASOS SANGUÍNEOS DA
RETINA**

Tese apresentada à Faculdade de Engenharia Elétrica da Universidade Federal de Uberlândia como parte das exigências para obtenção do título de Doutor em Ciências.

Orientador(a):
Gilberto Arantes Carrijo

Uberlândia, Fevereiro de 2022

Ficha Catalográfica Online do Sistema de Bibliotecas da UFU
com dados informados pelo(a) próprio(a) autor(a).

S237
2022 Santos, Jucelino Cardoso Marciano dos, 1979-
Técnicas de Aprendizado de Máquinas na Segmentação
Automática de Vasos Sanguíneos da Retina [recurso
eletrônico] / Jucelino Cardoso Marciano dos Santos. -
2022.

Orientador: Gilberto Arantes Carrijo.
Tese (Doutorado) - Universidade Federal de Uberlândia,
Pós-graduação em Engenharia Elétrica.
Modo de acesso: Internet.
Disponível em: <http://doi.org/10.14393/ufu.te.2022.129>
Inclui bibliografia.
Inclui ilustrações.

1. Engenharia elétrica. I. Carrijo, Gilberto Arantes,
1948-, (Orient.). II. Universidade Federal de
Uberlândia. Pós-graduação em Engenharia Elétrica. III.
Título.

CDU: 621.3

Bibliotecários responsáveis pela estrutura de acordo com o AACR2:

Gizele Cristine Nunes do Couto - CRB6/2091

JUCELINO CARDOSO MARCIANO DOS SANTOS

**TÉCNICAS DE APRENDIZADO DE
MÁQUINAS NA SEGMENTAÇÃO
AUTOMÁTICA DE VASOS SANGUÍNEOS DA
RETINA**

Tese defendida por Jucelino Cardoso Marciano dos Santos e aprovada em 18 de fevereiro de 2022, pela banca examinadora constituída pelos membros:

PhD. Gilberto Arantes Carrijo
Orientador

PhD. Júlio César Ferreira
Avaliador

Dr. Monica Sakuray Pais
Avaliador

Dr. Antônio Claudio Paschoarelli Veiga
Avaliador

Dra. Milena Bueno Pereira Carneiro
Avaliador

Uberlândia, 18 de fevereiro de 2022



UNIVERSIDADE FEDERAL DE UBERLÂNDIA
 Coordenação do Programa de Pós-Graduação em Engenharia Elétrica
 Av. João Naves de Ávila, 2121, Bloco 3N - Bairro Santa Mônica, Uberlândia-MG, CEP 38400-902
 Telefone: (34) 3239-4707 - www.posgrad.feelt.ufu.br - copel@ufu.br



ATA DE DEFESA - PÓS-GRADUAÇÃO

Programa de Pós-Graduação em:	Engenharia Elétrica				
Defesa de:	Tese de Doutorado, 296, PPGEELT				
Data:	Dezoito de fevereiro de dois mil e vinte e dois	Hora de início:	09:00	Hora de encerramento:	12:00
Matrícula do Discente:	11813EEL004				
Nome do Discente:	Jucelino Cardoso Marciano dos Santos				
Título do Trabalho:	Técnicas de aprendizado de máquinas na segmentação automática dos vasos sanguíneos da retina				
Área de concentração:	Processamento da informação				
Linha de pesquisa:	Processamento digital de sinais				
Projeto de Pesquisa de vinculação:	Coordenador do projeto: Gilberto Arantes Carrijo. Título do projeto: Segmentação de imagens aplicadas a imagens biomédicas. Agência financiadora: Número do processo na agência financiadora: Vigência do projeto: 2017- atual.				

Reuniu-se por meio de videoconferência, a Banca Examinadora, designada pelo Colegiado do Programa de Pós-graduação em Engenharia Elétrica, assim composta: Professores Doutores: Antônio Claudio Paschoarelli Viegas - FEELT/UFU; Milena Bueno Pereira Carneiro - FEELT/UFU; Monica Sakuray Pais - IFGOIANO; Júlio César Ferreira - IFGOIANO; Gilberto Arantes Carrijo - FEELT/UFU, orientador(a) do(a) candidato(a).

Iniciando os trabalhos o(a) presidente da mesa, Dr(a). Gilberto Arantes Carrijo, apresentou a Comissão Examinadora e o candidato(a), agradeceu a presença do público, e concedeu ao Discente a palavra para a exposição do seu trabalho. A duração da apresentação do Discente e o tempo de arguição e resposta foram conforme as normas do Programa.

A seguir o senhor(a) presidente concedeu a palavra, pela ordem sucessivamente, aos(às) examinadores(as), que passaram a arguir o(a) candidato(a). Ultimada a arguição, que se desenvolveu dentro dos termos regimentais, a Banca, em sessão secreta, atribuiu o resultado final, considerando o(a) candidato(a):

Aprovado(a).

Esta defesa faz parte dos requisitos necessários à obtenção do título de Doutor.

O competente diploma será expedido após cumprimento dos demais requisitos, conforme as normas do Programa, a legislação pertinente e a regulamentação interna da UFU.

Nada mais havendo a tratar foram encerrados os trabalhos. Foi lavrada a presente ata que após lida e achada conforme foi assinada pela Banca Examinadora.



Documento assinado eletronicamente por **Mônica Sakuray Pais, Usuário Externo**, em 18/02/2022, às 11:53, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Júlio César Ferreira, Usuário Externo**, em 18/02/2022, às 11:53, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Milena Bueno Pereira Carneiro, Professor(a) do Magistério Superior**, em 18/02/2022, às 11:54, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Antonio Claudio Paschoarelli Veiga, Professor(a) do Magistério Superior**, em 18/02/2022, às 11:55, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Gilberto Arantes Carrijo, Professor(a) do Magistério Superior**, em 18/02/2022, às 12:43, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



A autenticidade deste documento pode ser conferida no site

[https://www.sei.ufu.br/sei/controlador_externo.php?](https://www.sei.ufu.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0)

[acao=documento_conferir&id_orgao_acesso_externo=0](https://www.sei.ufu.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0), informando o código verificador **3386709** e o código CRC **9A2A444E**.

*Dedico este trabalho aos meus Pais, Maria de Fátima Marciano, e Juracy
Cardoso de Carvalho em sua memória.
Dedico também aos meus filhos Laura, Arthur, Heloisa e Henry.*

AGRADECIMENTOS

Agradeço a todos que contribuíram direta ou indiretamente para a realização deste trabalho. Em especial, a minha mãe Maria de Fátima Marciano e minha sogra Ormezida pelo cuidado com os netos, durante a ausência para cursar as disciplinas do doutorado.

Agradeço em especial ao amigo Pedro Moises de Sousa pelo companheirismo em todos os projetos e trabalhos que desenvolvemos ao longo de quatro anos.

Em especial gostaria de agradecer ao Professor Júlio Cesar Ferreira pelas orientações e ensinamentos e contribuições ao longo desta jornada.

Agradeço ao meu estimado orientador professor Gilberto Arantes Carrijo, pela confiança depositada no meu trabalho e por lições valiosas que me acompanharam por toda a vida.

Agradeço aos professores: Antônio Cláudio Paschoarelli Veiga, Júlio César Ferreira, Mônica Sakuray e Milena Bueno Pereira Carneiro pelas valiosas contribuições.

Agradeço aos amigos Davidson e Izabela pelos momentos de descontração e busca pela sanidade.

Agradeço em especial a minha companheira e parceira de todas as horas, Cristiane de Fátima dos Santos Cardoso.

Por fim, agradeço ao povo brasileiro e a Deus.

As imagens de retinografias, ou imagens de fundo de olho, são obtidas por meio da pupila e fornecem uma representação direta do sistema nervoso e vascular do corpo humano, são facilmente alteradas na presença de patologias diversas, por isso o grande interesse em imagens de retinografia. Doenças como lúpus, diabetes, zika, glaucoma, dentre outras produzem significativas alterações das estruturas anatômicas. Neste sentido, marcos anatômicos como o disco óptico, a mácula e vasos sanguíneos são marcadores importantes nos exames oftalmológicos. Consequentemente, muitos especialistas aplicam técnicas de processamento digital de imagens com objetivo de analisar, localizar e identificar tais estruturas. Em particular a segmentação dos vasos sanguíneos é um dos elementos chaves para compor sistemas de triagem ou diagnóstico automáticos. Por exemplo, a detecção de hemorragias presentes na estrutura vascular da retina pode indicar uma enfermidade como retinopatia diabética, doenças cardiovasculares entre outras. Diante disso, o objetivo geral deste trabalho é desenvolver e avaliar estratégias baseadas em *Artificial Neural Network* (**ANN**) e *Convolutional Neural Network* (**CNN**), a fim de aumentar o desempenho em segmentação de imagens de fundo de olho com ênfase nos vasos sanguíneos em termos das métricas de avaliação sensibilidade, especificidade e precisão utilizadas em imagens médicas. São propostas duas abordagens para segmentar os vasos sanguíneos em imagens de fundo de olho, de forma que as contribuições podem ser listadas como: (i) adaptação de um modelo *Multilayer Perceptron* (**MLP**) para a detectar os vasos sanguíneos, com resultados equivalentes ao estado da arte e (ii) adaptação do modelo *Convolutional Neural Network* (**CNN**), arquitetura U-Net, para segmentação dos vasos sanguíneos, com diferencial de utilizar todas as informações do modelo de cores rgb. A metodologia proposta foi avaliada nos bancos de dados publicamente disponíveis: *Digital Retinal Images for Vessel Extraction* (**DRIVE**) e *Structured Analysis of the Retina* (**STARE**). As arquiteturas **MLP** e **CNN** U-Net forneceram resultados expressivos em relação ao estado da arte, com medidas de precisão de 0.9487 e 0.9715 no banco de dados **DRIVE**.

Palavras-chave:

Imagens de fundo de olho. Pré-processamento digital de imagem. Segmentação dos vasos sanguíneos. Artificial Neural Network. Convolutional Neural Network.

The images of retinographies, or images of the fundus of the eye, are obtained through the pupil and provide a direct representation of the nervous and vascular system of the human body, they are easily altered in the presence of different pathologies, hence the great interest in retinography images. Diseases such as lupus, diabetes, Zika, glaucoma, among others, produce significant changes in anatomical structures. In this sense, anatomical landmarks such as the optic disc, the macula and blood vessels are important markers in ophthalmologies exams. Consequently, many specialists apply digital image processing techniques in order to analyze, locate and identify such structures. In particular, the segmentation of blood vessels is one of the key elements to compose automatic screening or diagnosis systems. For example, the detection of hemorrhages present in the vascular structure of the retina can indicate a disease such as diabetic retinopathy, cardiovascular diseases, among others. Therefore, the general objective of this work is to develop and evaluate strategies based on *Artificial Neural Network* (ANN) and *Convolutional Neural Network* (CNN), in order to increase the performance in fundus image segmentation with an emphasis on blood vessels in terms of sensitivity, specificity and precision assessment metrics used in medical imaging. Two approaches are proposed to segment blood vessels in fundus images, so that the contributions can be listed as: (i) adaptation of a model *Multilayer Perceptron* (MLP) to detect the blood vessels, with results equivalent to the state of the art and (ii) adaptation of the CNN model, UNet architecture, for blood vessel segmentation, with the differential of using all the information of the RGB color model. The proposed methodology was evaluated in the publicly available databases: *Digital Retinal Images for Vessel Extraction* (DRIVE) and *Structured Analysis of the Retina* (STARE). The MLP and CNN UNet architectures provided expressive results in relation to the state of the art, with precision measurements of 0.9487 and 0.9713 in the DRIVE database.

Key-words:

Background images. Digital image pre-processing. Segmentation of blood vessels. Artificial Neural Network. Convolutional Neural Network.

1.1	Processo de Segmentação no banco de imagens DRIVE. (a) Imagem original, (b) Segmentação do primeiro especialista, (c) Segmentação do segundo especialista.	20
2.1	Exemplo de corte longitudinal das estruturas do olho	25
2.2	Processo de aquisição da imagem da retina. (a) Aparelho denominado de Retinógrafo digital. (b) Imagem da retina adquirida do retinógrafo	27
2.3	Exemplos de imagens Normais e patológicas. (a) Imagem normal - DRIVE. (b) Imagem Patológica - DRIVE. (c) Imagem normal - STARE. (d) Imagem patológica - STARE	28
2.4	Sintomas visuais causados pelo glaucoma	30
2.5	Sintomas anatômicos na retina. (a) Os principais elementos da retina, vasos sanguíneos, disco óptico e mácula. (b) presença de microaneurismas. (c) exsudatos moles. (d) exsudatos duros. (e) combinação de exsudatos duros e hemorragias. (f) exsudatos duros afetando a mácula.	30
2.6	Sintomas anatômicos na retina. (a) Drusas grandes ou seja manchas sub-retinianas amareladas. (b) Progressão das drusas depois de cinco anos. (c)-(d) Degeneração macular neovascular relacionada a idade. (e) Desenvolvimento de fibrose densa sub-retiniana não tratada após um ano (região tracejada). (f) Desenvolvimento de áreas atroficas ao redor do novo vaso sanguíneo após três anos de não tratamento (região tracejada).	31
2.7	Aprendizado de máquina como um novo paradigma de programação	33
2.8	Diagrama geral mostrando Inteligência Artificial e seus subcampos	34
2.9	Estrutura e fluxo do MLP. (a) Estrutura de um neurônio biológico. (b) Estrutura do neurônio matemático com seta para direita indicando o fluxo de processamento dos dados.	35
2.10	Diagrama intuitivo sobre o processo de divisão do conjunto de dados	37
2.11	Composições de Arquiteturas. (a) ANN arquitetura básica. (b) ANN arquitetura profunda	39
2.12	Multilayer perceptron. (a) MLP com duas camadas de profundidade. (b) Fluxo das informações.	40
2.13	Exemplo do processo de otimização. (a) Função de duas variáveis. (b) Gráfico da descida do gradiente	44
2.14	Exemplos da aplicação do Gradiente descendente. (a), (b) e (c) Diferentes valores para α e $iter$ fixo em 50.	45

2.15	Arquitetura proposta no trabalho de Alex Khrizhevsky	52
2.16	Arquitetura Unet	53
2.17	Exemplo da operação de convolução	54
2.18	Convolução entre a imagem e o <i>kernel</i> resultando em uma mapa de características. Na primeira linha, a imagem sem <i>padding</i> e a imagem resultante com dimensão reduzida. Na segunda linha, a imagem com preenchimento com <i>padding</i> e a imagem resultante com dimensão aumentada	56
2.19	Convolução entre a imagem e o <i>kernel</i> resultando em uma mapa de características. Na primeira linha, o padding usando para imagem é $P = 1$, $k = 3$ e $S = 1$, a imagem resultante da convolução têm a mesma dimensão da original. Na segunda linha, o padding $P = 2$, $k = 3$ e $S = 1$, a imagem resultante da covolução têm dimensão maior que a original.	57
2.20	Estratégias de <i>Pooling</i> propostas para reduzir o conjunto de informações	58
2.21	Efeito do <i>Dropout</i> na rede neural	59
2.22	Aplicação das métricas uma imagem de fundo de olho.	60
3.1	Diagrama - Diagrama geral para classificação dos métodos que segmentam os vasos sanguíneos	63
3.2	Visão geral das abordagens que utilizam algoritmos não supervisionados para tarefa de semgntação de imagens de fundo de olho	67
3.3	Visão geral das abordagens que utilizam algoritmos supervisionados para tarefa de segmentação de imagens de fundo de olho	73
4.1	Principais causas da cegueira	79
4.2	Porcentagem de todas as mortes atribuíveis a níveis elevados de glicose no sangue para adultos com idade entre 20 e 69 anos, por região e sexo, nos anos de 2000 e 2012	79
4.3	Imagem colorida de fundo com estruturas anatômicas	80
4.4	Metodologia proposta para ajustar os parâmetros e aprimorar as imagens de fundo do olho	91
4.5	Metodologia proposta para segmentação dos vasos sanguíneos	92
4.6	Metodologia proposta para aprimoramento da imagem	93
4.7	Aprimoramento na imagem "test". (a) Canal verde original da imagem. (b) Imagem após o filtro de <i>Wiener</i> e fitro de abertura. (c) Imagem depois do CLAHE (d) Imagem depois do filtro da média	96
4.8	Resultados usando o parâmetro α	98
4.9	Resultado da segmentação dos vasos sanguíneos	101
5.1	Metodologia geral para segmentação dos vasos sanguíneos	113
5.2	Diagrama da fase de pré-processamento	116
5.3	Divisão do <i>ground thruth</i> em blocos	118
5.4	Arquitetura da CNN Unet utilizada	118
5.5	Esquema de segmentação da CNN Unet	119
5.6	Esquemas de entrada e saída da rede Unet proposta	119
5.7	Análise dos parâmetros CLAHE no processo de melhoria do contraste da imagem da retina	121
5.8	Análise dos parâmetros do filtro Wiener no processo de atenuação dos ruídos presentes na imagem da retina	121

5.9	Resultado do processo de aprimoramento da imagem usando CLAHE e filtro <i>Wiener</i>	122
5.10	Análise dos algoritmos de otimização <i>Adam</i> e <i>RMSPProp</i>	122
5.11	Comparação da segmentação utilizando a CNN Unet proposta neste trabalho	123
5.12	Aplicação da CNN Unet treinada na imagem de test 03-manual	123

LISTA DE TABELAS

2.1	Exemplos de algumas sinais oculares em imagens retinianas	29
2.2	Funções de Ativação	42
2.3	Exemplos da operação de convolução utilizando as fórmulas definidas nas equações 2.23 e 2.24	57
2.4	Anotações realizadas por um especialista.	60
4.1	Resumo das informações DRIVE e STARE	86
4.2	ANNs selecionadas na fase experimental. O termo $MLP_{(16-32-16)}$ significa 16 neurônios na camada de entrada, 32 neurônios na camada oculta, e 16 neurônios na camada de saída.	90
4.3	Parâmetros e valores para os filtros <i>Wiener</i> , CLAHE e Média	92
4.4	Análise de potência de ruído do filtro <i>Wiener</i> por métrica Acc	97
4.5	Análise de tamanho do bloco <i>Wiener</i> por medida Acc	97
4.6	Análise de distribuição CLAHE por medida Acc	97
4.7	Análise de bloco CLAHE por medida Acc	98
4.8	Análise de Cliplimit CLAHE por medida Acc	98
4.9	Análise do parâmetro α CLAHE por medida Acc	98
4.10	Análise CLAHE nBins por medida Acc	99
4.11	Análise do tamanho médio do bloco de filtro por medida Acc	99
4.12	Comparação do método proposto na base de dados DRIVE com métodos do estado da arte.	100
4.13	Tabela com as métricas de avaliação	100
4.14	<i>Accuracy</i> em função do número de neurônios na camada oculta.	102
5.1	Tabela de Métodos Supervisionados	111
5.2	Tabela de Métodos não supervisionados	112
5.3	Aumento do bando de dados	114
5.4	Número de paths para treinamento e teste	114
5.5	Parâmetros para os filtros <i>Wiener</i> e CLAHE	116
5.6	Comparação do método proposto com métodos supervisionados e não supervisionados.	124

LISTA DE ABREVIATURAS E SIGLAS

- Acc** *Accuracy.* Pag.59
- AdaDelta** *Adaptive Learning Rate.* Pag.46
- AdaGrad** *Adaptive Gradient.* Pag.46
- Adam** *Adaptive Moment Estimation.* Pag.46
- AHE** *Adaptive Histogram Equalization.* Pag.79
- AI** *Artificial Intelligence.* Pag.31
- ANN** *Artificial Neural Network.* Pag.7, 8
- ARMD** *Age-Related Macular Degeneration.* Pag.29
- AUC** *Area Under the ROC Curve.* Pag.59
- CAD** *Computer Aided Diagnosis.* Pag.16
- CCWT** *Complex Continuous Wavelet Transform.* Pag.67
- CLAHE** *Contrast Limited Adaptive Histogram Equalization.* Pag.21
- CNN** *Convolutional Neural Network.* Pag.7, 8
- DAEs** *De-noising Auto-Encoders.* Pag.73
- DRIVE** *Digital Retinal Images for Vessel Extraction.* Pag.7, 8
- DSM** *Dice Similarity Measure.* Pag.59
- F** *F-score.* Pag.59
- FN** *False Negative.* Pag.58
- FOV** *Field-Of-View.* Pag.19
- FP** *False Positive.* Pag.58
- FPR** *False Positive Rate.* Pag.59

GAN *Generative adversarial Networks.* Pag.72
HE *Histogram Equalization.* Pag.79
IAPB *International Agency for Prevention of Blindness.* Pag.25
MDCNN *Multiple Deep Convolutional Neural Network.* Pag.106
MLP *Multilayer Perceptron.* Pag.7, 8
MPMF *Multiscale Production of the Matched Filter.* Pag.66
MRF *Markov Random Field.* Pag.64
NE *noise estimate.* Pag.80
PACG *Primary Angle-Closure Glaucoma.* Pag.27
PCA *Principal Analysis Component.* Pag.70
POAG *Primary Open-Angle Glaucoma.* Pag.27
PSNR *Peak Signal to Noise Ratio.* Pag.80
RF *Randon Florest.* Pag.73
ROC *Receiver Operating Characteristic.* Pag.59
S *Overlapping score.* Pag.59
Se *Sensitivity.* Pag.58
SGD *Stochastic Gradient Descent.* Pag.45
SNR *signal-to-noise ratio.* Pag.80
Sp *Especificity.* Pag.58
SSIM *Structural Similarity Index.* Pag.80
STARE *Structured Analysis of the Retina.* Pag.7, 8
SVM *Support Vector Machine.* Pag.66
TN *True Negative.* Pag.58
TP *True Positive.* Pag.58
TPR *True Positive Rate.* Pag.59
WHO *World Health Organization.* Pag.17

I	Background	17
1	INTRODUÇÃO	18
1.1	Justificativa	19
1.2	Objetivo	22
1.2.1	Objetivo geral	22
1.2.2	Objetivos Específicos	22
1.3	Contribuições	23
1.4	Estrutura da tese	23
2	CONCEITOS BÁSICOS	24
2.1	Anatomia da Retina	24
2.1.1	Retinografia Digitais	26
2.1.2	Enfermidades oculares	27
2.1.3	Análise de imagens médicas	32
2.2	Inteligência Artificial e Redes Neurais Artificiais	33
2.2.1	Aprendizado Supervisionado	36
2.2.2	Perceptron Múltiplas Camadas	40
2.2.3	Funções de Ativação	41
2.2.4	Normalização de camadas	42
2.3	Técnicas de Otimização	43
2.3.1	Gradiente Descendente	43
2.3.2	Momentum	47
2.3.3	Gradiente Descente Estocástico	47
2.3.4	AdaGrad e RMSProp	49
2.3.5	Adam	50
2.4	Redes Neurais Convolucionais	51
2.4.1	Convolução	54
2.4.2	Pooling Layer	57
2.4.3	Dropout	58
2.4.4	Função Softmax	58
2.5	Métricas de Avaliação	59
2.6	Considerações finais deste capítulo	62

3	Trabalhos Relacionados	63
3.1	Metodologia Baseada em Rastreamento	64
3.2	Técnicas Baseadas em Modelo	65
3.3	Abordagem Híbridas	66
3.4	Abordagem de Reconhecimento de Padrões não supervisionada	67
3.4.1	Técnicas baseadas em Multiescala	67
3.4.2	Morfologia Matemática	69
3.4.3	Abordagem de Filtragem combinada	70
3.4.4	Técnica baseadas em crescimento de região	71
3.5	Abordagem de Reconhecimento de Padrões supervisionada	73
3.5.1	Técnicas Baseadas em Redes neurais	74
3.5.2	Técnicas baseadas em Classificação de Conjunto	75
3.6	Considerações finais deste capítulo	76
II	Contribuições	77
4	Aprimoramento da qualidade da imagem do fundo de olho para detecção de vasos sanguíneos por meio de uma rede neural usando filtro CLAHE e Wiener	78
4.1	Trabalhos Relacionados	82
4.2	Métodos	85
4.2.1	Banco de dados	85
4.2.2	Filtro de Wiener	86
4.2.3	CLAHE	87
4.2.4	Experimentos realizados	89
4.3	Aprimoramento de imagem para detecção dos vasos sanguíneos usando uma rede neural baseada em CLAHE e o filtro Wiener	91
4.3.1	Pré-processamento	93
4.3.2	Treinamento MLP	94
4.3.3	Pós-processamento	95
4.3.4	Saída	95
4.4	Resultados e Discussões	96
4.5	Considerações Finais	104
5	Segmentação Automática dos vasos sanguíneos combinando Filtragem de Domínio Duplo e uma Arquitetura CNN Unet reduzida.	105
5.1	Trabalhos Relacionados	108
5.2	Métodos	112
5.2.1	Banco de Dados	113
5.2.2	Pré-Processamento	114
5.2.3	Filtros	115
5.2.4	Esquema do pré-processamento	116
5.2.5	CNN Unet	117
5.3	Experimentos	119
5.4	Resultados e Discussões	120
5.5	Considerações Finais	124
	CONCLUSÃO	125

Parte I

Background

CAPÍTULO 1

INTRODUÇÃO

O conceito de transformação digital tem influenciado diversos setores da sociedade moderna, em especial a área da medicina (ABRÀMOFF et al., 2010). Com o desenvolvimento das técnicas de processamento digital de imagens e o aumento da capacidade computacional, a construção de sistemas inteligentes é uma realidade possível (KRIZHEVSKY et al., 2012). No contexto da medicina, os sistemas inteligentes são definidos como sistemas de Diagnóstico Auxilado por Computador (ou o termo em inglês CAD (*Computer Aided Diagnosis*)) e estão em crescente desenvolvimento (ZHANG et al., 2014). O principal objetivo de um CAD é acelerar os procedimentos de diagnósticos e atuar em larga escala tornando as etapas de triagem e monitoramento dinâmica (MA; LI, 2015). Portanto, diagnosticar precocemente uma enfermidade como retinopatia diabética ou glaucoma, significa melhoria na qualidade de vida do paciente, aumento da probabilidade de cura, e economia de recursos financeiros.

Neste sentido, o uso de imagens digitais em diferentes tipos de exames é uma estratégia que ganhou destaque no mundo moderno desde a invenção das câmeras (GONZALEZ; WOODS, 2009). O desenvolvimento tecnológico das últimas décadas influenciou de maneira positiva as formas de triagem e diagnóstico na medicina, em especial a realização de exames por meio de imagens digitais (RAMANI et al., 2013). Assim, o sistema circulatório composto por vasos sanguíneos é um dos sistemas bem investigados (FRAZ et al., 2012c). De fato, o funcionamento de todo o corpo humano depende do bom funcionamento dos diferentes vasos sanguíneos, por exemplo, artérias pulmonares e coronárias, artérias cerebrais, vasos retinianos (DORION, 2002). Desta forma, as alterações anormais

nos vasos sanguíneos podem ser indicadores de diferentes enfermidades como doenças cardiovasculares, hipertensão, acidente vascular cerebral, retinopatia diabética, glaucoma entre outros (ABRÀMOFF et al., 2010).

A proporção de enfermidades conhecidas tem aumentado significativamente nas últimas décadas, em particular, enfermidades como retinopatia diabética e glaucoma (FRAZ et al., 2012a). O relatório mais recente da Organização Mundial da Saúde (*World Health Organization* (WHO)) revela que 2.2 bilhões de pessoas têm deficiência visual e 50% dos casos poderiam ser evitados e tratados com um diagnóstico preventivo (World Health Organization, 2019). Em especial, a segmentação dos vasos sanguíneos permite avaliar as alterações nas principais veias e artérias (FRAZ et al., 2012a). Na medicina moderna, a segmentação pode ser realizada por um médico especialista ou realizada por um algoritmo computacional (ABDULLAH et al., 2016).

1.1 Justificativa

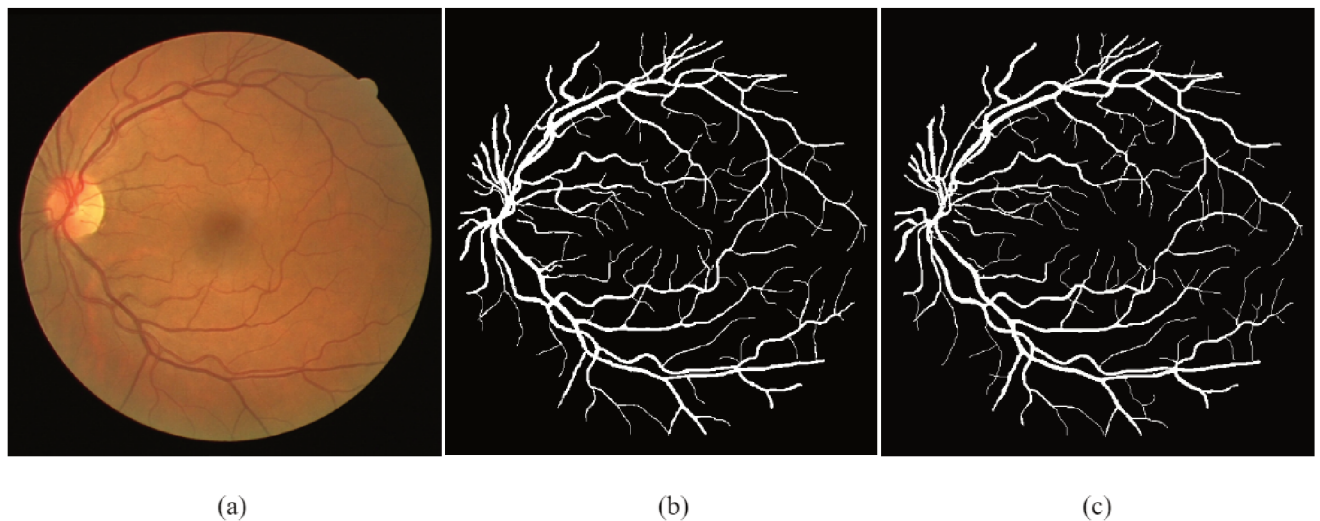
A descoberta precoce das enfermidades como retinopatia diabética e glaucoma são consideradas etapas cruciais para reduzir a ocorrência da morbidade visual (ZAKI et al., 2016). Contudo, em grande parte dos países subdesenvolvidos as opções de análise da retina são escassas ou inexistente (World Health Organization, 2019). Além disso, o procedimento clássico de avaliar as imagens da retina por métodos manuais são caros e trabalhosos (KUMAR et al., 2012), (ZAKI et al., 2016). A Figura 1.1 apresenta uma imagem de fundo de olho e a segmentação realizada por dois especialistas (STAAL et al., 2004). Muitos métodos surgiram ao longo do anos com objetivo de fornecer algum tipo de solução para segmentação dos vasos sanguíneos, como: (NEKOVEI; SUN, 1995), (AL-RAWI; KARAJEH, 2007), (FARNELL et al., 2008), (ASAD et al., 2013), (JEBASEELI et al., 2019), (KARAALI et al., 2021), (YAN et al., 2019), (HASHEMZADEH; Adlpour Azar, 2019).

Um dos primeiros trabalhos relacionado à detecção dos vasos sanguíneos foi proposto por (NEKOVEI; SUN, 1995). Nesse trabalho, as imagens de angiograma¹ são fornecidas a uma Artificial Neural Network (ANN) com abordagem *back-propagation* treinada para detectar a estrutura vascular. Em outra linha de pesquisa, a detecção dos vasos sanguíneos

¹angiograma ou arteriografia, é a observação da anatomia do coração e dos vasos sanguíneos, por meio de radiografia, após a administração intravascular de meio de contraste radiopaco (tintura) (ABRÀMOFF et al., 2010).

é realizada por meio de algoritmos genéticos e *matched filter* (AL-RAWI; KARAJEH, 2007). Em geral, o filtro pode ser ajustado de diferentes maneiras, alterando seus parâmetros, o que reflete diretamente na sua resposta. Assim, os autores utilizaram a *Area Under the Curve* como *fitness function* obtendo bons resultados (FARNELL et al., 2008). Em outra linha de investigação, o trabalho de (ASAD et al., 2013) segmenta os vasos sanguíneos por meio de um método não supervisionado. A segmentação dos vasos é composta por três fases distintas. Na fase um, as técnicas *wMean* e *gMean* são utilizadas para melhorar o brilho. Na fase dois, os vasos sanguíneos são segmentados usando um algoritmo baseada *water flooding*. Por fim, uma fase de pós-processamento visa melhorar os resultados da fase dois e determinam a segmentação final da estrutura vascular (ASAD et al., 2013).

Figura 1.1: Processo de Segmentação no banco de imagens DRIVE. (a) Imagem original, (b) Segmentação do primeiro especialista, (c) Segmentação do segundo especialista.



Fonte: Adaptado da (STAAL et al., 2004)

Em outra perspectiva, o trabalho de (JEBASEELI et al., 2019) realizou a segmentação automática dos vasos sanguíneos utilizando duas estratégias inovadoras. A primeira é uma arquitetura de rede neural acoplada denominada como *Tandem Pulse Coupled Neural Network* (TPCNN). Esta arquitetura segmenta os vasos sanguíneos da retina usando imagens patológicas inclusive os vasos menores. A segunda, uma *Deep Learning Based Support Vector Machine* (DLBSVM) melhora a precisão da classificação dos pixels que pertencem ao vaso e os que não pertencem. Em (KARALI et al., 2021) é apresentado um método automático, baseado em Convolutional Neural Network (CNN) para segmentar os vasos sanguíneos. A estratégia considera uma arquitetura de decodificador multi-codificador, composta por duas unidades codificadoras com camadas convolucionais

e *maxpooling*, e uma decodificadora com camadas convolucionais e deconvolucionais. Nesta abordagem a imagem de retina RGB é inserida diretamente como entrada, diferente das técnicas citadas acima que utilizam somente um canal. Na mesma linha de algoritmos de aprendizado de máquinas (YAN et al., 2019) propôs uma estrutura geral em três estágios: i) o *ThickSegmenter* segmenta os vasos sanguíneos mais espessos; ii) o *ThinSegmenter* segmenta os vasos sanguíneos mais finos e iii) o *FusionSegmenter* faz a fusão dos dois estágios anteriores. Em uma estratégia semelhante (HASHEMZADEH; Adlpour Azar, 2019) propôs a segmentação dos vasos sanguíneos combinando métodos supervisionado e não supervisionados. O algoritmo proposto é dividido em três etapas. Na primeira, o pré-processamento das imagens da retina como: o corte da Field-Of-View (FOV), a transformação no espaço de cores e aprimoramento do contraste. Na segunda, a extração do vasos utilizando métodos para extrair referências e orientações. Por fim, um vetor de características fornece as informações necessárias para o processo final de classificação.

Assim a segmentação constitui uma das etapas com maior relevância no processamento de imagens digitais (GONZALEZ; WOODS, 2009). Desta forma, o conceito primordial de segmentar uma imagem, significa dividi-la em diversas partes sem sobreposição (ZHANG et al., 2017). No entanto, o nível de detalhe da subdivisão na imagem depende do problema a ser resolvido, ou seja, a segmentação deve parar quando as regiões de interesse são detectadas (AGGARWAL et al., 2018). Grande parte dos algoritmos de segmentação baseia-se em propriedades de intensidade, como descontinuidade e similaridade. As técnicas baseadas em descontinuidade dividem a imagem com base nas mudanças abruptas de intensidade, por exemplo mudança na intensidade dos *pixels* quando caminhamos no sentido da borda da imagem (GONZALEZ; WOODS, 2009). Em outro sentido, as técnicas baseadas em similaridade dividem a imagem a partir de um conjunto de regras pré-definidas agrupando regiões que são semelhantes (PEDRINI; SCHWARTZ, 2007). Apesar de todos os avanços obtidos na tarefa de segmentação dos vasos sanguíneos, a construção de um sistema automatizado ainda é um grande desafio. Os sistemas CAD atuais que utilizam imagens com objetivo de diagnosticar alguma enfermidade, geralmente extraem algumas características relevantes, como disco óptico, fóvea, mácula e os vasos sanguíneos da retina (FAZAL et al., 2018). Contudo, a grande dificuldade é determinar com precisão quais as características devem ser usadas na etapa de classificação.

Além disso, existem condições adversas que afetam a segmentação dos vasos san-

guíneos em imagens de fundo de olho (ABRÀMOFF et al., 2010). Desta forma, os fatores considerados agravantes para segmentação podem ser listadas em: reflexo da luz central dos vasos, junções de baixo contraste, cruzamento de vasos estreitos, inclusão de vasos estreitos e paralelos, presença de patologias, contraste variável, ruído de fundo e falsos pixels dos vasos (SOOMRO et al., 2019). Embora, a segmentação dos vasos sanguíneos seja um problema bem investigado para imagens da retina, não existem sistemas que sejam capazes de lidar com todas as dificuldades listadas anteriormente (ALMOTIRI et al., 2018). Neste sentido, este trabalho investigou se as estratégias baseadas em ANN e CNN são capazes de aumentar o desempenho em segmentação de imagens de fundo de olho, com ênfase nos vasos sanguíneos em termos de métricas de avaliação, como sensibilidade, especificidade e precisão utilizadas em imagens médicas.

1.2 Objetivo

1.2.1 Objetivo geral

O objetivo geral deste trabalho é desenvolver e avaliar estratégias baseadas em ANN e CNN, a fim de aumentar o desempenho em segmentação de imagens de fundo de olho com ênfase nos vasos sanguíneos em termos das métricas de avaliação, como sensibilidade, especificidade e precisão utilizadas em imagens médicas.

1.2.2 Objetivos Específicos

Os objetivos específicos são:

- i) Estudar e implementar algoritmos que sejam capazes de reduzir o nível de ruído e preservar as principais informações da imagem;
- ii) Estudar e implementar arquiteturas ANN e CNN;
- iii) Desenvolver e avaliar estratégias baseadas em ANN e CNN capazes de aumentar o desempenho da tarefa de segmentação;
- iv) Comparar os métodos desenvolvidos com o estado da arte e métodos correlatos de segmentação de imagens de fundo de olho.

1.3 Contribuições

A segmentação dos vasos sanguíneos é uma das etapas mais importante para detectar alterações nas estruturas vasculares da retina. Neste sentido, abordamos o problema com o foco de aprimorar a imagem de fundo de olho e posteriormente construir uma ANN para segmentar os vasos sanguíneos. Assim, a primeira contribuição desta pesquisa resultou na construção de uma rede Multilayer Perceptron (MLP) para otimizar os valores do *Contrast Limited Adaptive Histogram Equalization* (CLAHE) e do filtro de Wiener para a segmentação dos vasos sanguíneos publicado em (SANTOS et al., 2020). Na segunda contribuição realizou-se uma adaptação para uma arquitetura CNN Unet para a tarefa de segmentação automática dos vasos sanguíneos. Um diferencial nesta proposta é o uso das informações de cores das imagens de fundo de olho, aprimoradas usando filtragem no domínio do tempo e frequência.

1.4 Estrutura da tese

Este trabalho está organizado em duas partes. Na primeira parte constam os antecedentes da pesquisa, situados na introdução, dos quais fazem parte os objetivos, justificativa, entre outras questões pertinentes ao desenvolvimento do trabalho. Ainda nessa parte constam, no capítulo 2, os conceitos básicos acerca do tema de pesquisa e, no capítulo 3, os trabalhos relacionados à pesquisa sobre segmentação. Na segunda parte, composta por três capítulos, tem-se a primeira contribuição, com uma rede MLP e etapas de filtragem e aprimoramento da imagem da retina (capítulo 4), logo após, apresenta-se uma CNN Unet combinando informações do modelo de cor e filtros no domínio do tempo e frequência (capítulo 5); e, por último, no capítulo 6, são apresentados as conclusões obtidas no estudo.

CAPÍTULO 2

CONCEITOS BÁSICOS

Este capítulo tem como objetivo apresentar ao leitor conceitos específicos que foram interligados no desenvolvimento do algoritmo de segmentação. Portanto, o objetivo neste capítulo é introduzir uma abordagem sucinta dos principais tópicos como: anatomia da retina, retinografias digitais, análise de imagens médicas, redes neurais artificiais, redes neurais convolucionais e métricas de avaliação. Neste capítulo não temos a pretensão de esgotar de modo profundo nenhum dos tópicos a serem abordados, mas sim fornecer uma introdução sucinta mas necessária para o entendimento do leitor.

2.1 Anatomia da Retina

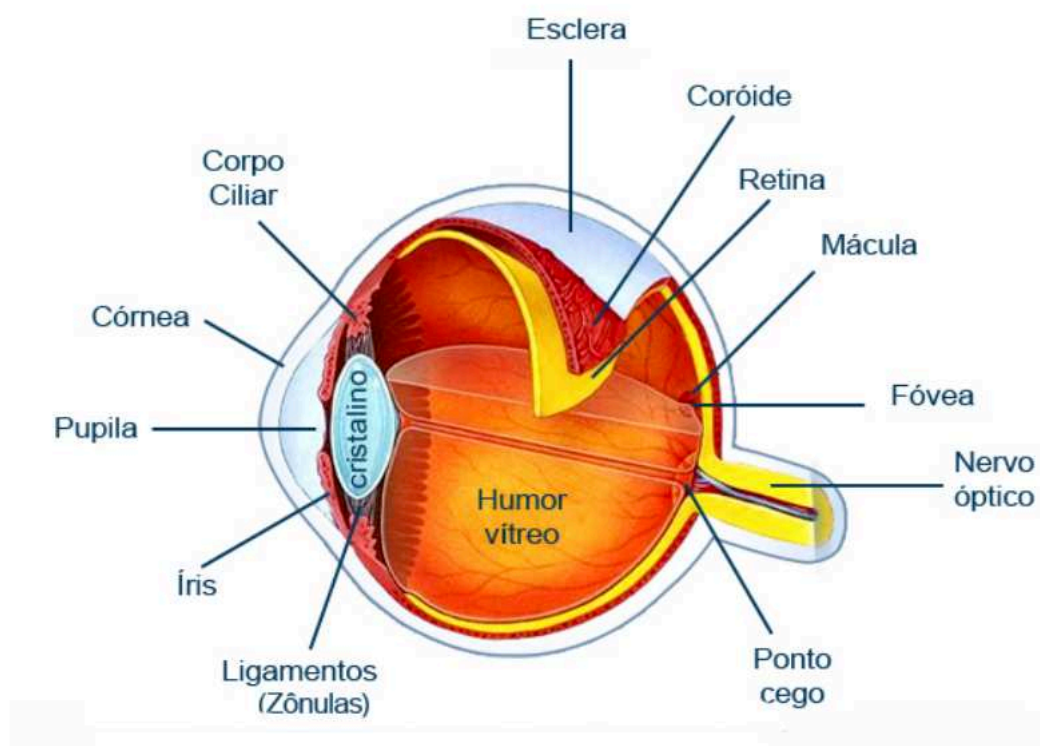
O sistema visual humano é uma estrutura complexa formada por diversos elementos combinados capazes de gerar uma percepção geral do ambiente (GONZALEZ; WOODS, 2009). Neste sentido, a visão humana é binocular e os dois olhos trabalham em conjunto para criar uma única imagem. O olho é composto por duas partes específicas, uma parte anterior formada pela córnea, íris e cristalino, e a parte posterior composta por corpo vítreo, retina, coróide e esclera (DORION, 2002).

O processo de formação da imagem é focado em uma área pequena da retina central, chamada mácula (ABRAMOFF et al., 2010). A luz refletida de um meio externo percorre o caminho passando pela córnea, pupila e lente do olho chegando à retina. Na retina, as informações mais relevantes podem ser mapeadas e interpretadas pelo cérebro por meio das células fotoreceptoras sensíveis a luz, denominadas de cones e bastonetes.

Na formação da imagem, a luz refletida de um meio externo percorre o caminho

passando pela córnea, pupila e lente do olho chegando à retina (GONZALEZ; WOODS, 2009). Em seguida, as informações mais relevantes são mapeadas e interpretadas pelo cérebro. Particularmente na retina, as células fotorreceptoras sensíveis a luz, denominadas de cones e bastonetes são responsáveis pelo processo de aquisição da imagem (GONZALEZ; WOODS, 2009). A Figura 2.1 apresenta uma seção transversal do olho humano e seus detalhes anatômicos.

Figura 2.1: Exemplo de corte longitudinal das estruturas do olho



Fonte: <https://www.optivista.com.br/pt/olho-humano/>

Algumas estruturas anômicas presentes na Figura 2.1 também são conhecidas como biomarcadores e são elementos importantes que devem ser destacados.

- **Retina** - a retina é considerada a parte mais interna das três camadas do globo ocular, as outras camadas são a coróide e a esclera. A retina é dividida em duas partes: a retina pigmentada (parte mais externa) e a retina sensorial localizada na parte mais interna. A retina é a principal responsável na formação da imagem, ela recebe as imagens formadas pelas lentes e as converte em sinais elétricos que chegam ao cérebro utilizando o nervo óptico. Coróide é a estrutura que reveste o globo ocular e está localizada entre a retina e a esclera (DORION, 2002).

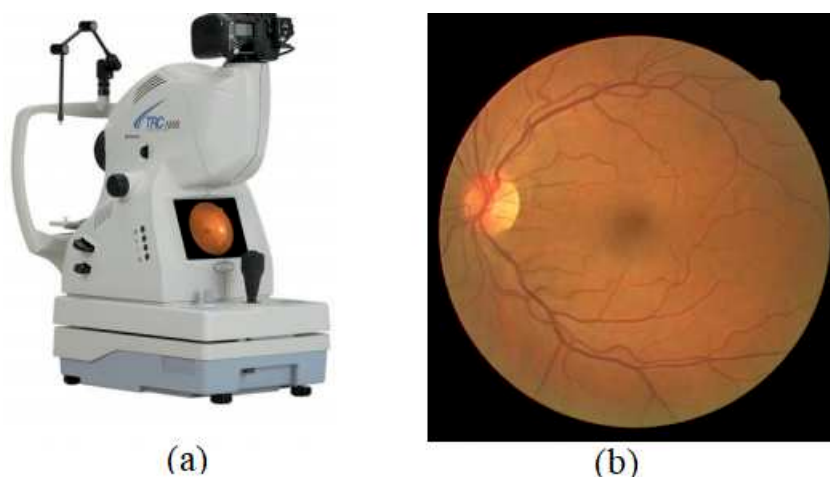
- **Disco óptico** - o disco óptico pode ser descrito como uma estrutura aproximadamente elíptica com um diâmetro médio de 25mm, em que as fibras nervosas do olho de toda a retina convergem aproximadamente para o centro. Com relação a sua tonicidade suas bordas são bem definidas e seu contorno em retinografias coloridas podem variar do rosa ao laranja com uma depressão amarelo-pálida no centro. O disco é normalmente dividido em duas partes, a temporal e nasal. A metade temporal é mais pálida que a metade nasal e mais transparente.
- **Mácula** - a mácula é considerada a parte ocular com maior sensibilidade visual e é a parte mais próxima ao centro da retina e permite visualizar os objetos com um grande nível de detalhes. Em relação aos seu aspecto ou tonalidade de cor é a parte mais escura que o restante do fundo de olho.
- **Fóvea** - a fóvea é uma área muito importante da retina que concentra muitas células fotorreceptoras, sua localização é do lado temporal do disco óptico, ou seja, na mácula, e sua distância até ele é de aproximadamente 2.5 vezes os diâmetros do disco óptico.
- **Vasos sanguíneos** - os vasos sanguíneos são formados por veia central e artéria central retiniana. A estrutura vascular é dividida em vasos retinianos, vasos coróides e capilares. Os vasos retinianos entram e saem da retina pelo nervo óptico e bifurcam em ramos superior e inferior que se dividem em nasal e temporal. Naturalmente as artérias se ramificam em ângulo agudo e na primeira das ramificações geram arteríolas de coloração vermelho clara. Além dos vasos apresentados acima, existem vasos capilares de calibre mais finos presentes nas camadas ganglionares.

2.1.1 Retinografia Digitais

No contexto atual de aquisição de imagens médicas, o fundo do olho é considerado uma representação em cores detalhada que permite visualizar as artérias, veias e outras estruturas anatómicas de forma não invasiva (FRAZ et al., 2012c). Antes da aquisição das imagens, as pupilas dos olhos são dilatadas com auxílio de um colírio aumentando assim o ângulo de observação (DORION, 2002). Em seguida, um microscópio de baixa potência com uma câmera acoplada é utilizado para capturar os detalhes anômicos da retina. Um feixe de luz é projetado no centro da íris, de forma que o interior dos olhos

e os detalhes anatômicos possam ser visualizados na retina (ABRÀMOFF et al., 2010). Muitos fabricantes de câmeras de fundo de olho especializadas como *Topcom*, *Zeiss* e *Canon* fornecem uma imagem de fundo de olho colorida ou monocromática (FRAZ et al., 2012a). Na Figura 2.2 têm-se um equipamento denominado de retinógrafo e uma imagem da retina, exemplificando o processo de aquisição das imagens (STAAL et al., 2004).

Figura 2.2: Processo de aquisição da imagem da retina. (a) Aparelho denominado de Retinógrafo digital. (b) Imagem da retina adquirida do retinógrafo



Fonte: adaptado de <http://www.topconmedical.com/products/trcnw8.htm> -maio. 2021

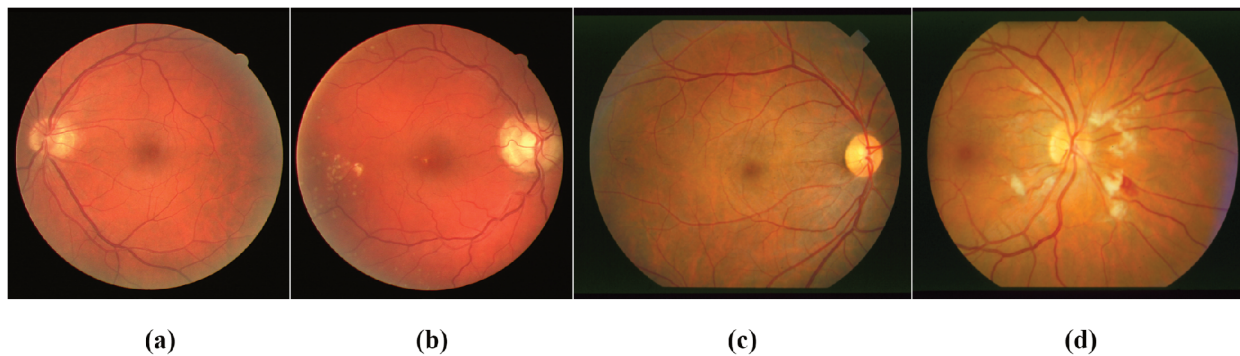
2.1.2 Enfermidades oculares

Nas últimas décadas as técnicas de processamento digital para imagens oftalmológicas procuram acelerar o diagnóstico das enfermidades visuais (MOCCIA et al., 2018). As deficiências visuais podem ser classificadas em reversíveis e irreversíveis. Os problemas refrativos que afetam a visão humana, como miopia, astigmatismo e catarata podem ser reversíveis (DORION, 2002). No entanto, glaucoma e distrofias retinianas, são consideradas irreversíveis pela gravidade dos danos causados as células fotoreceptoras (World Health Organization, 2019). Outras enfermidades como: doenças sistêmicas, diabets mellitus e hipertensão arterial sistêmica, vasculopatias retinianas podem ser evitadas (FRAZ et al., 2012c).

Organizações como a Agência Internacional para Prevenção da Cegueira (*International Agency for Prevention of Blindness* (IAPB)) e a Organização Mundial da Saúde têm trabalhando em conjunto por mais de quatro décadas apoiando diversos programas que proporcionam o cuidado com a saúde em geral (WHO, 2016). Em especial, o seu relatório mais recente revela que 2.2 bilhões de pessoas apresentam alguma deficiência visual ou

problemas oculares que frequentemente não são tratados ([World Health Organization, 2019](#)). Na Figura 2.3 é mostrado dois exemplos de imagens retinianas dos bancos de dados DRIVE e STARE evidenciando suas características anatômicas e patológicas.

Figura 2.3: Exemplos de imagens Normais e patológicas. (a) Imagem normal - [DRIVE](#). (b) Imagem Patológica - [DRIVE](#). (c) Imagem normal - [STARE](#). (d) Imagem patológica - [STARE](#)



Fonte: ([STAAL et al., 2004](#)) e ([HOOVER, 2000](#))

No cenário global, o relatório indica que aproximadamente um bilhão de pessoas poderiam ter a partir do diagnóstico prematuro evitado ou até mesmo se curado com o tratamento adequado ([World Health Organization, 2019](#)). Nos países com renda mais baixa e média os idosos e os moradores de comunidades rurais são os mais afetados por doenças oculares ([World Health Organization, 2019](#)).

Neste sentido, as projeções realizadas mostram um aumento da demanda global de cuidados com a visão. O aumento desta demanda global também está relacionada ao crescimento populacional, envelhecimento e mudanças no estilo de vida ([World Health Organization, 2019](#)). Algumas das enfermidades oculares responsáveis para o aumento dos casos de cegueira e baixa visão são o glaucoma, a retinopatia diabética, a degeneração macular relacionada à idade, entre outras ([MOOKIAH et al., 2013](#)), ([DORION, 2002](#)). Neste sentido, a Tabela 2.1 enumera alguns sinais oculares que podem indicar o início de algumas das enfermidades citadas anteriormente. A Figura 2.5 retrata os aspectos visuais nas imagens de fundo de olho relacionadas as enfermidades listadas na Tabela 2.1. As principais enfermidades que afetam um grande número de pessoas podem ser listadas em glaucoma, retinopatia diabética e degeneração macular relacionada à idade ([DORION, 2002](#)).

O glaucoma é classificado como uma enfermidade oftalmológica crônica e neurodegenerativa, que apresenta consequências irreversíveis se não for tratada a tempo, e pode levar o paciente à perda total da visão. Assim, o glaucoma ocupa uma das posições

Tabela 2.1: Exemplos de algumas sinais oculares em imagens retinianas

Sinais Oculares	Definição da enfermidade
Exsudatos moles	são microinfartos e estão relacionados a enfermidades como retinopatia diabética e hipertensão.
Exsudatos duros	são depósitos lipoprotéicos com aspecto brilhante e margens delimitadas.
Microaneurismas	são lesões em forma sacular que surgem na parede dos vasos capilares.
Hemorragias	podem ocorrer na camada pré-retiniana, na retina, ou no vítreo.
Tortuosidade	quando ocorre uma mudança na estrutura vascular retiniana.
Anormalidades microvasculares	segmentados de microvasos dilatados e tortuosos, geralmente na periferia do disco.
Edema Macular	causado pelo vazamento de plasma e como consequência ocorre o inchaço.
Neovascularização	crescimento anormal de novos vasos sanguíneos.

globais com maior número de pacientes com cegueira irreversível (FRAZ et al., 2012a). Uma das dificuldades para diagnosticar o glaucoma é sua características assintomática. O glaucoma pode dividido em dois casos específicos: o glaucoma primário de ângulo aberto (*Primary Open-Angle Glaucoma* (POAG)) e glaucoma primário de ângulo fechado (*Primary Angle-Closure Glaucoma* (PACG)). Nesta perspectiva, as projeções resultam em 76 milhões de pessoas afetadas globalmente em 2020 (World Health Organization, 2019). Os sintomas visuais podem ser percebidos como na indicação da Figura 2.4.

Um dos motivos mais evidentes segundo pesquisadores é o envelhecimento rápido das pessoas em todo o mundo (OVREIU et al., 2020). Em relação as imperfeições que ocorrem no disco óptico elas caracterizam uma enfermidade. Assim, a cabeça do nervo óptico é dividida em duas regiões. A primeira, uma zona periférica mais conhecida como borda neuroretiniana. A segunda, uma zona central com mais brilho chamada de copo óptico. Os danos causados à cabeça do nervo óptico é causado pela alteração da pressão

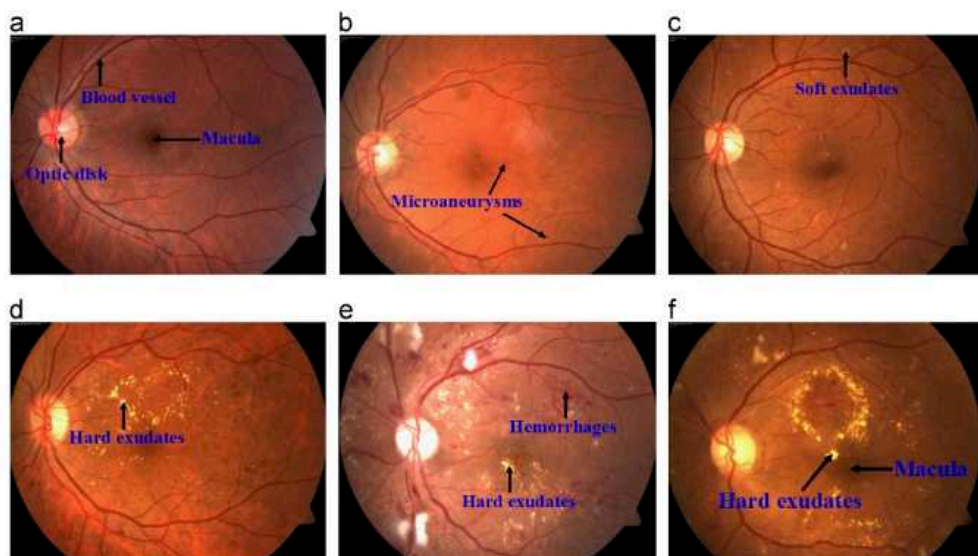
Figura 2.4: Sintomas visuais causados pelo glaucoma

Fonte: <https://www.aoa.org/healthy-eyes/eye-and-vision-conditions/glaucoma?sso=y>

intra-ocular do olho (FRAZ et al., 2012a).

Uma das principais causas da cegueira que afetam as pessoas em idade produtiva é atribuída a retinopatia diabética que é considerada uma complicação da diabetes mellitus, e além disso, cerca de 40% das pessoas com diabetes tipo 2 apresentam algum grau de retinopatia (JONES; EDWARDS, 2010). A retinopatia é dividida em dois estágios, no primeiro temos a retinopatia não proliferativa em que os danos aos vasos sanguíneos são evidentes (DORION, 2002). No segundo, a retinopatia proliferativa caracterizada pelo aparecimento de novos vasos sanguíneos anormais (DORION, 2002). A gravidade da retinopatia pode ser avaliada mediante a identificação de micro-aneurismas, hemorragias, exsudatos (duros e moles), todos listados na Figura 2.5 (FRAZ et al., 2012b).

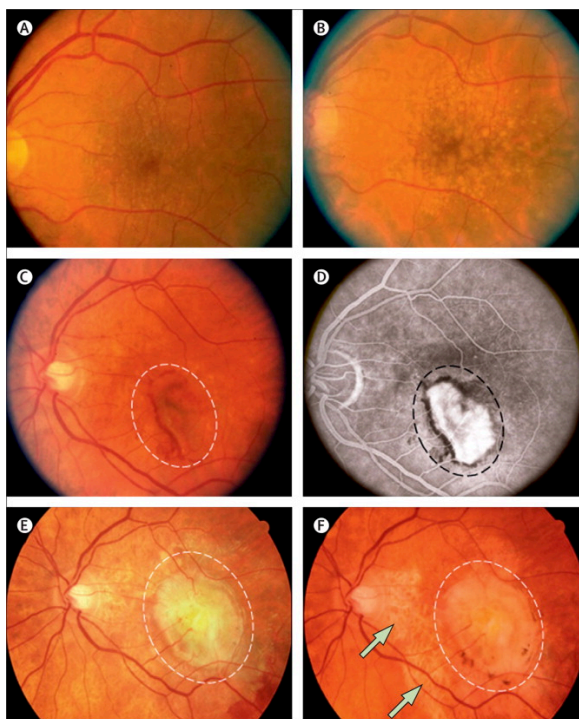
Figura 2.5: Sintomas anatômicos na retina. (a) Os principais elementos da retina, vasos sanguíneos, disco óptico e mácula. (b) presença de microaneurismas. (c) exsudatos moles. (d) exsudatos duros. (e) combinação de exsudatos duros e hemorragias. (f) exsudatos duros afetando a mácula.



Fonte: (FRAZ et al., 2012c)

Age-Related Macular Degeneration (**ARMD**) é uma lesão degenerativa relacionada a mácula e ocorre bilateralmente de forma simultânea ou não (**World Health Organization, 2019**). A **ARMD** é classificada em dois tipos clínicos: seco e úmido. Na degeneração seca (não vascular) é caracterizada pela mácula pálida, com manchas irregulares, áreas bem demarcadas de pigmentação mosqueadas por acúmulo e dispersas, e despigmentação (**World Health Organization, 2019**). Na Figura 2.6 é apresentado alguns casos do degeneração macular relacionada à idade.

Figura 2.6: Sintomas anatômicos na retina. (a) Drusas grandes ou seja manchas sub-retinianas amareladas. (b) Progressão das drusas depois de cinco anos. (c)- (d) Degeneração macular neovascular relacionada a idade. (e) Desenvolvimento de fibrose densa sub-retiniana não tratada após um ano (região tracejada). (f) Desenvolvimento de áreas atróficas ao redor do novo vaso sanguíneo após três anos de não tratamento (região tracejada).



Fonte: (**LIM et al., 2012**)

Em alguns casos pode existir drusa que se apresenta como manchas pequenas amarelas, discretas, redondas, ligeiramente elevadas, duras e com bordas não definidas (**DORION, 2002**). Na degeneração macular úmida (ou exsudativa) a mácula pode ser vista com uma aparência vitrificada composta por uma superfície irregular e é elevada com acúmulo de fluido sub-retiniano. Assim uma consequência grave deste sintomas pode levar ao descolamento macular seroso ou hemorrágico (**DORION, 2002**).

2.1.3 Análise de imagens médicas

O número de aplicações que utilizam a tecnologia de processamento digital de imagens cresceu de forma exponencial nos últimos anos. Em especial, nas áreas de visão computacional, processamento de sinais médicos e morfológicos, classificação óptica, realidade aumentada, detecção de recursos entre outras (AGGARWAL et al., 2018). Imagens médicas podem ser aprimoradas para fornecer uma perspectiva melhor para o profissional que realiza a avaliação. Neste sentido, os algoritmos de análise de imagem de baixo nível são utilizados para transformar as imagens em representações de alto nível, com objetivo de preservar as informações principais, sejam elas em escala de cinza ou imagens coloridas (GONZALEZ; WOODS, 2009).

O Aprimoramento de imagens é um termo que tem ganhado destaque em muitas aplicações, pois seu principal objetivo é melhorar a aparência visual das imagens e torná-las interpretáveis para os avaliadores humanos (GONZALEZ; WOODS, 2009). Muitos recursos em imagens, tais como bordas, limites e contrastes podem ser aprimorados (WANG et al., 2021). Estes aspectos podem ser avaliados no domínio espacial ou no domínio da frequência. Entretanto existe uma grande diferença na abordagem de algoritmos baseados no domínio do tempo ou frequência (GONZALEZ; WOODS, 2009). Máscaras espaciais como filtros de suavização e nitidez atuam sobre cada pixel da imagem no domínio espacial. No domínio da frequência, a transformada de Fourier é aplicada nos pixels da imagem por meio de uma função de transformação (AL-BANDER et al., 2018).

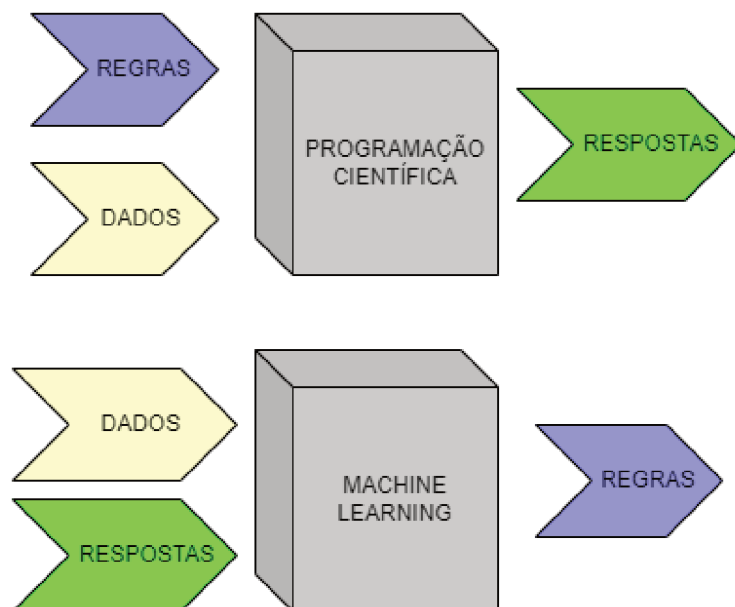
Imagens médicas são propensas a vários tipos de ruídos. Os instrumentos de captura podem inserir ruídos gaussianos, *poisson*, sal e pimenta, *speckle noise* entre outros (AL-BANDER et al., 2018). As interferências provocam um processo de degradação nas imagens adquiridas. Entretanto, muitos métodos foram propostos para aprimorar as imagens com ruído. Muitos métodos são baseados em abordagem estatística, e exploram a classificação dos *pixels* em uma determinada janela (AGGARWAL et al., 2018). Os mais conhecidos são o filtro mediano padrão, filtro de mediana ponderado, filtro da média entre outros. O algoritmo popular no domínio espacial, denominado de CLAHE, é frequentemente utilizado para melhorar o contraste de diferentes imagens médicas e tem sido usado com sucesso em diferentes aplicações (LIDONG et al., 2015), (KUMAR et al., 2018), (JEBASEELI et al., 2019). No domínio da frequência o filtro *Wiener* é um algoritmo versátil e pode aprimorar sinais uni-dimensionais ou bidimensionais (GONZALEZ; WOODS, 2009).

2.2 Inteligência Artificial e Redes Neurais Artificiais

Inteligência Artificial (*Artificial Intelligence* (AI)) é um campo geral que engloba *machine learning* e *deep learning* (BISHOP, 2006). Sua primeira abordagem ficou conhecida como AI simbólica, que se desenvolveu entre 1950 e 1980, cujo objetivo era elaborar uma grande quantidade de regras, codificadas por programadores, para resolver uma tarefa específica (BENGIO et al., 2017). Um exemplo prático desta abordagem foram os primeiros programas para jogar xadrez.

Entretanto, o sucesso da AI simbólica não conseguiu transpor problemas como a classificação de imagens, reconhecimento de fala e a tradução de idiomas (AGGARWAL et al., 2018). Desta forma, o conceito de *machine learning* foi concebido como novo paradigma (JONES, 2017). A questão natural a ser respondida é se as máquinas são capazes de aprender por si mesmas para executar uma tarefa específica (GRAVES, 2012). A Figura 2.7 apresenta o conceito geral em torno da área de *machine learning* e destaca em termos gerais a mudança de paradigma para compreender e resolver problemas científicos.

Figura 2.7: Aprendizado de máquina como um novo paradigma de programação

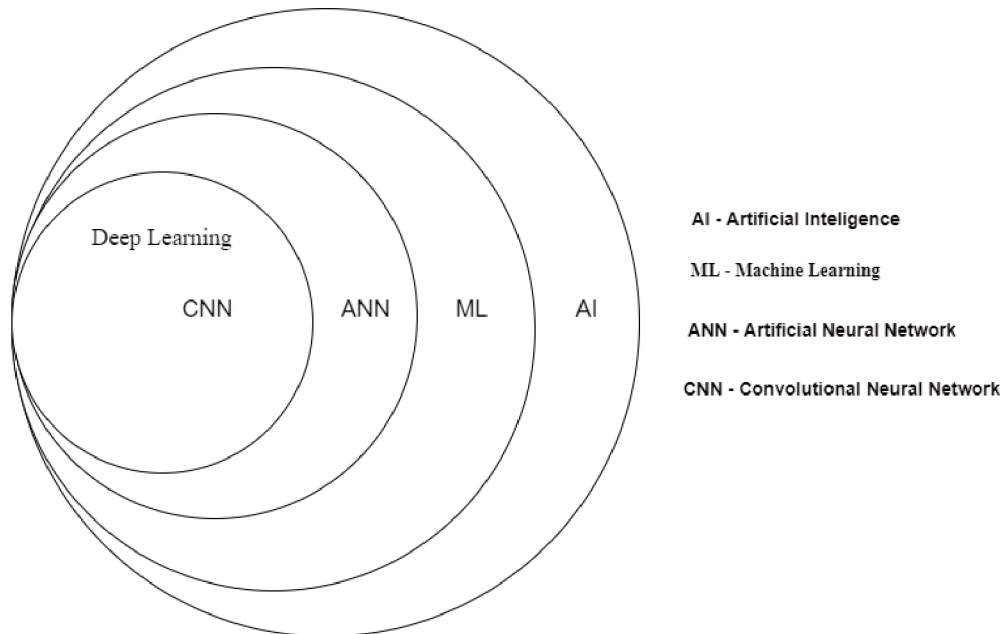


Fonte: Elaborada pelo autor

Os métodos de *machine learning* constituem um subcampo da área de AI. A Figura 2.8 mostra as subdivisões da AI (GOODFELLOW et al., 2016). Na área de visão computacional o avanço científico desta abordagem pode ser notado com maior evidência (AGGARWAL et al., 2018). Os motivos evidentes são a disponibilidade de

grandes conjuntos de imagens com milhares de exemplos, e computadores capazes de realizar cálculos complexos (BENGIO et al., 2017).

Figura 2.8: Diagrama geral mostrando Inteligência Artificial e seus subcampos

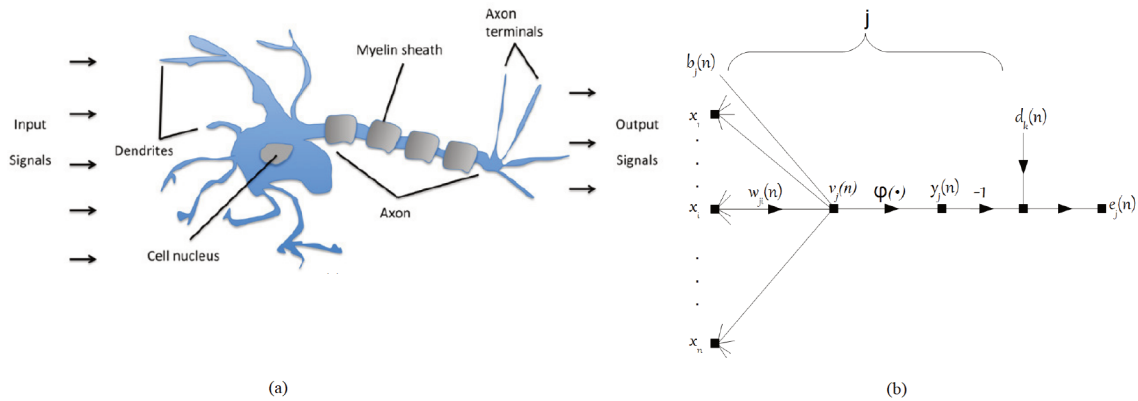


Fonte: Adaptado de (GOODFELLOW et al., 2016)

As primeiras menções das **ANNs** surgiram por volta de 1940, mas precisamente no ano de 1943 (GOODFELLOW et al., 2016). O neurofisiologista Warren McCulloch e o matemático Walter Pitts, pesquisador do Universidade de Illiniois, construíram um novo entendimento sobre o sistema nervoso e fizeram uma analogia entre as células nervosas (Figura 2.9 (a)) e a representação matemática de um neurônio (Figura 2.9 (b)) (MCCULLOCH; PITTS, 1943). A pesquisa foi publicada no Bulletin of Mathematical Biophysics com o seguinte título: *A Logical Calculus of the Ideas Immanent in Nervous Activity* (MCCULLOCH; PITTS, 1943). No ano de 1949, o psicólogo e biólogo Donald Hebb publicou um livro reforçando as teorias de que o condicionamento psicológico estava presente em qualquer parte dos animais, e que este fato estava relacionado aos neurônios e suas propriedades individuais (HEBB, 1949).

Em 1951, Marvin Minsky construiu o primeiro simulador denominado de SNARC, ele podia ajustar seus pesos sináptico de forma automática (MINSKY, 1956). No ano de 1956, o pesquisador da IBM Nathaniel Roschester durante a primeira Conferência Internacional de **AI**, apresenta um modelo que simulava centenas de neurônios interconectados. Este modelo era capaz de verificar como a rede responderia a estímulos ambientais (MINSKY, 1956).

Figura 2.9: Estrutura e fluxo do MLP. (a) Estrutura de um neurônio biológico. (b) Estrutura do neurônio matemático com seta para direita indicando o fluxo de processamento dos dados.



Fonte: (HAYKIN, 2010)

Em 1959, o trabalho de Frank Rosenblatt apresentou uma rede de múltiplos neurônios do tipo discriminadores lineares, trabalho baseado nos trabalhos de McCulloch para desenvolver o modelo da sinapse humana (ROSENBLATT, 1960). A obra publicada por Minsky e Seymour Papert, no final da década de 50, demonstrava que o modelo apresentado por Rosenblatt não era muito promissor, devido ao uso de técnicas empíricas para determinar as soluções (BENGIO et al., 2017). Outro ponto importante era a grande dificuldade computacional da época com recursos muito limitados. De fato, esta publicação esfriou as pesquisas e os aportes financeiros da época e poucos pesquisadores continuaram seguindo esta linha de investigação (GOODFELLOW et al., 2016).

As novas técnicas computacionais contam com a capacidade das unidades de processamento gráfico (*Graphical Processing Units* - GPUs) para executar algoritmos complexos de forma paralela (GOODFELLOW et al., 2016). Por exemplo, as arquiteturas ANN e CNN usadas no processamento de imagens utilizam o alto desempenho das GPUs para implementar seus algoritmos (PATTERSON; GIBSON, 2017). O uso das GPUs permite treinar sistemas grandes e complexos baseados nos algoritmos de *machine learning* e assim compor diferentes arquiteturas com várias camadas de profundidade (JONES, 2017). Neste sentido, os algoritmos de *machine learning* são desenvolvidos utilizando grandes conjuntos de dados, como conjuntos de imagens, conjuntos de textos e áudio (GRAVES, 2012). Contudo, em um conjunto de dados rotulados é preciso usar o recurso para treinar, ajustar e testar a precisão do modelo construído (AGGARWAL et al., 2018).

2.2.1 Aprendizado Supervisionado

O problema de *machine learnig* em que os dados rotulados podem ser traduzidos como um par de entrada e saída, ou seja, para cada entrada é especificada uma saída ou resposta desejada são definidos como aprendizado supervisionado (GRAVES, 2012). A etapa prática do aprendizado supervisionado de uma ANN deve ser elaborada com alguns cuidados em relação a compensação do tempo e sua aplicação. Os problemas estão associados a escolha do modelo e as escolhas subjetivas para o processo de aprendizado supervisionado (BISHOP, 2006). Um exemplo simples seria sintonizar a ANN com os mesmos dados usados no treinamento. Neste caso, o que ocorre com a rede treinada é que os resultados obtidos sofrem um sobreajuste (*overfitting*), ou seja, a ANN se torna extremamente especialista em reconhecer os padrões dos dados que compõem o treinamento (GOODFELLOW et al., 2016). Portanto, uma estratégia é treinar a rede utilizando uma divisão dos dados. Assim, um determinado conjunto de dados (Figura 2.10) deve sempre ser dividido em três partes definidas de acordo com a forma como os dados são usados:

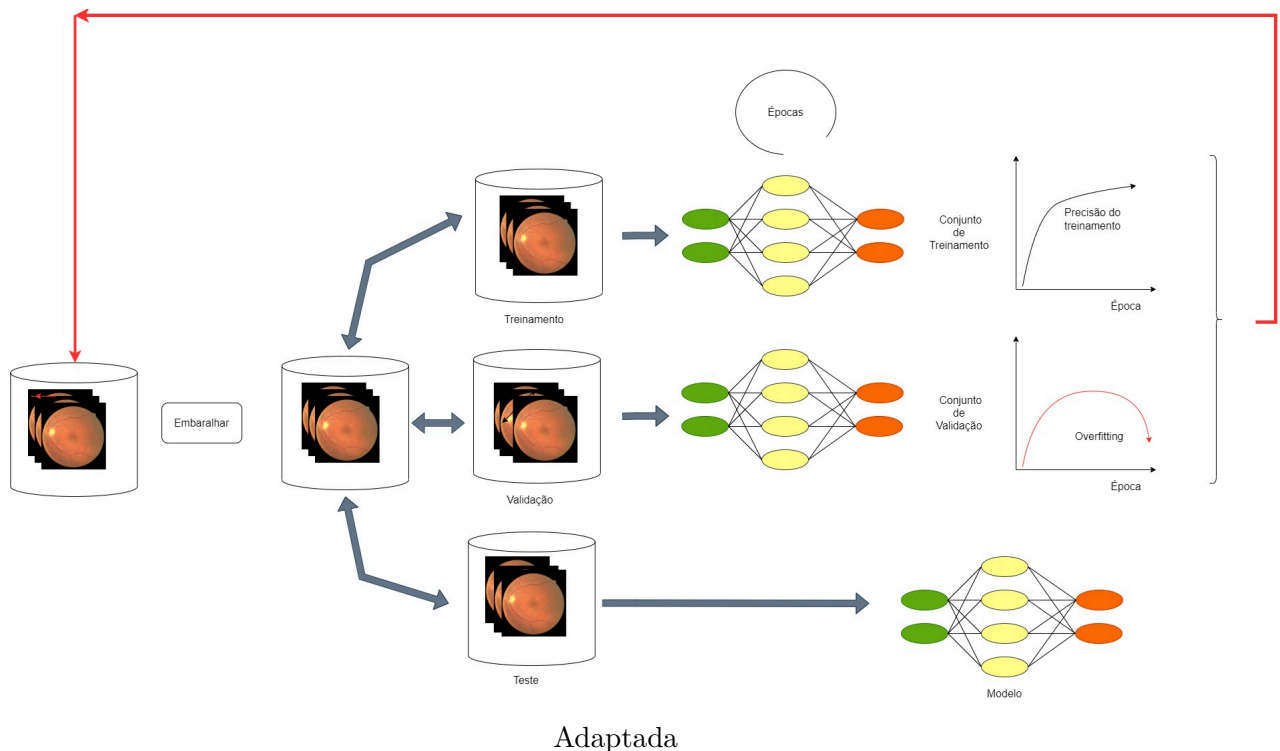
- **Conjunto de treinamento** – o conjunto de treinamento é utilizado para construir o modelo de treinamento, ou seja os pesos atribuídos a arquitetura da rede devem ser apreendidos durante o processo (HAYKIN, 2010). Diferentes projetos podem ser concebidos na construção de um modelo ANN e CNN. Nesta fase o conjunto de treinamento pode ser tentado várias vezes, e diferentes escolhas para taxas de aprendizagem ou regularização que compõe a arquitetura da ANN ou CNN (PATTERSON; GIBSON, 2017). Em tese, o processo seleciona o melhor algoritmo entre as possíveis combinações. No entanto, a avaliação real do modelo proposto não é realizada nos dados de treinamento, mas em um conjunto de validação separado para evitar que o modelo obtenha um desempenho com vícios (AGGARWAL et al., 2018).
- **Conjunto de validação** - esta parte é usada para selecionar o melhor modelo e ajustar os hiperparâmetros (PATTERSON; GIBSON, 2017). Por exemplo, a taxa de aprendizagem é ajustada repetindo o modelo várias vezes no conjunto de treinamento. Cada ciclo completo repetido na arquitetura da rede é denominado de época (GOODFELLOW et al., 2016). Na sequência o modelo é submetido ao conjunto de validação para estimar a precisão alcançada. De modo geral, a conjunto de

validação é utilizado para ajustar: a taxas de aprendizagem, o número de camadas ou unidades em cada camada, e ainda selecionar o melhor design frente aos algoritmos de otimização baseados em gradientes (HAYKIN, 2010).

- **Conjunto de teste** – este conjunto é responsável para executar um teste de precisão no modelo ANN ou CNN proposto (CHOLLET, 2021). É importante garantir que o conjunto de teste não tenha sido utilizado em nenhuma das fases anteriores. Assim, o conjunto de teste é usado apenas uma vez no final do processo para verificar a precisão do modelo.

Portanto, o processo de treinamento para ANN e CNN é composto por treinamento, validação e teste (AGGARWAL et al., 2018). A Figura 2.10 apresenta todas as fases a serem desenvolvidas em projeto de *machine learning* usando ANN ou CNN. Os ajustes realizados nos neurônios das camadas ocultas são realizados por funções de ativação seção 2.2.3. Contudo, a verificação se o modelo aprendeu uma boa generalização é realizado por meio de algoritmos de otimização baseados em gradientes que serão discutidos na seção 2.3.

Figura 2.10: Diagrama intuitivo sobre o processo de divisão do conjunto de dados



Outro fator importante relacionado a essa etapa de treinamento para ANN e CNN são as definições de hiperparâmetros e parâmetros (PATTERSON; GIBSON, 2017).

- Os hiperparâmetros segundo (PATTERSON; GIBSON, 2017) são definidos como qualquer configuração que pode ser escolhida pelo usuário e que afeta de forma direta o desempenho da rede. A classificação dos hiperparâmetros se enquadram em categorias como: tamanho da camada, magnitude (inicialização, taxa de aprendizagem), regularização (*dropout*, norma L_1 e L_2), ativações (famílias de funções de ativação), estratégia de inicialização de peso, funções de perda, configurações para épocas durante o treinamento (tamanho do lote de treinamento), esquema de normalização para os dados de entrada.
- Parâmetros são responsáveis por transformar a entrada e determinar as classificações que uma rede infere. Em outras palavras, os parâmetros representam os pesos das conexões na ANN ou CNN já treinadas (AGGARWAL et al., 2018).

Desta forma, a escolha do algoritmo para o treinamento depende do entendimento do problema, da adequação e design de todos os detalhes (GOODFELLOW et al., 2016). A Figura 2.9 (b) apresenta o fluxo de entrada e saída da rede para um neurônio. Na saída final da rede temos a função $e_j(n)$ que realiza o cálculo do erro entre os dados oriundos da rede e os dados do *ground truth*. Os termos $x_1, \dots, x_j$ representam os dados de entrada, $w_{ji}(n)$ representam o peso na camada ji e $v_j(n)$ representa os pesos na camada j , e o termo $b_j(n)$ representa o *bias* na camada j , e o termo $\varphi(\cdot)$ representa a aplicação de um função de ativação, para inserir a não linearidade no modelo, o termo $y_j(n)$ representa a saída da rede, e o termo $d_k(n)$ representa a resposta desejada.

Assim, as informações propagadas na fase de treinamento da ANN estão presentes na Figura 2.9 (b), em que os termos d_k e y_k são submetidas a uma função que calcula o erro, para um neurônio na camada j e a saída é da forma $e_j(w_k) = d_j(k) - y_j(k)$ como observado na Equação 2.1

$$E_k(w) = \frac{1}{N} \sum_{n=1}^N (d_k - y_k)^2 \quad (2.1)$$

em que d_k representa os dados de rotulados de saída, e y_k representa os dados de entrada já regularizados após à aplicação da função de ativação. Desta forma, para cada rodada de treinamento é possível ajustar os pesos sinápticos do modelo ANN e CNN. Assim é possível construir uma curva de aprendizagem, Figura 2.10, plotando o erro médio em relação ao número de épocas e medir precisão do modelo proposto (HAYKIN, 2010).

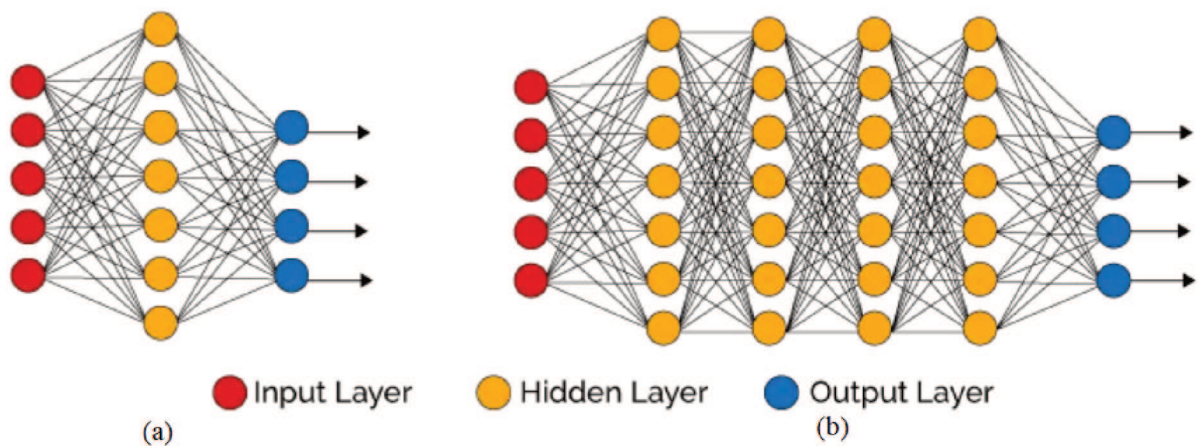
Uma **ANN** pode ser vista como uma abstração computacional que busca emular o funcionamento do sistema nervoso do ser humano (HAYKIN, 2010). As ligações entre cada um dos neurônios da rede é denominado de pesos das conexões e simulam teoricamente a força das sinapses entre os neurônios (FAUSETT, 1994). A rede é ativada quando fornecemos uma entrada para alguns neurônios ou para todos os neurônios, assim esta ativação se propaga ao longo das conexões ponderadas (GOODFELLOW et al., 2016). A Equação 2.2 generaliza a saída de um neurônio mostrada na Figura 2.9 e seu fluxo de propagação das informações.

$$\hat{y}_k = \left(\sum_{j=1}^m w_j x_j + b \right) \quad (2.2)$$

em que os valores $\hat{y}_k$ representam as saídas dos neurônios artificiais combinados para uma entrada $x = [x_1, x_2, x_3, \dots, x_m]^T$. E os valores $w = [w_1, w_2, \dots, w_m]$ são os pesos associados para cada uma das m dimensões do vetor de entrada e o termo b que representa o *bias*.

Uma classificação importante para as arquiteturas de **ANN** está relacionada à quantidade de camadas existentes entre a entrada e a saída da rede. A Figura 2.11 (a) e (b) demonstra a diferença sutil entre **ANN** com uma única camada oculta e com mais camadas ocultas, caracterizando sua profundidade. Muitas arquiteturas exploram a profundidade das **ANN** para resolver os mais diversos problemas (KHAN et al., 2020).

Figura 2.11: Composições de Arquiteturas. (a) **ANN** arquitetura básica. (b) **ANN** arquitetura profunda



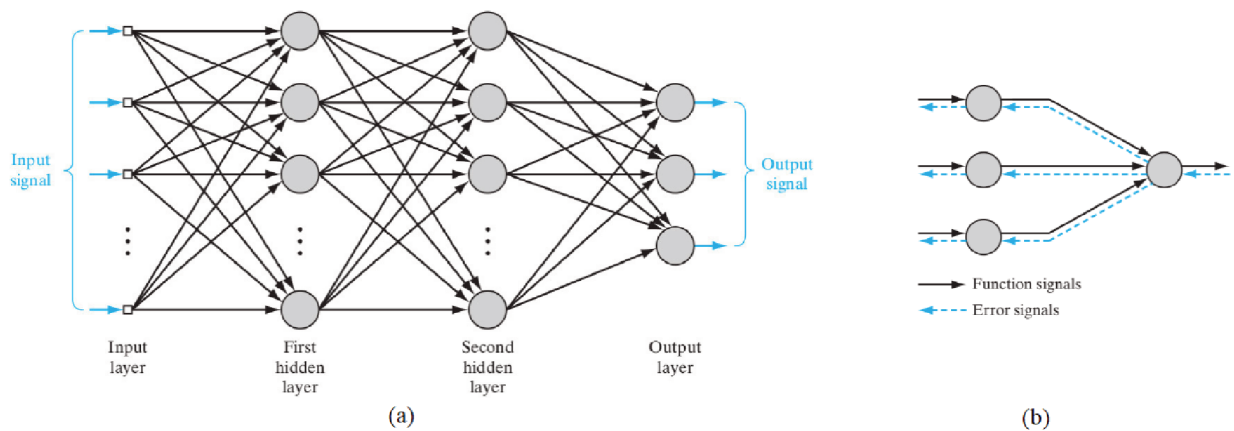
Fonte: (PICCINELLI, 2020)

2.2.2 Perceptron Múltiplas Camadas

Os modelos baseados em ANN podem ser classificados pela arquitetura (padrão de conexões entre os neurônios), pelo treinamento (método de determinar os pesos nas conexões) e pela função de ativação (GOODFELLOW et al., 2016). Assim, uma ANN simples é composta basicamente por uma camada de entrada, uma camada oculta e uma camada de saída (FAUSETT, 1994) como pode ser observado na Figura 2.12 (a). A primeira camada é constituída de neurônios de entrada conectados às camadas posteriores. Os modelos baseados em ANN também são caracterizados pelo número de camadas definidos entre a entrada e a saída de rede, caracterizando a profundidade (HAYKIN, 2010).

O MLP foi aplicado com sucesso em diferentes cenários devido a forma de treinamento supervisionada e a retro-propagação do erro (*error backpropagation*) (AGGARWAL et al., 2018). O algoritmo tem como base a regra de aprendizado por correção de erro. Assim o aprendizado com retro-propagação de erro é composto de duas etapas distintas. Na primeira etapa, as informações são propagadas para a frente (*forward*). Na segunda etapa, as informações são propagadas para trás, ou seja retro-propagação (HAYKIN, 2010). A Figura 2.12 (a) representa a arquitetura de MLP com duas camadas ocultas e (b) representa o fluxo do algoritmo retro-propagação. Ou seja, setas pretas indicam o fluxo para frente e setas azuis indicam o fluxo para trás.

Figura 2.12: Multilayer perceptron. (a) MLP com duas camadas de profundidade. (b) Fluxo das informações.



Adaptada (HAYKIN, 2010)

2.2.3 Funções de Ativação

A função de ativação representa uma soma ponderada dos dados de entrada que são transformados em uma saída que compõem a rede neural (BISHOP, 2006). Na prática a escolha destas funções tem um grande impacto na capacidade e no desempenho da rede neural (HAYKIN, 2010). Assim, as funções que compõem as redes neurais são cuidadosamente projetadas com base em suas propriedades e sua capacidade de inserir a não-linearidade ao processo realizado pelas ANN e CNN (BENGIO et al., 2017). As funções de ativação devem ter algumas características como:

- **Custo computacional** - Operações matemáticas são computacionalmente baratas. Assim, uma função de ativação incrementa o custo na medida que confere a característica desejada (HAYKIN, 2010).
- **Diferenciabilidade** - Na fase de treinamento da ANN e CNN seus pesos e *bias* são ajustados pelo mecanismo de descida do gradiente, que exige que cada operação nas camadas tenha sua derivada calculada (GOODFELLOW et al., 2016). Desta forma, as operações devem ser diferenciáveis.
- **Centradas em zero** - A função de ativação fornece um espaço determinado de resultados. Estes resultados são expressos em forma de probabilidades e classificam em positivo ou negativo (AGGARWAL et al., 2018). Assim, a origem dos valores esta no cálculo do gradiente atualizado pela função de ativação. De fato, se o resultado do gradiente for sempre positivo ou negativo, o valor atribuído pelo cálculo do gradiente será positivo ou negativo, o que pode dificultar a convergência da rede (AGGARWAL et al., 2018).
- **Não produzir platôs** - No âmbito das ANNs um platô é uma região que contém o universo de resultados que tem uma tendência a ser constante (HAYKIN, 2010). Em regiões constantes o valor da derivada que compõe o gradiente tendem a zero, o que causa o problema de *vanishing gradient* (desaparecimento do gradiente) (GOODFELLOW et al., 2016).

A Tabela 2.2 apresenta seis funções de ativação, a função linear, as funções logística e sigmoide, a função tangente hiperbólica, a função *ReLU*, a função *PReLU* e a função *Softplus*. Cada uma destas funções de ativação podem ser escolhidas com base na natureza

dos dados e as distribuições assumidas nas variáveis alvo (AGGARWAL et al., 2018). Segundo (BISHOP, 2006), existem muitas possibilidades para realizar a escolha de uma função de ativação, em especial em arquiteturas profundas (HAYKIN, 2010).

Tabela 2.2: Funções de Ativação

Nome	Função	Derivada $f'(x)$
Identidade	$f(x) = x$	$f'(x) = 1$
Logística, sigmoide	$f(x) = \frac{1}{1 + e^{-x}}$	$f'(x) = f(x)(1 - f(x))$
Tangente hiperbólica	$\tanh(x) = \frac{2}{1 + e^{-2x}}$	$f'(x) = 1 - f(x)^2$
Unidade Linear Retificada (ReLU)	$f(x) = \begin{cases} 0 & \text{se } x < 0 \\ x & \text{se } x \geq 0 \end{cases}$	$f'(x) = \begin{cases} 0 & \text{se } x < 0 \\ 1 & \text{se } x \geq 0 \end{cases}$
Unidade Linear Retificada Paramétrica (PReLU)	$f(x) = \begin{cases} \alpha x & \text{se } x < 0 \\ x & \text{se } x \geq 0 \end{cases}$	$f'(x) = \begin{cases} \alpha & \text{se } x < 0 \\ 1 & \text{se } x \geq 0 \end{cases}$
Softplus	$f(x) = \ln(1 + e^x)$	$f'(x) = \frac{e^x}{1 + e^x}$

Fonte: Elaborado pelo autor

2.2.4 Normalização de camadas

Em algoritmos de *machine learning* aprendizagem um dos problemas a ser contornado é a grande variabilidade dos dados de entrada que podem comprometer o processo de convergência da ANN ou CNN. Uma etapa que pode garantir que estes modelos continuem seu treinamento é a normalização das camadas. Em (IOFFE; SZEGEDY, 2016), a estratégia para obter uma padronização nos dados de entrada é realizada com base na entrada *mini-batch* durante a etapa de treinamento, onde é aplicada uma normalização de lote. As estatísticas referentes às amostras são utilizadas como um regularizador e assim mantém os pesos sinápticos da rede pequenos, com objetivo de não ocorrer o estouro da função gradiente (PATTERSON; GIBSON, 2017).

Considere que a camada $x = (x_1, \dots, x_d)$ seja normalizada levando em consideração as dimensões. Portanto, é possível normalizar a k -ésima dimensão usando a seguinte Equação 2.3 :

$$\hat{x}^k = \frac{x^k - E[x^k]}{\sqrt{Var[x^k]}} \quad (2.3)$$

onde o termo $E[.]$ representa a esperança e o termo $Var[.]$ é a variância da variável aleatória x^k . Desta forma, é possível dimensionar e deslocar a entrada normalizada $\hat{x}$ a partir da seguinte transformação apresentada na Equação 2.4:

$$y^k = \gamma^k \hat{x} + \beta^k \quad (2.4)$$

Em que γ e β são os parâmetros a serem aprendidos. A camada normalizada é colocada antes da função de ativação. Assumindo $z = \sigma(Wx + b)$ em que os elementos são os pesos W , a entrada x , o *bias* b e $\sigma(.)$ são parâmetros que são atualizados. Assim, a função $\sigma(BN(wx))$ com normalização em lote, em que o parâmetro β de BN assume o papel do *bias* b .

2.3 Técnicas de Otimização

Os algoritmos de aprendizagem de máquinas tem como objetivo determinar o melhor conjunto de valores para um conjunto de parâmetros condicionados ao modelo ANN ou CNN proposto (JONES, 2017). Durante o processo de otimização a tarefa específica é minimizar ou maximizar uma determinada função $f(.)$ percorrendo os valores do seu domínio. Estas funções são denominadas de função objetivo, e quando o objetivo principal é a minimização dos valores as funções também são chamadas de função de custo, função de perda, ou função de erro. A forma tradicional de otimizar os treinamentos em ANN ou CNN é utilizar o método da descida de gradiente (GOODFELLOW et al., 2016). Desta forma, a implementação do algoritmo gradiente descendente permite encontrar os valores de máximo ou de mínimo de uma função de erro.

2.3.1 Gradiente Descendente

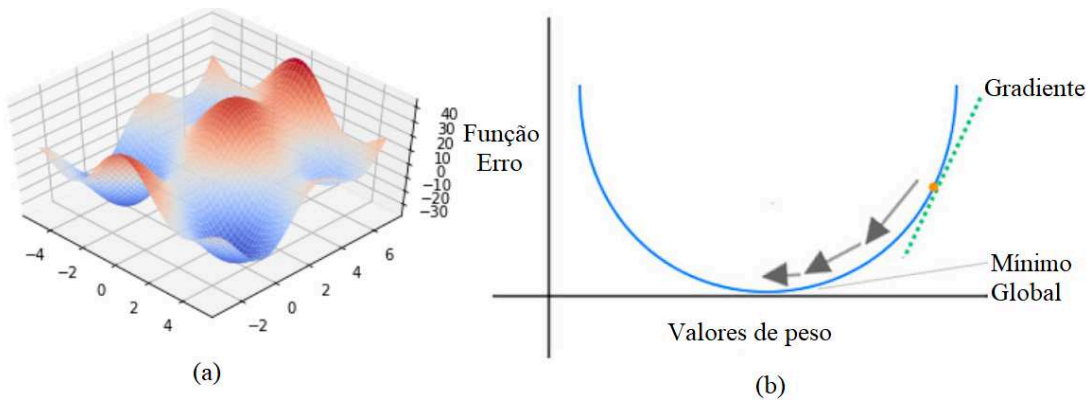
O método do gradiente descendente calcula a perda geral em todos os exemplos de treinamento e busca encontrar o mínimo de uma função de erro $f(.)$ e prevê a qualidade dos prognósticos que as arquiteturas ANN ou CNN estão desempenhando (PATTERSON; GIBSON, 2017). No contexto geral de minimização da função de erro $f(.)$ pode existir apenas um mínimo global ou vários mínimos locais da função. Em muitos problemas

práticos é preciso minimizar funções que possuem múltiplas entradas, ou seja, $f : \mathbb{R}^n \rightarrow \mathbb{R}$ (GOODFELLOW et al., 2016). Assim, para essa classe de funções é preciso fazer uso do conceito de derivadas parciais. A derivada parcial $\frac{\partial f}{\partial x_i}$ mede a variação da função f em um determinado ponto x_i do seu domínio.

Assim, o gradiente generaliza a noção de derivada para o caso em que a derivada é com relação a um vetor. Logo o gradiente da função f é um vetor contendo todas as derivadas parciais, denotado por $\nabla_w f(w)$. Cada elemento i do gradiente é a derivada parcial de f em relação a w_i . Em dimensões maiores os pontos críticos são pontos onde cada elemento do gradiente é igual a zero, ou seja, $\nabla_w f(w) = 0$ (GOODFELLOW et al., 2016). O método se inicia em um ponto e a partir desse ponto é calculado o gradiente, de tal forma que seja possível caminhar na direção do menor valor possível da função de erro (PATTERSON; GIBSON, 2017).

Em seguida, a Figura 2.13 têm como objetivo de ilustrar o funcionamento do algoritmo do gradiente descendente. O gráfico apresentado na Figura 2.13 (a) foi gerada por meio da função $f(x, y) = (\cos(0.8x))(25 - x^2) - (\cos(0.8y))(25 - y^2)$ em conjunto com sua representação gráfica e seus diferentes pontos de máximo e mínimos para um intervalo específico. Na Figura 2.13 (b) é apresentado os passos de minimização da função de erro em relação aos pesos da rede.

Figura 2.13: Exemplo do processo de otimização. (a) Função de duas variáveis. (b) Gráfico da descida do gradiente



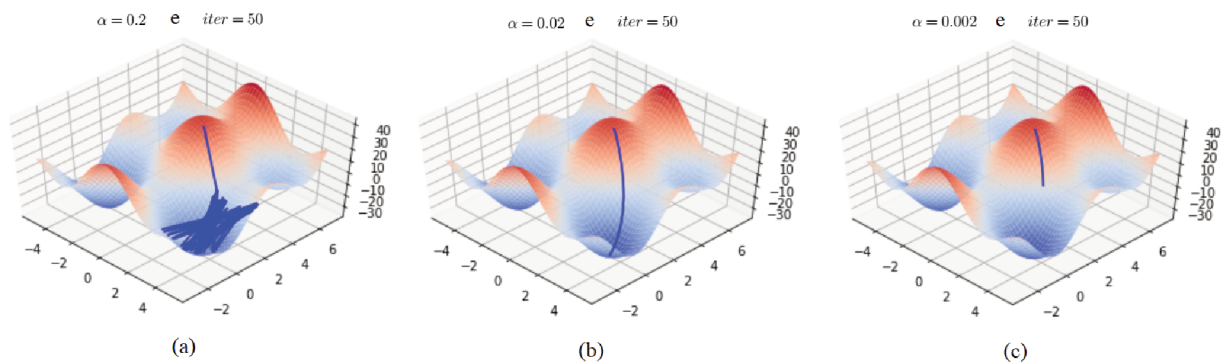
Fonte: Elaborado pelo autor

Em seguida, a Figura 2.14 apresenta um teste de convergência para o gradiente descendente e as possibilidades de escolhas dos hiperparâmetros: taxa de aprendizagem α e número de interações $iter$. De fato determinadas escolhas para os hiperparâmetros α e $iter$ influenciam diretamente no desempenho da rede. Na Figura 2.14 (a) a escolha de

$\alpha = 0.2$ e $iter = 50$ não foi adequado para guiar a convergência da rede para um mínimo local da função. Assim, a função de erro não consegue chegar ao valor mínimo e seus valores oscilam em torno da solução ótima. Na Figura 2.14 (b) a escolha de $\alpha = 0.02$ e $iter = 50$ possibilitou convergir para o valor de mínimo com poucas interações. Por fim, na Figura 2.14 (c) a escolha de $\alpha = 0.002$ e $iter = 50$ não foi o suficiente para convergir para o mínimo e neste caso é preciso de 100 iterações para convergir. Em geral, as escolhas destes hiperparâmetros é realizada de forma empírica e depende do entendimento do problema.

Um algoritmo de otimização como a descida de gradiente pode detectar a inclinação real da função em relação a cada peso (AGGARWAL et al., 2018). Assim o algoritmo mede a inclinação (a mudança no erro causada por uma mudança no peso) e leva o peso um passo em direção ao mínimo. Isso é feito tomando uma derivada da função de erro para produzir o gradiente (GOODFELLOW et al., 2016).

Figura 2.14: Exemplos da aplicação do Gradiente descendente. (a), (b) e (c) Diferentes valores para α e $iter$ fixo em 50.



Fonte: Elaborado pelo autor

Portanto, os hiperparâmetros que compõem o gradiente descendente são taxa de aprendizagem que controla o tamanho do passo dado em direção ao valor de mínimo local e o número de interações repetidas pelo o algoritmo (PATTERSON; GIBSON, 2017).

Muitos problemas de *machine learning* podem ser reformulados como problemas de otimização de função objetivo (AGGARWAL et al., 2018). A função objetiva para uma ANN ou CNN podem ser definidas em termos de otimização da função de erro. Por exemplo, em uma tarefa de regressão linear, minimiza-se a soma dos erros de predição ao quadrado sobre os pontos dos dados de treinamento disponíveis. Em outra tarefa como a redução de dimensão, minimiza-se a soma dos erros de representação ao quadrado nos pontos dos dados de treinamento reconstruídos. Assim, a função de erro, Equação 2.5,

pode ser escrita da seguinte forma:

$$E = \sum_{i=1}^n E_i \quad (2.5)$$

Aqui o termo E_i é a contribuição da função de erro para o i -ésimo dado de treinamento. Portanto, a interpretação da descida do gradiente minimiza a função de erro movendo os parâmetros ao longo da direção negativa do gradiente, por exemplo, para um **MLP** genérico, esses parâmetros correspondem a $w = (w_1, w_2, \dots, w_d)$. Desta forma, o processo de cálculo iterativo para algoritmo tradicional da descida do gradiente é apresentado na Equação 2.6:

$$w = w - \alpha \left(\frac{\partial E}{\partial w_1}, \frac{\partial E}{\partial w_2}, \dots, \frac{\partial E}{\partial w_d} \right) \quad (2.6)$$

usando a notação matricial obtemos a Equação 2.7:

$$w = w - \alpha \left(\frac{\partial E}{\partial w} \right) = w - \alpha \nabla f(w) \quad (2.7)$$

A formulação matemática para otimizar a função de erro é equivalente a resolver um problema de otimização representado na Equação 2.8:

$$w^* = \min_{w \in \mathbb{R}^n} f(w) \quad (2.8)$$

onde f é uma função que contabiliza o erro e w^* representa o valor mínimo determinado pela função $f(w)$. Assim, as iterações podem ser definidas na Equação 2.9:

$$w_{k+1} = w_k - \eta_k \hat{\nabla} f(w_k) \quad (2.9)$$

w_k denota a k -ésima iteração, η_k é uma sequência de valores que representam o tamanho do passo do gradiente em direção ao ponto ótimo, ou a taxa de aprendizado. Por fim, o termo $\hat{\nabla} f(w_k)$ representa o gradiente calculado para os w_k pesos.

Para redes **MLP** o algoritmo do gradiente descendente é calculado em relação a w . Contudo, para redes maiores como **CNN**, todos os parâmetros precisam ser atualizados (PATTERSON; GIBSON, 2017). Neste caso, o número pode ser na ordem de milhões de parâmetros e no caso de redes com etapas de retropropagação é preciso executar simultaneamente todos os dados de treinamento para frente e para trás e calcular as atualizações (BENGIO et al., 2017). O exemplo citado na Figura 2.14 ilustra bem

o comportamento do gradiente descendente, com base na Equação 2.6 cada dado de treinamento deve ser armazenado na memória até terminar as atualizações (AGGARWAL et al., 2018). No entanto, é impraticável utilizar o algoritmo de gradiente descendente em todo o conjunto de dados. Assim outras versões deste algoritmo foram propostas para trabalhar com redes compostas por muitos parâmetros (GOODFELLOW et al., 2016).

2.3.2 Momentum

O termo *momentum* pode ser conceituado como um novo hiperparâmetro associado à rede com objetivo de controlar a velocidade das mudanças que acontecem no peso w da rede. Na abordagem usando o *momentum*, o vetor w_{t+1} é modificado com base em uma suavização exponencial, em que o termo $\beta \in (0,1)$ (AGGARWAL et al., 2018). Em outras palavras, a escolha do novo hiperparâmetros acrescenta um peso na atualização da direção do gradiente e ao mesmo tempo controla as novas atualizações dos parâmetros dos novos pesos w_{t+1} para que a direção seja cada vez mais precisa no espaço de busca (GOODFELLOW et al., 2016). Observe que se o valor de $\beta = 0$, a Equação 2.10 se torna a atualização baseada no gradiente descendente.

$$w_{t+1} = \beta w_t - \alpha \sum_{j=1}^B \hat{\nabla} f(w_t, x_j^B) \quad (2.10)$$

O objetivo principal nos métodos de momentum • são baseados no comportamento das derivadas parciais. Os parâmetros com grandes valores de derivadas parciais têm oscilação forte, enquanto os parâmetros com valores pequenos de derivadas parciais têm oscilação fraca (AGGARWAL et al., 2018).

2.3.3 Gradiente Descente Estocástico

O treinamento das redes neurais é realizado por algoritmos especializados. O algoritmo denominado gradiente descendente estocástica (*Stochastic Gradient Descent* (SGD)) surgiu no trabalho de (ROBBINS; MONRO, 1951). O SGD usa como estratégia selecionar um número menor de dados de treinamento, mas ainda sim caminhar para o mínimo (AGGARWAL et al., 2018). De fato, usar essa quantidade menor que o conjunto de treinamento pode não conduzir os pesos da rede na direção de mínimo da função (GOODFELLOW et al., 2016). Entretanto essa nova seleção de amostras de treinamento fornece um indício do caminho para o mínimo, ou seja, uma aproximação grosseira do

gradiente.

Durante o processo de treinamento da ANN ou CNN deve-se atualizar os pesos entre as camadas à medida que usamos os dados rotulados (GOODFELLOW et al., 2016). Desta forma, o SGD analisa apenas uma amostra do conjunto de treinamento denominada de *mini-batch* (AGGARWAL et al., 2018). A quantidade mínima de amostras a serem escolhidas para cada interação do algoritmo SGD é chamada de *mini-batch* (GOODFELLOW et al., 2016). A escolha do tamanho do *mini-batch* depende da avaliação do custo-benefício para o seu modelo (BENGIO et al., 2017). Se o *mini-batch* escolhido têm um grande número de amostras, ele vai ser mais preciso, contudo o tempo de processamento vai ser maior (AGGARWAL et al., 2018). No sentido oposto, se o *mini-batch* for menor, ele vai ser menos preciso, contudo o tempo de treinamento para o gradiente vai ser mais rápido (AGGARWAL et al., 2018). A representação deste tipo de atualização dos pesos da rede pode ser representada na Equação 2.11 :

$$w_{t+1} = w_t - \alpha \sum_{j=1}^B \hat{\nabla} f(w_t, x_j^B) \quad (2.11)$$

onde α é o hiperparâmetro que determina a taxa de aprendizagem. Os termos w_{t+1} e w_t são as atualizações dos pesos no tempo $t+1$ e t . O termo B representa o conjunto aleatório das amostras de treinamento. De forma geral, os valores escolhidos para α decaem exponencialmente em função do número de interações calculadas e o termo $\hat{\nabla} f(w_t)$ representa o gradiente calculado para as amostras do conjunto B e x_j^B representa o conjunto *mini-batch* escolhido. A Equação 2.12 representa os pesos calculados com base no gradiente para o *mini-batch* escolhido (GOODFELLOW et al., 2016).

$$w_{t+1} = \alpha \sum_{j=1}^B \hat{\nabla} f(w_t, x_j^B) \quad (2.12)$$

Assim o processo de ajuste da taxa de aprendizado α se torna trabalhosa. Nesse sentido surgiram novas versões com objetivo de otimizar a função de erro como: o RMSProp, o *Adaptive Gradient* (AdaGrad), *Adaptive Learning Rate* (AdaDelta) e *Adaptive Moment Estimation* (Adam) (KINGMA; BA, 2015), (DUCHI; SINGER, 2009), (KESKAR; SOCHER, 2017).

2.3.4 AdaGrad e RMSProp

O algoritmo **AdaGrad** é definido como um método de otimização estocástica que adapta a taxa de aprendizagem em relação aos hiperparâmetros (GOODFELLOW et al., 2016). Um detalhe importante sobre este algoritmo é a execução das atualizações do gradiente. Para informações mais frequentes a atualização atribuída é menor. Entretanto, para informações mais raras as atualizações são maiores (AGGARWAL et al., 2018). Esta é uma característica presente em problemas com dados esparsos. O que caracteriza uma sistema esparsos é o número muito grande de equações e variáveis tornando a resolução manual e o armazenamento na memória do computador um grande desafio (BENGIO et al., 2017). Assim, o método mantém o histórico da influência de cada uma das atualizações. Os gradientes podem ser representados como $g_{(t+1)} = g_{(t)} + \nabla f(w_t)^2$. Portanto, o processo utilizado para normalizar o tamanho do passo dado em cada parâmetro é expresso na Equação 2.13:

$$w_{t+1} = w_t - \frac{\alpha}{\sqrt{g_{(t+1)} + \epsilon}} \hat{\nabla} f(w_t)^2 \quad (2.13)$$

onde o valor de g_t é um tipo de normalização, pois o termo mede a magnitude histórica do gradiente ao invés do seu sinal. O termo $\sqrt{g_{(t+1)}} + \epsilon$ serve para evitar o mau condicionamento com valores pequeno. O termo ϵ é um número real pequeno. O fator $\hat{\nabla} f(w_t)^2$ representa o valor do gradiente calculado elevado a segunda potência.

Este algoritmo é similar ao método **AdaGrad**. A estratégia para modificar a taxa de aprendizado é calcular médias da magnitude dos gradientes mais recentes antes de aplicar os gradientes. A diferença em relação ao **AdaGrad** é que o cálculo do gradiente $g_{(t)}$ usa uma média com decaimento exponencial. Assim obtém-se a representação dos gradientes por meio da Equação 2.14:

$$g_{(t+1)} = \gamma g_{(t)} + (1 - \gamma) \hat{\nabla} f(w_t)^2 \quad (2.14)$$

onde $g_{(t)}$ é denominado momento de segunda ordem de $\hat{\nabla} f(w_t)$ e os valores de γ e $1 - \gamma$ são um tipo de *momentum* inserido nas atualizações dos gradientes. Os momentos são muito importantes em estatística para caracterizar distribuições de probabilidade de uma variável aleatória. Em resumo o primeiro momento para uma variável aleatória é a média da variável aleatória. O segundo momento de uma variável aleatória é a sua variância

(AGGARWAL et al., 2018). Desta forma, a atualização dos hiperparâmetros é realizada adicionando um novo termo, resultando na Equação 2.15:

$$w_{(t+1)} = w_{(t)} + \eta(w_{(t)} - w_{(t-1)}) + (1 - \eta) - \frac{\alpha}{\sqrt{g_{(t+1)}} + \epsilon} \hat{\nabla} f(w_t) \quad (2.15)$$

em $w_{(t+1)}$ é o peso a ser atualizado. O termo $w_{(t)} - w_{(t-1)}$ é a diferença dos pesos calculados nos instantes de tempo t e $t - 1$. As variáveis η e $1 - \eta$ controlam as atualizações dos parâmetros adicionando um tipo de momentum.

2.3.5 Adam

Os algoritmos de otimização precisam superar o próprio ruído existente nos dados de treinamento sub-amostrados. Assim, o algoritmo de otimização denominado de Adam combina vantagens de dois métodos clássicos: o AdaGrad, que trabalha bem com gradientes esparsos, e o RMSProp, que funciona bem em configurações não estacionárias (DUCHI; SINGER, 2009). Portanto, o método para otimização estocástica Adam requer apenas gradientes de primeira e segunda ordem para adaptar a taxa de aprendizado para cada peso da ANN e CNN. O primeiro momento é a média e o segundo momento é a variância não centrada, ou seja, não subtraímos o valor da média durante o cálculo da variância.

Durante o processo de otimização o gradiente da função de erro pode ser considerado uma variável aleatória, pois é avaliada em *mini-batch* aleatório do dados. Assim seja β_1 e β_2 os valores que controlam w e g , respectivamente, a atualização dos pesos é o cálculo do gradiente. De fato, a influência dos parâmetros diminui ao longo do tempo, a medida que o tamanho do passo se aproxima-se do mínimo. Seja m uma variável auxiliar, assim as Equações 2.16 e 2.17 são:

$$m_{t+1} = \alpha_{t+1} g_t + (1 - \alpha_{t+1}) \hat{\nabla} f(w_t) \quad (2.16)$$

$$\widehat{m}_{t+1} = \frac{m_{t+1}}{1 - \alpha_{t+1}} \quad (2.17)$$

onde m é definido como momento de primeira ordem de $\hat{\nabla} f$ e $\widehat{m}$ é m depois da aplicação do fator de decaimento. Assim, o cálculo dos gradientes g são usados para normalização, e as Equações 2.18 e 2.19 são:

$$g_{t+1} = \gamma_{t+1} g_t + (1 - \gamma_{t+1}) \nabla \varphi(w_t)^2 \quad (2.18)$$

$$\hat{g}_{t+1} = \frac{g_{t+1}}{1 - \gamma_{t+1}} \quad (2.19)$$

g é definido como momento de segunda ordem de $\nabla\varphi$. Desta forma, a atualização dos parâmetros é então calculada pela seguinte Equação 2.20:

$$w_{t+1} = w_t - \eta \frac{\hat{m}_{t+1}}{\sqrt{\hat{g}_{t+1} + \epsilon}} \quad (2.20)$$

2.4 Redes Neurais Convolucionais

Os métodos de *machine learning* tradicionais usavam abordagens manuais baseadas em engenharia de recursos (AGGARWAL et al., 2018). A engenharia de recursos é o processo de criação de novos recursos a partir dos dados brutos e tem como objetivo aumentar a potência de predição do algoritmo de aprendizado. A extração de recursos visa tornar mais simples os modelos de *machine learning* a distinguirem entre classes diferentes (BENGIO et al., 2017). Contudo, os recursos são projetados à mão e isso torna o processo de treinamento uma tarefa difícil (GOODFELLOW et al., 2016). Assim, a concepção da CNN pode lidar com excelência com a extração de recursos que compõe a engenharia de recursos (BENGIO et al., 2017).

Neste sentido, os algoritmos baseados em CNN são uma ferramenta poderosa, pois são capazes de aprender recursos automaticamente (PATTERSON; GIBSON, 2017). Assim, tarefas como: reconhecimento de imagem e previsão são desempenhadas com excelência usando CNN (BENGIO et al., 2017). As arquiteturas CNN são projetadas para funcionar com entradas estruturadas em padrão de grade retangular, que possuem fortes dependências espaciais em regiões locais da grade retangular (GOODFELLOW et al., 2016). As imagens bidimensionais são um exemplo estruturado que exibe dependências espaciais, porque as localizações espaciais adjacentes, presentes na imagem, têm valores de cor semelhantes dos *pixels* individuais (AGGARWAL et al., 2018). Neste sentido, o número de unidades por camada aumentam de forma significativa. O que torna o modelo mais complexo que as arquiteturas ANN tradicionais (HAYKIN, 2010).

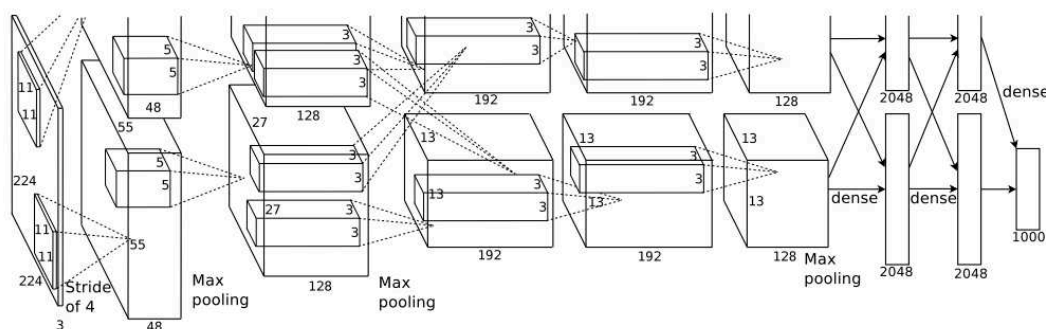
Em geral, a organização da CNN é formada por módulos ou blocos, constituídos por operações de convoluções, funções de ativação, funções de regularização, e camadas totalmente conectadas (GOODFELLOW et al., 2016). Os módulos são empilhados uns sobre os outros para formar um modelo profundo. Na arquitetura CNN, o fluxo da

informação ocorre somente em uma direção, ou seja, de suas entradas para suas saídas. Desta forma, vale destacar algumas vantagens ao utilizar uma arquitetura **CNN** como: (i) a capacidade de extrair características relevantes através de aprendizado de transformações (*kernels*) e (ii) o número reduzido de parâmetros de ajustes de redes totalmente conectadas com o mesmo número de camadas ocultas. Como cada unidade de uma camada não é conectada com todas as unidades da camada seguinte, há menos pesos para serem atualizados, facilitando assim o treinamento (BENGIO et al., 2017).

A arquitetura LeNet é considerada a primeira arquitetura **CNN** proposta por Yann LeCun no ano de 1988 (LECUN et al., 1998). A LeNet foi utilizada para o reconhecimento de caracteres, tais como código postal e dígitos numéricos. A arquitetura LeNet é composta por três camadas, *Kernels* convolucionais seguidos de operações de sub-amostragem com não linearidades na forma de tangente hiperbólica ou sigmoide (LECUN et al., 1998). As CPUs utilizadas eram lentas e o treinamento da rede se tornava complicado e despendia um grande período de tempo (LECUN et al., 1998). Praticamente não houve grande avanços entre 1998 e 2010.

O trabalho de (CIREŞAN et al., 2010) utilizou uma abordagem inovadora usando a primeira implementação de rede neural por meio de GPU para reconhecimento de dígitos manuscritos iniciando uma nova era de investigações no uso das arquiteturas **CNNs**. Assim, o uso da **CNN** foi marcado pelo trabalho de Alex Krizhevsky no ano de 2012 (KRIZHEVSKY et al., 2012). O desafio proposto no concurso ImageNet ILSVRC-2010 consistia em classificar 1,2 milhão de imagens de alta resolução em 1000 classes diferentes. A estratégia desenvolvida foi projetar uma **CNN** com 60 milhões de parâmetros e 650,000 neurônios distribuídas em cinco camadas convolucionais e três camadas totalmente conectadas, como mostra a Figura 2.15 (KRIZHEVSKY et al., 2012).

Figura 2.15: Arquitetura proposta no trabalho de Alex Krizhevsky

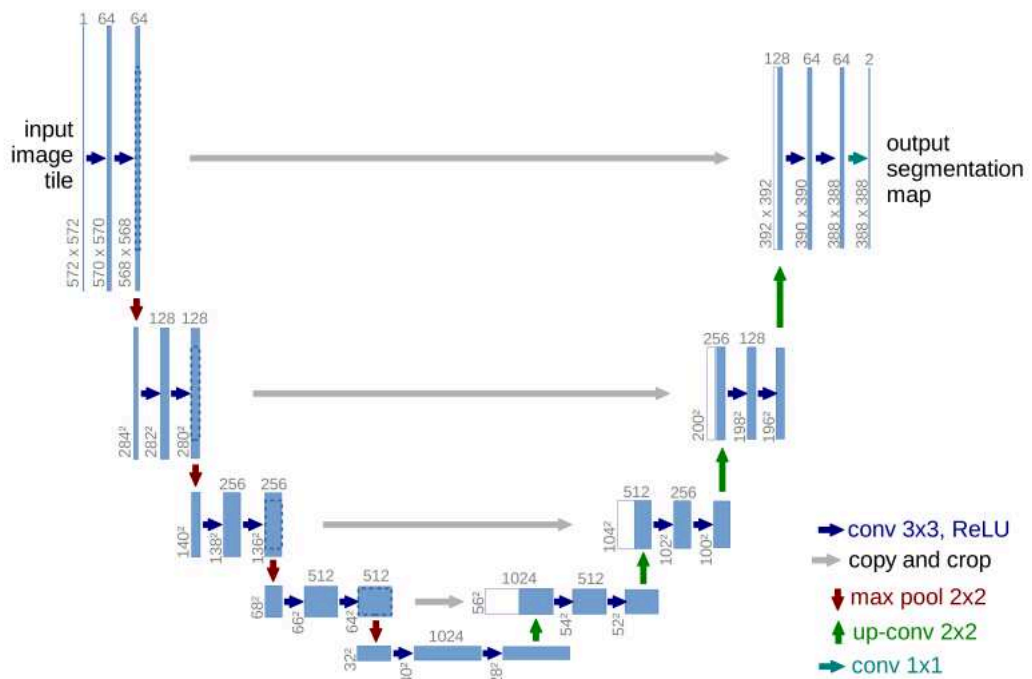


Fonte: (KRIZHEVSKY et al., 2012)

O treinamento da **CNN** é implementado em uma GPU que trabalha de forma eficiente a operação de convolução. Nas camadas totalmente conectadas foi proposto o uso de um método de regularização denominado de *dropout*. O aumento dos dados é realizado extraindo *patches* de forma aleatória e gerando translações e reflexões dos dados, e também é realizado o aumento alterando as intensidades dos canais RGB. Outro fator importante é a presença da função de ativação não linear *Relu*, para inserir a não linearidade durante o processo de treinamento da rede. Na saída da rede existe uma camada totalmente conectada que é alimentada por uma função *Softmax* que calcula a probabilidade logarítmica do rótulo correto para cada uma das mil classes. O resultado obtido da arquitetura alcançou uma taxa de erro de 15.3%, que é um resultado muito superior, ao melhor resultado obtido anteriormente de 26.2% (KRIZHEVSKY et al., 2012).

Uma evolução do trabalho de (CIREŞAN et al., 2012) foi proposto por (RONNEBERGER et al., 2015) vencendo a competição da *imageNet*. A proposta apresenta uma nova concepção de arquitetura profunda denominada de Unet capaz de lidar com bancos de dados pequenos e com poucos exemplos disponíveis. A arquitetura consiste em um caminho de contração para capturar as informações de contexto e um caminho de expansão simétrico que permite uma localização precisa com mostra a Figura 2.16.

Figura 2.16: Arquitetura Unet

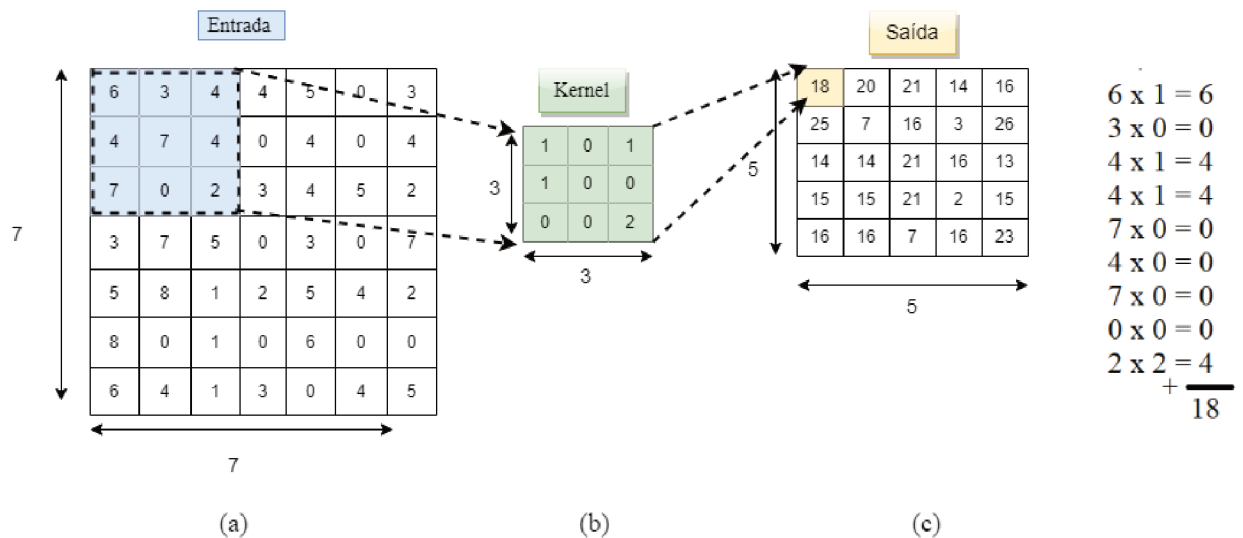


Fonte: (RONNEBERGER et al., 2015)

2.4.1 Convolução

As **CNNs** representam uma categoria de algoritmos baseados em **ANN**. A diferença principal entre **ANN** e **CNN** é a operação de convolução definida para **CNN**. A operação de convolução é uma operação matemática que envolve um conjunto de pesos estruturados em grade, denominados de *kernel*, e entradas estruturadas em grade extraídas de diferentes localidades espaciais do volume de entrada (**AGGARWAL et al., 2018**). Portanto, as **CNNs** são definidas como redes que usam a operação de convolução em pelo menos uma camada (**GOODFELLOW et al., 2016**). Entretanto, a grande maioria das arquiteturas **CNN** usam essa mesma operação de convolução em várias camadas sucessivas (**BENGIO et al., 2017**). A Figura 2.17 apresenta o processo de convolução entre uma imagem (a) e um *kernel* predefinido (b). A operação de convolução é um produto escalar sobre todo o volume do *kernel* que é repetido em todas as posições espaciais válidas (i,j) da imagem (a) e o *kernel* (b), e como resultado final a imagem de saída (c).

Figura 2.17: Exemplo da operação de convolução



Fonte: Elaborado pelo autor

A operação de convolução, exemplificada na Figura 2.17, pode ser expressa de forma matemática. Assim, seja $x(i,j)$ uma imagem bidimensional (ou um volume de entrada) em um espaço discreto e $w(i,j)$ um *kernel* (ou filtro) definido no mesmo espaço. A operação de convolução entre $x(i,j) * w(i,j)$ é definido na Equação 2.24 onde $y(i,j)$ é o resultado da operação de convolução. Em outras palavras, o *kernel* se comporta semelhante a uma pequena janela ou região que realiza operações ao longo da imagem. Assim, em cada passo, cada elemento do *kernel* é multiplicado pelo elemento que esta na mesma região que o

kernel está sendo aplicado naquele passo.

$$y[i, j] = x[i, j] * w[i, j] = \sum_m \sum_n x[m, n] w[i - m, j - n] \quad (2.21)$$

Uma propriedade extremamente importantes em relação a operação de convolução é o fato dela ser comutativa, ou seja, a ordem entre os termos podem ser alternados sem prejuízo. A operação é mostrada na Equação 2.22 (GOODFELLOW et al., 2016).

$$y[i, j] = w[i, j] * x[i, j] = \sum_m \sum_n x[i - m, j - n] w[m, n] \quad (2.22)$$

O sucesso das arquiteturas CNNs esta relacionada a sua capacidade de adaptar a rede com uma compreensão semântica do problema em investigação, ou seja, a semântica neste contexto se refere à forma como a imagem se relaciona às questões do mundo para ganhar significado próprio (BENGIO et al., 2017). A operação de convolução pode extrair recursos de forma hierárquica, ou seja, ela é capaz de codificar formas primitivas nas camadas iniciais (como bordas por exemplo) e formas complexas em camadas posteriores (combinações das bordas para obter formas mais complexas como curvas e cantos) (GOODFELLOW et al., 2016). Um detalhe importante nos projetos práticos usando CNN é que os *kernels* são inicializados de forma aleatórias (PATTERSON; GIBSON, 2017). Desta forma, o objetivo é que a rede aprenda os valores ideais dos *kernels* de forma automática. Isso significa na prática que a rede decide quais características são importantes para realizar a tarefa para qual foi projetada (GOODFELLOW et al., 2016).

Portanto, a operação de convolução pode ser considerada a parte mais importante de uma CNN (GONZALEZ; WOODS, 2009), (PEDRINI; SCHWARTZ, 2007), (GOODFELLOW et al., 2016). As Equações 2.23, 2.24 representam a convolução entre a imagem e um *kernel* de tamanho predefinido, em que o seu resultado gera vários mapas com as características de uma determinada região da imagem, extraindo assim informações importantes para o processamento digital (AGGARWAL et al., 2018). Além disso, a abordagem utilizando CNN contornam as questões do número elevado de parâmetros.

Na Figura 2.18 são apresentados dois exemplos em que a operação de convolução é aplicada a uma imagem e um *kernel* predefinido. Na primeira linha da Figura 2.18, uma imagem quadrada de tamanho 4×4 e um *kernel* de tamanho 3×3 são submetidos à operação de convolução. Observe que o resultado gera uma imagem com dimensões menores 2×2 . Uma formulação matemática para caracterizar o mapa de característica é

dada em respeito ao volume de entrada.

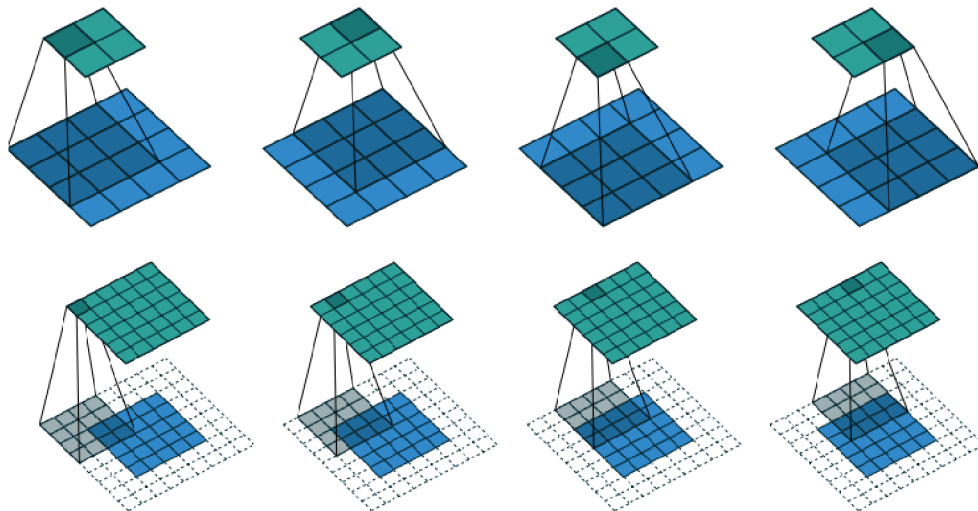
Contudo, três parâmetros controlam o tamanho da imagem resultante da camada convolucional: a profundidade (*depth*), o passo (*stride*) e preenchimento com zero (*padding*) (GOODFELLOW et al., 2016). A altura (AC) e a largura LC do volume convolucional são calculados utilizando as seguintes Equações 2.23 e 2.24, onde os parâmetros m e n são a altura e a largura da imagem de entrada, k é o tamanho do *kernel*, S é o tamanho do passo realizado pelo *kernel* e P é o *padding*.

$$AC = \frac{(m - k + 2P)}{S} + 1 \quad (2.23)$$

$$LC = \frac{(n - k + 2P)}{S} + 1 \quad (2.24)$$

Na segunda linha da Figura 2.18, uma imagem de tamanho 5×5 e um *kernel* de tamanho 4×4 são submetidos a operação de convolução. Neste caso foi aplicado o preenchimento com zeros na imagem. O objetivo deste preenchimento é que depois de realizada a operação de convolução a imagem mantenha ou aumente o seu tamanho original. Neste caso a imagem resultante apresenta uma dimensão a mais que a imagem de entrada resultando em uma imagem 6×6 .

Figura 2.18: Convolução entre a imagem e o *kernel* resultando em uma mapa de características. Na primeira linha, a imagem sem *padding* e a imagem resultante com dimensão reduzida. Na segunda linha, a imagem com preenchimento com *padding* e a imagem resultante com dimensão aumentada



Fonte: (DUMOULIN; VISIN, 2016)

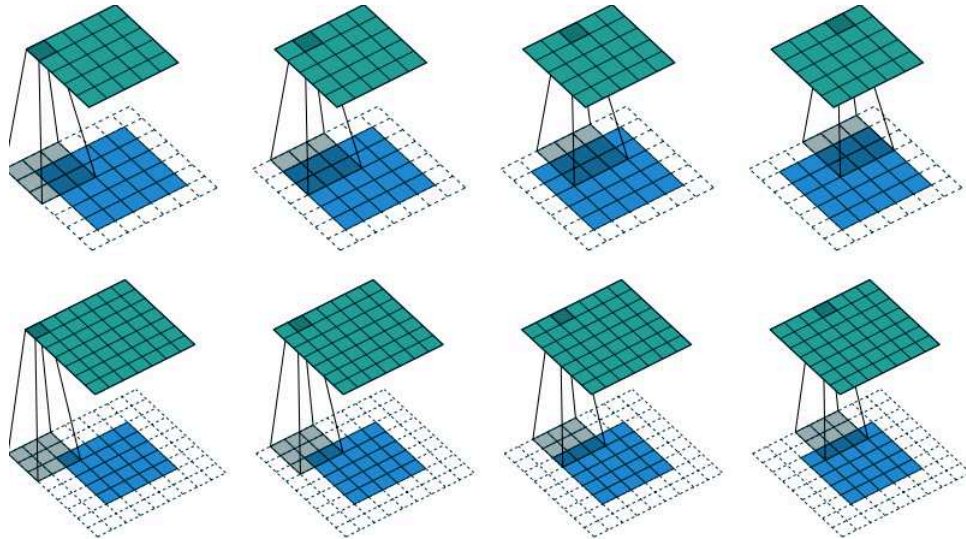
A Figura 2.19 apresenta mais dois exemplos de convolução com um *kernel* predefinido

de tamanho 3×3 e duas formas de preenchimento com *padding* e com o tamanho do passo $S = 1$. Os resultados destas operações aplicados são resumidas na Tabela 2.3.

Tabela 2.3: Exemplos da operação de convolução utilizando as fórmulas definidas nas equações 2.23 e 2.24

Dimensão da Imagem	Dimensão do <i>Kernel</i>	<i>zero-padding</i>	Dimensão da Imagem
5 x 5	3 x 3	1 x 1	5 x 5
5 x 5	3 x 3	2 x 2	7 x 7

Figura 2.19: Convolução entre a imagem e o *kernel* resultando em uma mapa de características. Na primeira linha, o padding usando para imagem é $P = 1$, $k = 3$ e $S = 1$, a imagem resultante da convolução têm a mesma dimensão da original. Na segunda linha, o padding $P = 2$, $k = 3$ e $S = 1$, a imagem resultante da covolução têm dimensão maior que a original.



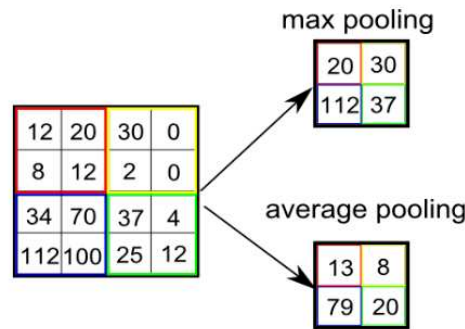
Fonte: (DUMOULIN; VISIN, 2016)

2.4.2 Pooling Layer

A estratégia denominada como *pooling* tem como objetivo reduzir o número de parâmetros de um mapa de características e acelerar a convergência da rede (GOODFELLOW et al., 2016). Neste sentido, o uso da operação de *pooling* é capaz de criar a invariância necessária para um modelo ANN ou CNN. A operação reduz o custo computacional, pois a medida que a profundidade (d) da rede aumenta é recomendado reduzir a dimensão espacial dos mapas de características (BENGIO et al., 2017). Assim, a operação pode ser resumida a partir de uma janela de tamanho $A \times A$ não sobreposta que é deslocada ao longo de toda imagem (AGGARWAL et al., 2018). Na janela é aplicado uma operação estatística (média, média ponderada, mediana, máximo ou mínimo) com relação aos elementos presentes. Na

estratégia denominada *Max-pooling* os valores dentro de uma região são analisados e o valor mais alto é escolhido dentro desta região. Uma estratégia semelhante é assumida na *Average-pooling* em que os valores que estão presentes na janela $A \times A$ são calculados e o valor médio é considerado (Figura 2.20). Assim, na operação de *pooling*, os valores de uma região são substituídos por alguma métrica. Portanto, o número de neurônios nos mapas de saída são reduzidos.

Figura 2.20: Estratégias de *Pooling* propostas para reduzir o conjunto de informações



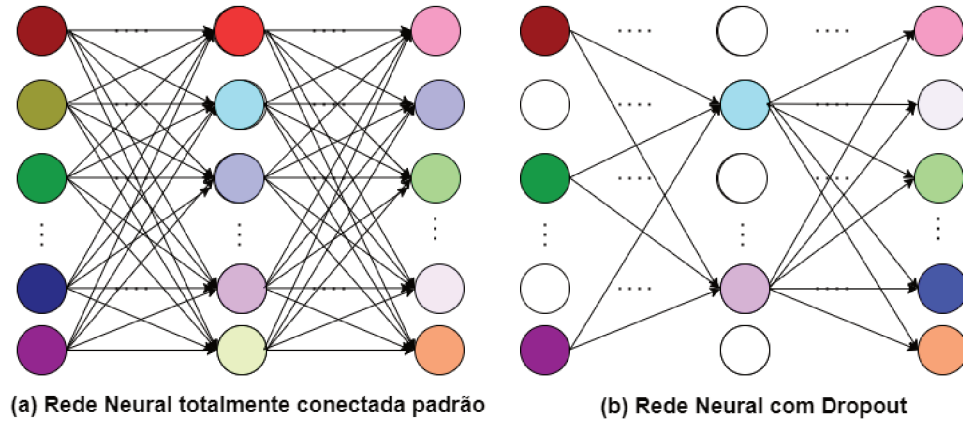
Fonte: (AL-BANDER et al., 2018)

2.4.3 Dropout

A técnica denominada de *Dropout* é computacionalmente barata, mas poderosa e capaz de regularizar uma ampla gama de funções projetadas para o aprendizado supervisionado (GOODFELLOW et al., 2016). Deste modo, o *dropout* treina o conjunto determinado de sub-redes que é formado a partir do momento que se remove de modo aleatório as unidade que estavam inicialmente conectadas. A Figura 2.21 evidência o efeito de aplicar a técnica de *dropout*. A Figura 2.21 (a) apresenta uma arquitetura genérica de rede totalmente conectada e na parte (b) a mesma rede com neurônios desativados.

2.4.4 Função Softmax

A camada final das arquiteturas ANN e CNN podem ser projetadas para realizar a tarefa de classificação (PATTERSON; GIBSON, 2017). Assim, a função *softmax* é utilizada para representar a probabilidade para cada classe presente na imagem. Então a função *softmax* é uma probabilidade exponencial normalizada de observações que atribui probabilidade da entrada x pertencer à classe I (BENGIO et al., 2017). Uma visão geral sobre a função *softmax* pode ser considerada como uma generalização da forma binária de

Figura 2.21: Efeito do *Dropout* na rede neural

Fonte: Elaborado pelo autor

regressão logística. Normalmente a função é uma generalização da tarefa de classificação multi-classe. Assim, a função pode ser dada pela Equação 2.25 como:

$$p(\sigma_i) = \frac{\exp^{z_i}}{\sum_{l=1}^L \exp^{z_l}} \quad (2.25)$$

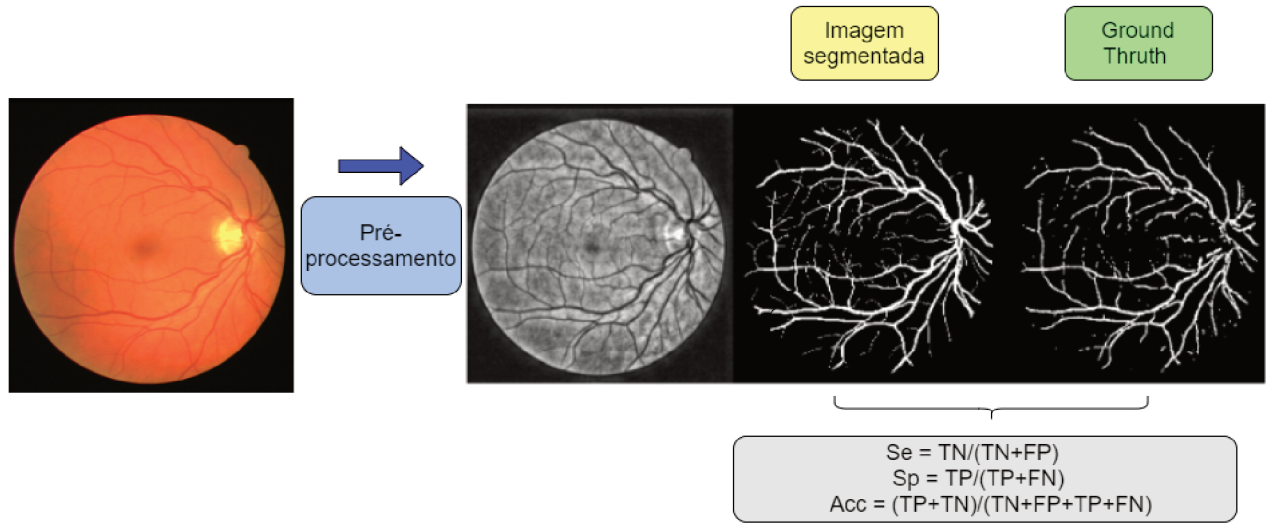
onde z é um vetor de entradas para camada de saída, e os valores que indexam os neurônios de saída, são $i = 1, 2, \dots, L$.

2.5 Métricas de Avaliação

A segmentação é uma das etapas importantes para concepção de sistemas automatizados (ABRÀMOFF et al., 2010). Uma etapa indispensável é avaliar o desempenho da segmentação realizada por um médico especialista, e a resposta do algoritmo proposto para realizar a segmentação. A Figura 2.22 ilustra o uso das métricas para uma imagem de fundo de olho e seu *ground truth*¹ (GOODFELLOW et al., 2016).

Assim, a Tabela 2.4 resume as possibilidades quando sobrepomos as duas imagens.

¹marcação realizado por um médico especialista em retina.

Figura 2.22: Aplicação das métricas uma imagem de fundo de olho.

Elaborado pelo autor

Tabela 2.4: Anotações realizadas por um especialista.

Classificação	Descrição
<i>True Positive</i> (TP) Verdadeiro positivo	pixels classificados pelo método e pelo especialista como pertencentes ao objeto.
<i>True Negative</i> (TN) Verdadeiro negativo	pixels classificados pelo método e pelo especialista como não pertencentes ao objeto.
<i>False Positive</i> (FP) Falso positivo	<i>pixels</i> classificados pelo método como pertencentes ao objeto e classificados pelo especialista como não pertencentes ao objeto.
<i>False Negative</i> (FN) Falsos negativo	<i>pixels</i> classificados pelo método como não pertencentes ao objeto e classificados pelo especialista como pertencentes ao objeto.

Com base neste conjunto de medidas é possível formular e estimar de forma matemática os diferentes parâmetros de desempenho, como sensibilidade (Se), especificidade (Sp), a precisão (Acc), Recall, F1 e área sob a curva (AUC) (ABRÀMOFF et al., 2010). Suas definições e fórmulas serão exemplificadas a seguir.

- *Sensitivity* (Se) - sensibilidade (Se) é a proporção de verdadeiros positivos, sua intenção é quantificar a capacidade do método em detectar o alvo. Normalmente é utilizada em conjunto com a especificidade. Pode ser definida na Equação 2.26:

$$Se = \frac{TP}{TP + FN} \quad (2.26)$$

Especificity (Sp) - especificidade, também designada por *precision*, é a proporção de

verdadeiros negativos, ou seja, é a capacidade do método em encontrar os *pixels* não pertencentes ao alvo. Assim pode ser definida na Equação 2.27:

$$Sp = \frac{TN}{TN + FP} \quad (2.27)$$

- *Accuracy* (**Acc**) - *accuracy* é o grau de proximidade das medições de uma quantidade com o valor real das quantidades. As medidas de verdadeiro positivos e verdadeiro negativos (**TP** e **TN**) refletem o total de acertos na segmentação:

$$Acc = \frac{TP + TN}{TN + FP + TP + FN} \quad (2.28)$$

- *Receiver Operating Characteristic* (**ROC**) - é a curva que plota a fração de *pixels* do alvo corretamente classificados (True Positive Rate (**TPR**)) versus a fração de *pixels* erroneamente classificados como pertencentes ao alvo (False Positive Rate (**FPR**));
- Area Under the ROC Curve (**AUC**) - é a área sob a curva **ROC**;
- *F-score* (**F**)² - é a média harmônica entre a especificidade e sensibilidade³:

$$F = 2 \frac{SP \cdot SN}{SP + SN} \quad (2.29)$$

- *Dice Similarity Measure* (**DSM**) - uma reformulação para a F-score:

$$DSM = \frac{2||TP||}{2||TP|| + ||FN|| + ||FP||} \quad (2.30)$$

- *Overlapping score* (**S**) - é a taxa de sobreposição entre os pixels detectados pelo método e os pixels classificados pelo especialista:

$$S = \frac{TP}{FP + TP + FN} \quad (2.31)$$

²Também chamada de *F₁-score* ou *F-measure*.

³Normalmente é definida em termos de *precision* e *recall*, o que é apenas uma questão de terminologia

2.6 Considerações finais deste capítulo

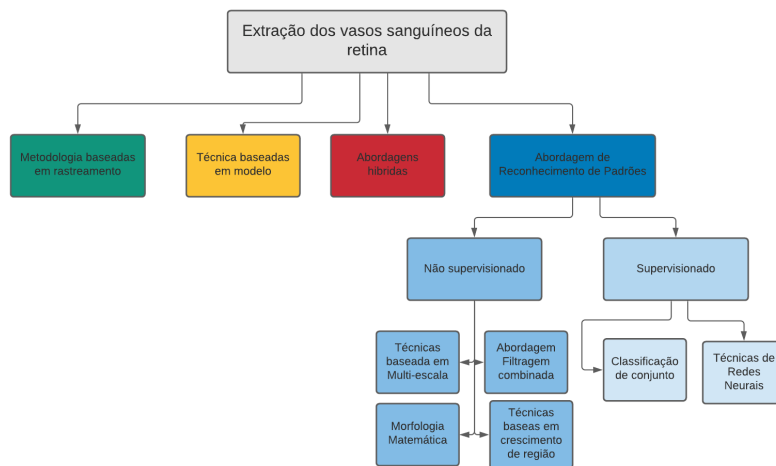
Este capítulo apresentou os principais conceitos para o desenvolvimento desta tese. Abordamos assuntos como anatomia da retina, inteligência artificial, ANN, CNN, técnicas de otimização e métricas de avaliação. Todos os tópicos são muito importante para desenvolver algoritmos automáticos para a tarefa de segmentação dos vasos sanguíneos. Em geral, os algoritmos ANN e CNN são o estado da arte em diferentes áreas de pesquisa.

CAPÍTULO 3

TRABALHOS RELACIONADOS

Neste capítulo é apresentada uma revisão dos trabalhos relacionados a segmentação dos vasos sanguíneos em imagens de fundo de olho. Destacamos que este estudo não visa esgotar o assunto por completo, mas sim realizar uma revisão consistente sobre os trabalhos relacionados à tarefa de segmentação. Muitas pesquisas classificam de maneiras diferentes os algoritmos e metodologias propostos para realizar a tarefa de segmentação dos vasos sanguíneos. Assim, os artigos são categorizados em: metodologia baseada em rastreamento, técnicas baseadas e modelo, abordagens híbridas, abordagem baseada em reconhecimento de padrões supervisionados e não supervisionados. A Figura 3.1 mostra um diagrama geral com as principais metodologias que tem como objetivo segmentar os vasos sanguíneos em imagens do fundo de olho.

Figura 3.1: Diagrama - Diagrama geral para classificação dos métodos que segmentam os vasos sanguíneos



Fonte: Elaborado pelo autor

3.1 Metodologia Baseada em Rastreamento

As metodologias de rastreamento atuam de modo recursivo para encontrar uma aproximação da rede vascular. O rastreio dos vasos sanguíneos pode ser realizado com base na seleção manual ou automática de algumas posições estratégicas denominadas de sementes (*seed*). Desta forma, um pixel pertencente ao vaso deve ser determinado e um processamento recursivo baseado em alguma técnica (como crescimento de regiões) agrega os novos *pixels* pertencentes a estrutura vascular (KHAN et al., 2019).

Uma das vantagens deste tipo de técnica é a sua competência, pois somente os *pixels* próximos aos pontos preliminares são inspecionados e avaliados. Outra vantagem são as informações estatísticas detectadas a partir da aproximação da rede vascular, como largura dos vasos e padrão de bifurcações. Entretanto, se os padrões de bifurcações ou limites não estiverem bem definidos, o processo tende a terminar nestas posições, e assim a obtenção da estrutura vascular fica incompleta.

Neste sentido, um trabalho pioneiro que usa a metodologia para rastrear recursivamente os vasos em imagens de angiogramas na modalidade de raio-x foi proposto por (LIU; SUN, 1993). O método proposto é verificado em três estágios. No primeiro, uma posição inicial e uma direção são escolhidas, e assim um segmento da estrutura vascular é determinado. No segundo estágio, a área onde os vasos se sobrepõem são excluídas para evitar o problema de reentrada dos mesmos segmentos. No estágio final, os pontos de ramificação são detectados usando a filtragem combinada (*matched filtering*) ao longo das bordas dos vasos.

Outra metodologia para o rastreamento automático dos vasos sanguíneos foi proposta por (ZHOU et al., 1994). As imagens de angiogramas são utilizadas para identificar e quantificar os segmentos dos vasos sanguíneos. A filtragem combinada em conjunto com o conhecimento a priori e as propriedades dos vasos sanguíneos permite detectar de forma automática os limites dos vasos e simultaneamente extrair parâmetros de interesse clínico.

No trabalho (XU et al., 2011) é proposto um algoritmo baseado em gráfico para medir a largura dos vasos sanguíneos da retina. O objetivo principal do trabalho é apresentar um método capaz de determinar os limites dos vasos sanguíneos com base na análise gráfica. Desta forma, as respostas são agrupadas em dois conjuntos. O primeiro contém a linha central do vaso, que é derivada da estrutura vascular. No segundo, as duas bordas do vaso são segmentadas por meio da técnica de derivadas da intensidade de *pixels*.

Posteriormente, os *pixels* obtidos nestes processos compõem um problema de segmentação de superfície 3D em duas fatias.

No trabalho (FRAZ et al., 2012a), uma combinação de técnicas é utilizada para desenvolver um método automático de segmentação dos vasos sanguíneos. As linhas centrais dos vasos são extraídas usando a derivada de primeira ordem de um filtro gaussiano com quatro orientações. Um operador morfológico *top-hat* com um elemento estruturante linear melhora a imagem em escala de cinza.

3.2 Técnicas Baseadas em Modelo

As técnicas baseados em modelos tentam incorporar a morfologia da estrutura vascular existente, interpretando limites de bordas, perfil de alongamento e reflexo axial. Assim, o trabalho proposto por (VERMEER et al., 2004) utiliza um método baseado em modelo para realizar a segmentação dos vasos sanguíneos. Desta forma o desempenho do método baseado em modelo é dividido em três etapas. Na primeira é utilizado a transformada de Laplace para o aprimoramento, em seguida uma segmentação baseada no limite, por fim, uma categorização.

Em outra abordagem, o trabalho (MAHADEVAN et al., 2004) apresenta um conjunto de três algoritmos para realizar a segmentação dos vasos sanguíneos em imagens retinianas de vídeo. O primeiro é baseado em verossimilhança, o segundo em estatística de teste *alpha* cortado, o terceiro na seleção de modelos robustos. Os algoritmos considerados são dotados de um modelo matemático para medir as principais características dos vasos.

Em outras perspectiva, o método proposto por (LAM; YAN, 2008) tem como principal objetivo detectar os vasos sanguíneos em imagens patológicas da retina. A estratégia utiliza um operador Laplaciano capaz de extrair objetos semelhantes a vasos sanguíneos, e campo vetorial normalizado para reduzir o nível de ruído presente nos vasos.

Na abordagem de (DELIBASIS et al., 2010) é proposto um algoritmo automático capaz de segmentar amplamente a estrutura vascular e medir o diâmetro e orientação dos vasos sanguíneos da retina. O vaso sanguíneo é modelado por meio de um modelo paramétrico que quantifica o modelo do vaso que corresponde a imagem de um angiograma.

Outra estratégia para a segmentação dos vasos sanguíneos é proposta no trabalho de (KOVÁCS; HAJDU, 2016) utilizando dois métodos, o casamento de modelo e a reconstrução de contorno. Operadores de segmentação baseados em função Gabor são

usados para detectar as linhas médias dos vasos sanguíneos e na sequência as características da intensidade e forma dos vasos são calculados em um conjunto de dados de treinamento.

3.3 Abordagem Híbridas

Um conjunto de novos métodos que fazem uso de várias técnicas ou classificadores para realizar uma tarefa específica são denominados como métodos híbridos. Este tipo de abordagem vem sendo utilizado em diferentes cenários de pesquisa e tem obtido bons resultados. Desta forma, uma abordagem híbrida diferenciada proposta por (CONDURACHE; MERTINS, 2012) segmenta os vasos sanguíneos utilizando um esquema *hysteresis-classifier*. A estratégia para obter uma boa segmentação é composta por dois classificadores, um classificador pessimista e um classificador otimista. No primeiro, a taxa FPR é considerada quase nula, assim as classes sobrepostas apresentam uma alta FNR. No segundo, a taxa de falsos negativo próxima de zero e consequentemente a taxa de falsos positivos é alta. Desta forma, é possível utilizar as classificações otimistas e pessimista para determinar a estrutura vascular. O resultado da metodologia é aplicado nas bases DRIVE e STARE.

Outra algoritmo híbrido proposto por (DHARMAWAN et al., 2019) apresenta um método de realce dos vasos sensível a mudanças de direções. Assim, as imagens geradas desta etapa são entregues a uma rede neural totalmente convolucional derivada da Unet, que realiza a tarefa de extração dos mapas de características e treina a rede para segmentar os vasos sanguíneos da retina.

Seguindo uma outra abordagem, o trabalho de (VILLALOBOS-CASTALDI et al., 2010) implementa uma técnica que depende do resultado da segunda entropia local e da matriz de concorrência em níveis de cinza para extrair a estrutura vascular presentes nas imagens da retina. Um diferencial desta abordagem é a determinação de limiares automaticamente, com base apenas nas informações apresentadas pela imagem original em avaliação.

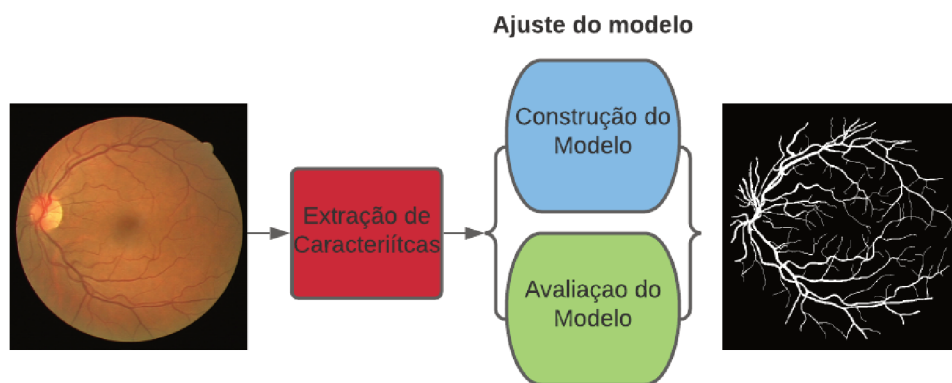
No trabalho (SALAZAR-GONZALEZ et al., 2014) a segmentação de duas estruturas anatômicas importantes, como disco óptico e vasos sanguíneos são elaboradas. Primeiro, a estrutura vascular é inicialmente segmentada utilizando *graph cut technique*. Em seguida um método de reconstrução de imagens *Markov Random Field* (MRF) é utilizado para descartar os vasos que estão no interior do disco óptico e um método do fator de compensação segmenta o disco óptico baseada na intensidade local dos vasos.

3.4 Abordagem de Reconhecimento de Padrões não supervisionada

A abordagem utilizando as técnicas de reconhecimento de padrão são abundantes, os principais objetivos são trabalhar na identificação dos padrões e na consistência das informações presentes nos dados. Desta forma, as técnicas de reconhecimento de padrões no cenário da segmentação dos vasos sanguíneos lida com os atributos relacionados aos vasos, que podem ser morfológicos ou geométricos. De modo geral, as técnicas podem ser divididas em dois métodos: supervisionadas e não supervisionadas (KHAN et al., 2019).

Os algoritmos derivados da abordagem não supervisionada têm a capacidade de criar modelos a partir das imagens não rotuladas do banco de imagens escolhido. Assim informações com intensidade local e gradiente são utilizadas para construção dos seus modelos. Na fase de construção do modelo de segmentação a qualidade é avaliada por meio de uma função de minimização, elaborada com a finalidade de encontrar a melhor separação entre as classes consideradas. Para obter a separação entre as classes geralmente a função é definida com base em métricas euclidiana ou métricas probabilísticas. A Figura 3.2 representa o diagrama geral deste tipo de metodologia.

Figura 3.2: Visão geral das abordagens que utilizam algoritmos não supervisionados para tarefa de segmentação de imagens de fundo de olho



Fonte: Elaborado pelo autor

3.4.1 Técnicas baseadas em Multiescala

A extração dos vasos pode ser obtida com uso de técnicas multiescala, que são baseadas na variação das escalas da imagem (ALMOTIRI et al., 2018). Consequentemente,

a cada imagem de entrada é possível extrair diferentes estruturas com base nesta variação das escalas da imagem (KHAN et al., 2019). Outro ponto importante é que este tipo de técnica permite uma redução no tempo de execução. Os elementos importantes relacionados as imagens de fundo de olho, como as artérias mais visíveis podem ser extraídas facilmente em imagens com um bom contraste. No entanto, os vasos finos são afetados por baixo contraste e ruídos que se assemelham as informações dos vasos finos (ALMOTIRI et al., 2018). De modo geral, os métodos múltiplas escalas podem resolver de modo positivo os problemas relacionados a largura do vaso em imagem do fundo de olho.

Outro método baseado em múltipla escala foi proposto por (FRANGI et al., 1998). O aprimoramento dos vasos é obtido por meio de atributos geométricos. As fases de filtragem tem dois objetivos: procurar por atributos geométricos e aumentar o contraste dos mesmos. A matriz hessiana da imagem são os valores próprios extraídos para melhoria do contraste. A mudança de escala é usada para obter os parâmetros estruturais dos vasos.

A segmentação da estrutura vascular é obtida a partir de um algoritmo multiescala determinístico proposto por (VLACHOS; DERMATAS, 2010). A estratégia baseada em multiescala facilita a detecção dos vasos que apresentam diferentes diâmetros.

Na mesma linha de investigação o trabalho proposto por (LI et al., 2012) denominado de *Multiscale Production of the Matched Filter* (MPMF) elabora um esquema de extração dos vasos sanguíneos em três escalas. O resultado do processo de filtragem possui o destaque forte em relação aos vasos, e um resultado menos evidente em relação ao ruído de fundo.

No trabalho de (WANG et al., 2016) é construído um vetor com trinta características de cada pixel da imagem da retina. As características que constituem os vetores são extraídas utilizando *Gaussian Scale Space*, *Multi-Scale Gabor Filter* e *Divergence of Vector Fields*. Os vasos sanguíneos são classificados por um classificador *Support Vector Machine* (SVM), com base no vetor de característica e nas informações da segmentação manual.

O método proposto por (NGUYEN et al., 2013) se concentra no fato que o comprimento de um detector de linha em escalas variadas pode detectar a linha individual dos vasos em diferentes escalas. Assim, é possível compor um método para extração automática dos vasos sanguíneos em imagens coloridas da retina. Por fim, as imagens das respostas de linha em escalas diferentes são combinadas linearmente para produzir a segmentação final.

No trabalho de (FATHI; NAGHSH-NILCHI, 2013) uma técnica multiescala para aprimorar os vasos e uma transformada *wavelet* contínua complexa (*Complex Continuous Wavelet Transform* (CCWT)) é usada para segmentação automática dos vasos sanguíneos. A CCWT composta da parte real e imaginária é capaz de diferenciar as estruturas alongadas e os contornos dos vasos sanguíneos presentes nas imagens da retina. A segmentação final é obtida integrando os valores do módulo dos coeficientes *wavelet* nas escalas adequadas em conjunto com um método de limiar adaptativo.

3.4.2 Morfologia Matemática

Morfologia matemática em um contexto geral tem como objetivo varrer toda imagem com um elemento denominado estruturante, e assim realizar pequenas ou grandes correções na aparência global da imagem. O elemento estruturante é capaz de armazenar uma forma predefinida enquanto modifica formas presentes na imagem. Assim, uma das aplicações dos filtros morfológicos é reduzir o nível de ruído, seja ele linear ou não linear (GONZALEZ; WOODS, 2009). Geralmente, a morfologia matemática é aplicada em imagens binárias e podem ser estendidas para imagens em escala de cinza. Considerando uma abordagem matemática, uma imagem I é representada como um conjunto $I \in R^2$, onde os *pixels* em primeiro plano pertencem a imagem I , e os *pixels* em segundo plano pertencem a imagem I^c (PEDRINI; SCHWARTZ, 2007). Desta forma, as operações matemáticas podem ser resumidas em: operações de erosão, dilatação, abertura e fechamento. Cada operação é utilizada com algum objetivo específico para destacar elementos que estão presentes na imagem. Assim, a partir da escolha do elemento estruturante as operações modificam a imagem.

Na abordagem proposta por (ZANA; KLEIN, 2001) o uso de morfologia matemática e a extração de curvatura é utilizada para extrair estruturas semelhantes aos vasos. Os dois principais focos são extrair por meio de operações morfológicas os vasos protuberantes com base em sua morfologia, e avaliar a curvatura cruzada para destacar os *pixels* dos vasos e do fundo da imagem.

Outra abordagem que combina filtros diferenciais, para extração das linhas centrais, e operadores morfológicos para preenchimento dos vasos sanguíneos é proposto por (??). O algoritmo tem como objetivo realizar a segmentação automatizada dos vasos sanguíneos em três passos. Primeiro, um banco de filtros de detecção de linha com diferentes orientações

aprimora os vasos finos em uma etapa de pré-processamento. Segundo, a extração da linha central é obtida para definir um conjunto de segmentos conectados. Por fim, a segmentação dos vasos rotula os *pixels* que pertence aos vasos.

Em outro trabalho (ROSSANT et al., 2011) foi proposto uma metodologia que considera à análise morfológica e topológica dos vasos sanguíneos para construir um algoritmo de segmentação. A estratégia da abordagem tem como objetivo extrair duas imagens com contraste aprimorado, obtido por um filtro morfológico e uma transformada *top hat*. A segmentação resultante é obtida por um processo de limiarização da fusão das imagens resultantes.

3.4.3 Abordagem de Filtragem combinada

As técnicas baseadas em núcleo (*Kernel-based* ou *matched filter*) têm como objetivo segmentar os vasos sanguíneos utilizando a distribuição das intensidades dos *pixels* que representam os vasos sanguíneos e assim construir um núcleo do *kernel* para detectar os limites da estrutura vascular. O *kernel* pode ser representado por uma forma matemática específica ou ser deformável de acordo com os limites dos vasos. A ideia principal é obter mapas representativos obtidos da distribuição de intensidades dos vasos.

Entretanto, o método proposto por (VILLALOBOS-CASTALDI et al., 2010) é composto de cinco etapas para extrair os vasos. Na primeira etapa, as informações do canal verde são utilizadas. Na segunda, uma máscara automática para evitar o processamento fora da parte de interesse é projetada. Na terceira, a aplicação do *matched filter* para realçar as bordas das imagens da retina. Na quarta é executado o cálculo da matriz de coocorrência. Por fim, a segmentação automática dos vasos sanguíneos é obtida por meio da entropia de segunda ordem. Dentro da mesma perspectiva, o trabalho (MA; LI, 2015) considera a assimetria parcial e a variação da iluminação nas imagens da retina. Assim, um modelo gaussiano composto por quatro partes é utilizado para ajustar o perfil de distribuição da intensidade da secção transversal do vaso.

No trabalho desenvolvido por (SINGH; SRIVASTAVA, 2016) o desempenho da segmentação dos vasos sanguíneos é melhorado a partir de uma nova abordagem que combina *matched filter* e uma função de distribuição de probabilidade *Gumbel*. Na fase que antecede a aplicação do *matched filter* as imagens são convertidas em escala de cinza utilizando análise de componentes principais PCA e na sequência o contraste da imagem é

equilibrado usando **CLAHE**. Em seguida os parâmetros do *matched filter* são determinados de forma empírica e na fase final a escolha de um limiar baseado em entropia e filtragem do comprimento dos vasos é usado para obter a imagem segmentada final.

3.4.4 Técnica baseadas em crescimento de região

Os métodos baseados na teoria do crescimento de região tentam agrupar os *pixels* que apresentam características semelhantes com objetivo de formar uma nova região que agrupa este conjunto de *pixels*. Desta forma, o bom desempenho deste tipo de método é baseado na localização dos pontos sementes e nas condições de crescimento. Assim, os algoritmos baseados neste tipo de abordagem segmentam os vasos sanguíneos de modo manual ou automático. A estratégia para detectar utiliza uma semente (*seed*) para extrair a linha central. Uma das vantagens desta estratégia é analisar os *pixels* próximos aos pontos preliminares. Contudo existe uma grande diversidade na anatomia dos vasos e suas características morfológicas e outras características relacionadas a qualidade da imagem da retina. Neste sentido, existem duas linhas de investigação que podem ser seguidas. Na primeira linha de investigação a morfologia dos vasos é considerada, e problemas como o tamanho pequeno dos vasos, o padrão de ramificação, a tortuosidade e o estreitamento dos vasos são os objetivos a serem alcançados.

Na segunda classe de problemas é considerado as características globais da imagem, como baixo contraste, ruídos, artefatos entre outros (**JIA; ZHUANG, 2021**). Uma característica importante para imagens médicas de fundo de olho é que os algoritmos deste tipo não requerem que o ponto semente esteja localizado no centro do vaso sanguíneo. Neste sentido, o trabalho de (**CHUTATAPE et al., 2002**) considera a extração dos pontos sementes em duas fases. Na primeira fase, os pontos sementes que se aproximam da circunferência que representa o contorno do disco óptico são extraídos. Na sequência, os centros dos vasos são marcados usando um filtro de *Kalman* estendido. Por fim, a estrutura dos vasos é obtida.

Em uma linha semelhante de investigação, o trabalho (**YEDIDYA; HARTLEY, 2008**) aborda a segmentação dos vasos sanguíneos do fundo de olho usando o filtro *Kalman* para rastrear os centros dos vasos. Assim, o filtro tem a capacidade de detectar vasos largos e finos definindo um modelo linear, mesmo em condições adversas em que as imagens são ruidosas. Portanto, a metodologia é composta de quatro fases. Na primeira, um conjunto

de pontos sementes que abrange toda imagem e usa filtros correspondentes (*matched filter*) em diferentes larguras e orientações. Na segunda, o rastreio do centro do vaso sanguíneo com o filtro *Kalman*, a partir dos pontos sementes encontradas na fase anterior. No terceiro, o rastreamento (*tracing*) é parado se a probabilidade de rastrear um vaso for pequena dentro de um número de movimentos consecutivos. Por fim, a segmentação é reiniciada se o rastreamento falha.

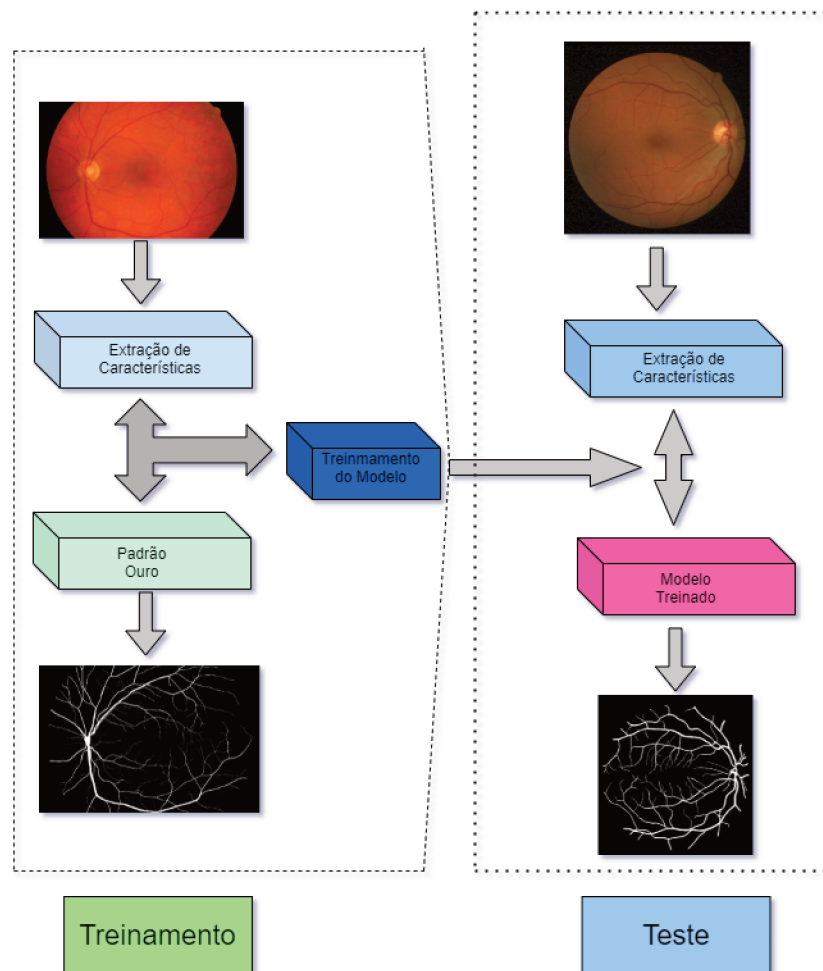
Outra estratégia proposta por (WU et al., 2007) apresenta uma metodologia de rastreamento dos vasos (*vessel tracking*) com objetivo de extrair os vasos sanguíneos da retina. A estratégia combina a saída das imagens pré-processadas por um filtro de difusão não linear, que suaviza as imagens. Na sequência um filtro de aprimoramento combina as vantagens obtidas com o *Hessian-based filters* e *matched filter*, que agregam as restrições das bordas dos vasos sanguíneos. Por fim, as cristas são obtidas por meio das orientações de tamanho, a partir do realce da imagem, que deixa mais evidente a diferença entre os vasos sanguíneos e o fundo da imagem.

Em uma abordagem diferenciada no trabalho de (LI et al., 2013) o rastreamento dos vasos foi investigado usando uma abordagem estatística de probabilidade máxima a posteriori. Na primeira parte do método o disco óptico é detectado por meio do algoritmo *Principal Analysis Component* (PCA). Em seguida, os pontos que pertencem a extremidade dos vasos podem ser obtidos de modo interativo utilizando estatísticas locais dos níveis de cinza e propriedades de continuidade dos vasos sanguíneos. No próximo passo o perfil do nível de cinza e suas propriedades geométricas melhoram a precisão do rastreamento. Na abordagem proposta por (FADZIL et al., 2007) é apresentado o método para detecção e reconstrução dos vasos sanguíneos da retina. O método tem duas fases importantes. Na primeira fase, a detecção das estruturas dos vasos sanguíneos é alcançada utilizando o algoritmo CLAHE combinada com extração dos vasos por meio de uma transformada morfológica *bottom-hat* (*bottom-hat morphological transformation*). Para a fase de reconstrução dos vasos, uma técnica de crescimento de região baseada na derivada de primeira ordem da função gaussiana que modelo o perfil da estrutura dos vasos.

3.5 Abordagem de Reconhecimento de Padrões supervisionada

As abordagens de reconhecimento de padrões supervisionados pertencem a grande área denominada como *machine learning*. Os modelos desenvolvidos neste tipo de abordagem precisam de um conjunto de informações já rotuladas para extrair os recursos necessários para aprender como novos modelos. De modo geral, o desenvolvimento dos algoritmos e métodos seguem os passos apresentados na Figura 3.3 para desenvolver suas aplicações, sejam elas segmentação, detecção ou classificação.

Figura 3.3: Visão geral das abordagens que utilizam algoritmos supervisionados para tarefa de segmentação de imagens de fundo de olho



Fonte: Elaborado pelo autor

3.5.1 Técnicas Baseadas em Redes neurais

As **ANNs** são métodos que pertencem ao area de *machine learning*. Assim, a implementação de algoritmos de segmentação precisa de dados rotulados para criar um modelo de representação. Nas imagens de fundo de olho as principais características são rotuladas por um médico especialista em retina (**ABRÀMOFF et al., 2010**). Assim, as imagens originais e as imagens rotuladas servem para o desenvolvimento de algoritmos de aprendizagem supervisionados (**BENGIO et al., 2017**). Muitos trabalhos usam **ANN** ou **CNN** para extrair características relacionadas aos detalhes anatômicos que estão presentes nas imagens de fundo de olho (**ABDULLAH et al., 2016**).

Na abordagem proposta por (**JIN et al., 2020**) cinco modelos pré-treinados *LeNet*, *AlexNet*, *ZF-Net*, *VGG* e *Deformable-ConvNet* são modificados para realizar a segmentação dos vasos sanguíneos. No primeiro passo, as arquiteturas **CNNs** são pré-processadas no canal verde e as correções usando **CLAHE** e correção gama aprimoraram a qualidade da imagem. Em seguida, um conjunto de *patches* é ajustado para todos os modelos e a nova rede é treinada para a segmentação dos vasos sanguíneos.

Em outra perspectiva o trabalho (**SON et al., 2019**) analisa o desempenho estatístico para arquitetura denominada de *Generative adversarial Networks* (**GAN**) para realizar a tarefa de segmentação dos vasos sanguíneos. Na arquitetura **GAN** existem duas sub-redes, uma que discrimina e tenta descobrir o padrão ouro, e outra denominada de geradora, que têm objetivo de gerar saídas mais realistas possíveis. No trabalho de (**GUO et al., 2019**) a segmentação dos vasos sanguíneos é proposto utilizando uma rede multi-escala profundamente supervisionada com conexões curtas denominada de (*BTS-DSN*). O principal objetivo das conexões curtas e transferir informações de baixo nível para alto nível.

No trabalho de (**RIBEIRO et al., 2019**) a segmentação dos vasos sanguíneos é abordada utilizando um modelo *Unet* e duas técnicas de *ensemble*, uma denominada de *Stochastic Weight Averaging* e a outra *Snapshot Ensembles*. A proposta de (**FENG et al., 2020**) considera uma **CNN** de conexão cruzada (*CcNet*) que é elaborada para realizar a tarefa de segmentação automática. As imagens do canal verde são consideradas para as etapas de treinamento e teste do modelo, e as camadas de convolução são capazes de extrair e prever as classes de *pixels* relacionadas aos vasos. Em (**NOH et al., 2019**) uma nova proposta de **CNN** que incorpora a teoria do espaço-escala é proposta para

segmentar os vasos sanguíneos. A estratégia da rede denominada de *SSANet* é combinar as etapas de *upsampling* e *down-sampling* em conjunto com os blocos residuais. Uma nova arquitetura denominada de *ReNet* é proposta para segmentação dos vasos sanguíneos (LU et al., 2019). A arquitetura é baseada em *Encoder-decoder* totalmente convolucional. Na parte de *encoder* os recursos hierárquicos são extraídos. Na parte do *decoder* é usado para reestruturar a entrada da rede. A arquitetura proposta trabalha com duas conexões de atalho que capturam as informações semânticas e podem superar as estratégias de pré-processamentos.

Outra estratégia que utiliza rede neural baseada em múltiplos níveis que visa segmentar os vasos sanguíneos foi proposta por (NGO; HAN, 2017). Para obter uma melhoria no processo de segmentação é proposta duas técnicas, o *dropout* e o *dropout* espacial. Essas combinações visam fornecer a rede neural profunda mais informações nos mapas de características. Uma proposta que usa De-noising Auto-Encoders (DAEs) combinada com rede neural profunda foi avaliada por (FAN; MO, 2016). Nesta abordagem, considerada multi-nível, para cada parte da imagem da retina, o mapa dos vasos é aprendido a partir de uma função de perda que guia o processo de aprendizado.

3.5.2 Técnicas baseadas em Classificação de Conjunto

Uma estratégia com objetivo de reunir diferentes classificadores é denominada de classificação baseada em conjuntos (*Ensemble Classification-Based*). O principal objetivo deste tipo de classificação é combinar vários classificadores para tratar uma tarefa específica da área de *machine learning* (BENGIO et al., 2017). Neste sentido, a segmentação supervisionado dos vasos sanguíneos foi proposta no trabalho de (FRAZ et al., 2012b). A metodologia proposta utiliza um sistema de conjuntos baseados em árvores de decisão aumentadas e ensacadas. Desta forma, é proposto um vetor de característica 9-D que consiste nos mapas dos vasos sanguíneos adquiridos a partir de um conjunto de métodos. Assim o procedimento combina análise do gradiente orientado, transformação morfológica, medidas de força de linha e filtro *Gabor*. Na mesma perspectiva, o trabalho (WANG et al., 2015) segmenta os vasos sanguíneos combinando dois classificadores robustos. Um classificador baseado em CNN e outra baseado em *Randon Florest* (RF). Nesta abordagem os recursos treináveis são extraídos usando uma CNN e como classificadores treináveis são usados os RFs. A partir da integração do aprendizado de características e do classificador

tradicional, o método aprende de forma automática.

Em uma abordagem semelhante, no trabalho de (GEETHARAMANI; BALASUBRAMANIAN, 2016a) é proposto combinação de metodologias supervisionadas e não supervisionadas para extrair os vasos sanguíneos das imagens da retina. A tarefa de segmentação é dividida em três etapas consecutivas, o pré-processamento, a extração dos vetores de características e uma técnica de agrupamento. Na etapa de pré-processamento são utilizadas transformação e extração do modelo de cores da imagem, filtragem de *Gabor*, retificação de meia onda de forma sequencial. Na etapa de extração de características, a partir das imagens pré-processadas são extraídos os vetores de características usando análise de componentes principais. Na última etapa, os *pixels* são agrupados usando uma técnica de agrupamento *k-means* e outros algoritmos de agrupamento.

Outro trabalho que também segmenta os vasos sanguíneos de forma automática foi proposto por (BALASUBRAMANIAN; ANANTHAMOORTHY, 2019). As orientações médias baseadas em super-pixel são utilizadas para segmentar os vasos. Adicionalmente, uma *CNN* é aplicada para extrair um conjunto de vetores de características presentes nas regiões segmentadas. Por fim um classificador baseado em *SVM* realiza a classificação dos recursos extraídos para classificar em regiões com a presença dos vasos e sem a presença dos vasos.

3.6 Considerações finais deste capítulo

Este capítulo apresentou os principais trabalhos encontrados em nossa pesquisa, levando em consideração os diferentes tipos de classificação do processo de segmentação de imagens da retina. Diferentes técnicas como metodologia baseada em rastreamento, técnicas baseada em modelo, abordagens híbridas, abordagem de reconhecimento de padrões não supervisionado e supervisionados. Apesar de cada técnica apresentar vantagens e desvantagens, as técnicas baseadas em *ANN* e *CNN* são o estado da arte em muitas tarefas, como detecção, localização e segmentação. Assim, uma diferença que pode ser ressaltada entre as arquiteturas *ANN* e *CNN* é o fluxo das informações. No modelo clássico de *ANN* o fluxo da informações é propagado para frente e no sentido inverso. Contudo, no modelo *CNN* as informações são propagadas em um único sentido devido o uso das operações de convolução. Outra, vantagem desta abordagem é a capacidade de inserir novas informações e treinar novamente o classificador.

Parte II

Contribuições

CAPÍTULO 4

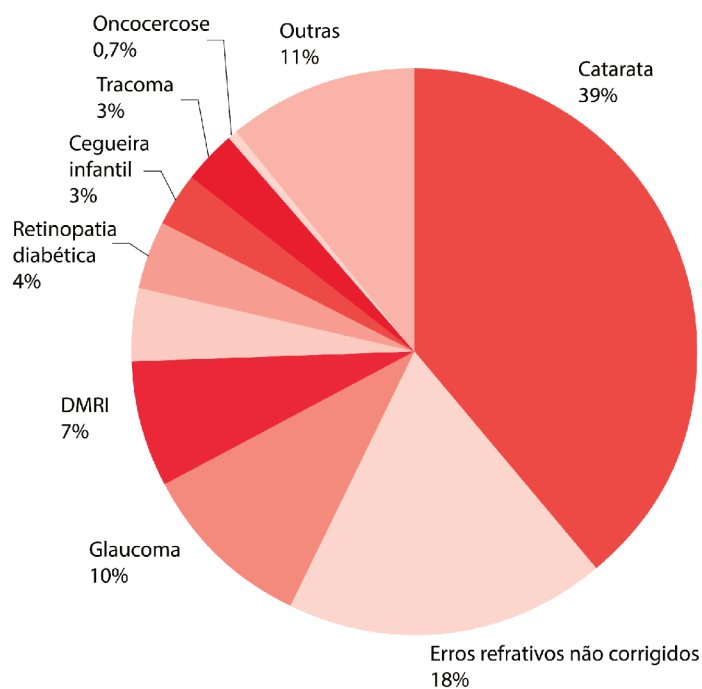
APRIMORAMENTO DA QUALIDADE DA IMAGEM DO FUNDO DE OLHO PARA DETECÇÃO DE VASOS SANGUÍNEOS POR MEIO DE UMA REDE NEURAL USANDO FILTRO CLAHE E WIENER

A retinografia digital é um método não invasivo para rastreamento e monitoramento de doenças visuais e tem sido amplamente utilizada devido ao aumento da incidência de diabetes. A International Diabetes Federation estima que 425 milhões de pessoas atualmente têm diabetes e que esse número pode aumentar para 629 milhões em 2045 (FOROUZANFAR, 2016). A glicose alta no sangue causa dano capilar retinal que leva a doenças diabéticas dos olhos, como catarata, glaucoma, retinopatia diabética (RD), edema macular diabético e outros, que por sua vez causam perda gradual da visão e eventual cegueira e nos casos mais extremos a morte (FOROUZANFAR, 2016).

Na representação gráfica da Figura 4.1 a catarata representa 39% dos casos de cegueira, o glaucoma corresponde a 10% e a retinopatia diabética representa 4%, se considerarmos apenas essas três enfermidades, elas representam um total de 53% dos casos de cegueira. Outro recorte importante, agora relacionada a idade, e que depois dos 50 anos, os países que apresentam uma renda média maior têm a maior proporção de mortes atribuídas à alta glicemia, tanto para homens quanto para mulheres Figura 4.2. Estes dados são extraídos da região Africana, Região da América Ocidental, região do Mediterrâneo Oriental, região Europeia, região do Sudeste Asiático Ocidental, Região do Pacífico Ocidental (WHO, 2016).

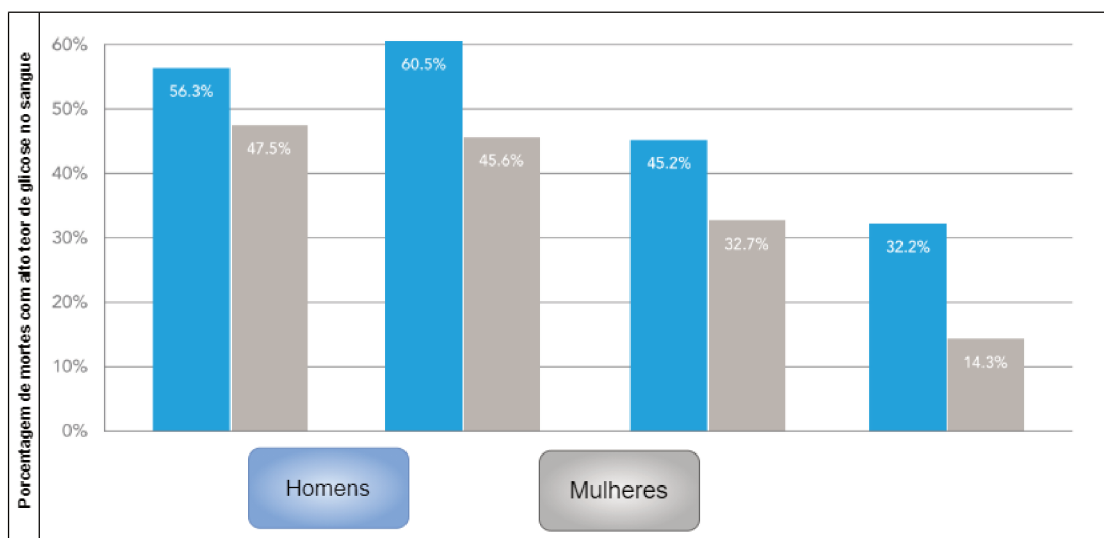
A detecção precoce de retinopatia diabética e edema macular é uma etapa crucial na

Figura 4.1: Principais causas da cegueira



Fonte:(WHO, 2016)

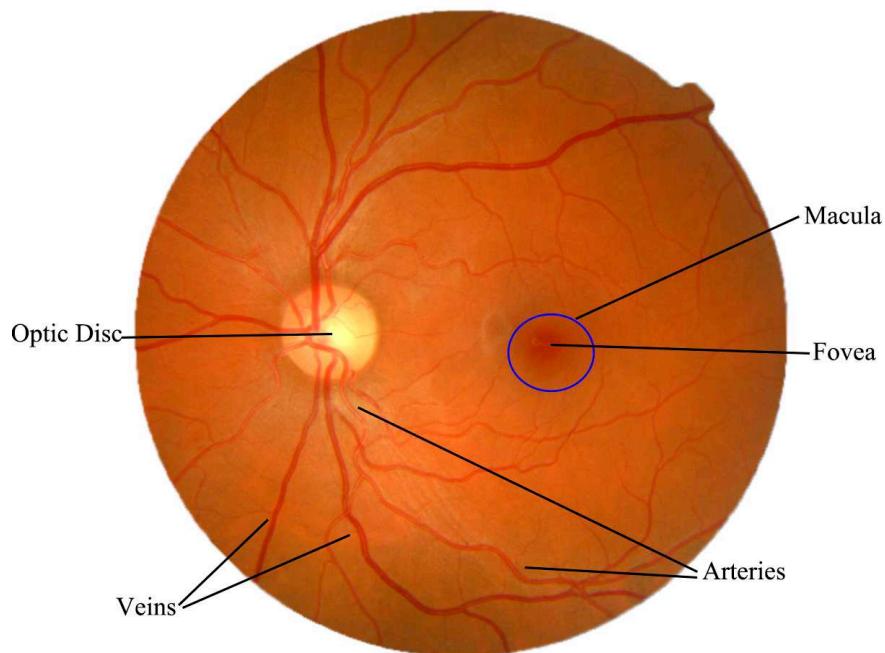
Figura 4.2: Porcentagem de todas as mortes atribuíveis a níveis elevados de glicose no sangue para adultos com idade entre 20 e 69 anos, por região e sexo, nos anos de 2000 e 2012



Fonte:(WHO, 2016)

redução da incidência de morbidade visual; no entanto, na maioria dos países, as imagens da retina (Figura 4.3) são analisadas por métodos manuais relativamente ineficientes e caros (KUMAR et al., 2012), (ZAKI et al., 2016). O diagnóstico automático depende da extração de sinais oculares e marcos anatômicos das imagens do fundo. Atenção especial também deve ser dada à detecção e remoção de vasos sanguíneos, que se estendem pela retina e complicam a análise dos elementos retiniais. Por exemplo, os graus de RD são definidos por sinais de alerta ocular, como número de microaneurismas, número de hemorragias, neovascularização e tortuosidade dos vasos sanguíneos. Uma hemorragia não pode ser confundida com um vaso sanguíneo. Portanto, a segmentação automática dos vasos sanguíneos retiniais é um assunto importante na análise de imagens médicas.

Figura 4.3: Imagem colorida de fundo com estruturas anatômicas



Fonte:(STAAL et al., 2004)

Segmentar vasos sanguíneos em imagens de fundo de olho significa separar as estruturas vasculares de outras estruturas anatômicas e elementos patológicos. A estrutura vascular do fundo do olho é composta por veias e artérias. Os principais vasos sanguíneos são a artéria retiniana central e a veia retiniana central e, às vezes, uma artéria ciliarretiniana. Esses vasos originam-se no disco óptico, bifurcam-se em ramos superior e inferior e se dividem em ramos nasal e temporal. As artérias retinianas têm paredes escuras e refletem a luz paralelamente ao eixo da artéria, enquanto as veias são mais escuras e largas do que as artérias (DORION, 2002).

Devido às características das paredes dos vasos sanguíneos, os métodos de detecção de bordas não são eficazes na extração de estruturas vasculares, o que requer métodos de segmentação mais sofisticados (CHAUDHURI et al., 1989). A segmentação de imagens é uma das tarefas mais complexa no processamento de imagens (GONZALEZ; WOODS, 2009). Um dos fatores é a baixa qualidade das imagens do fundo de olho, que apresentam baixo contraste, ruídos, defeitos de iluminação e todos estes problemas dificultam o processo de distinguir estruturas (KUMAR et al., 2012), (SUERO et al., 2013), (XIE et al., 2015). Consequentemente, a identificação automática das estruturas vasculares em imagens retinianas é um problema em aberto (VEGA et al., 2015).

Vários tipos de métodos têm sido usados para a segmentação dos vasos sanguíneos. Os mais importantes deles são baseados em rastreamento, baseados em modelo (filtragem combinada), classificação de *pixels* supervisionados, modelos deformáveis e morfologia matemática (MOOKIAH et al., 2013). Neste tese foi usado um método baseado na classificação supervisionada de *pixels*, em que a segmentação manual é realizada por um especialista. Essa imagem rotulada (*ground truth*) é usada para treinar um modelo que é então usado para segmentar outras imagens. A entrada de uma imagem adiciona um alto grau de complexidade à rede, o que exige o uso de técnicas de processamento de imagem convenientes para melhorar sua capacidade de aprendizado. Essas técnicas criam uma variabilidade ainda maior que pode aumentar ou diminuir os erros durante o treinamento. Pequenas modificações afetam muito os resultados. Consequentemente, ajustes baseados no julgamento humano normalmente não são confiáveis. As técnicas de aprimoramento de imagem também podem ser problemáticas. Quando uma imagem é processada por uma máquina, os resultados são quantificados mais facilmente (GONZALEZ; WOODS, 2009). Portanto, essas técnicas devem ser examinadas e ajustadas a casos de uso específicos, como segmentação de vasos sanguíneos por meio de classificação de *pixels* supervisionada.

Diversas técnicas podem ser empregadas para aumentar o contraste da imagem, como a transformação de extensão decorrelação, *Histogram Equalization* (HE), *Adaptive Histogram Equalization* (AHE) e CLAHE. CLAHE é um dos métodos mais usados para segmentação de vasos sanguíneos (SINGH; SRIVASTAVA, 2016), (GEETHARAMANI; BALASUBRAMANIAN, 2016b), (AZZOPARDI et al., 2015), (AGURTO et al., 2015), (VOSTATEK et al., 2017). O ruído é outro problema nas imagens de fundo do olho (SONALI et al., 2019) e é causado pela vibração térmica dos átomos e pela radiação de

objetos quentes durante a aquisição, armazenamento e processamento da imagem. Apesar desses problemas, os autores raramente relatam como os parâmetros **CLAHE** e do filtro de redução de ruído são escolhidos para a segmentação dos vasos sanguíneos.

Neste trabalho é proposto uma estratégia para melhorar a qualidade da imagem para a detecção de vasos sanguíneos usando **CLAHE**, um filtro de *Wiener* e uma **MLP**. Esse processo depende de um ajuste cuidadoso do pré-processamento da imagem para que os vasos sanguíneos possam ser classificados pela **ANN** quanto à precisão, especificidade e sensibilidade. O estudo atual difere de outros porque usa medidas de qualidade da imagem depois de comparar o comportamento do método de classificação de pixel supervisionado com o processo de pré-processamento, e não apenas *Peak Signal to Noise Ratio* (**PSNR**) e outras medidas de qualidade. O restante do capítulo está organizado da seguinte forma: A seção “Trabalhos relacionados” descreve o corpo do trabalho que fornece a base para o método proposto. A seção “Métodos” detalha a metodologia, descreve o banco de dados usado e fornece uma base para a compreensão do método. Finalmente, essas seções são seguidas pelas seções “Resultados e Discussão” e por fim a seção com as “Considerações Finais”.

4.1 Trabalhos Relacionados

Esta seção apresenta uma breve visão geral de alguns dos estudos que estão intimamente ligados ao presente. Alguns desses documentos negligenciam indicar se as técnicas são adequadas para análise humana ou de máquina e, portanto, apresentam métricas quantitativas para confirmar a eficiência dos métodos de aprimoramento de imagem, por exemplo, (**KUMAR et al., 2018**), (**SONALI et al., 2019**), (**LIDONG et al., 2015**). As métricas quantitativas mais comuns para avaliar os aprimoramentos da imagem são **PSNR**, signal-to-noise ratio (**SNR**), noise estimate (**NE**) e o Structural Similarity Index (**SSIM**). Outros estudos usaram aprimoramento de imagem para segmentar sinais oculares ou estruturas anatômicas e, em seguida, avaliaram a metodologia usando medidas de eficiência de segmentação, como precisão, especificidade e sensibilidade (**ELBALAOUI et al., 2016**), (**SRENG et al., 2017**). Ainda outros estudos segmentaram sinais oculares e estruturas anatômicas sem detalhar o processo de realce de imagem (**FRAZ et al., 2012a**), (**HASSANIEN et al., 2015**), (**LISKOWSKI; KRAWIEC, 2016**), (**MAJI et al., 2016**), (**ORLANDO; BLASCHKO, 2014**), (**SOOMRO et al., 2017**), (**VEGA et al., 2015**),

(VOSTATEK et al., 2017) (MORANO et al., 2021), (DHARMAWAN et al., 2019). Alguns desses estudos são discutidos em mais detalhes abaixo.

(KUMAR et al., 2018) usou técnicas de filtragem em conjunto com CLAHE. Nesse processo, o filtro adaptativo de ruído fuzzy switching median (NAFSM), a correspondência de blocos e a filtragem 3D (BM3D) foram combinados para diminuir o ruído de uma imagem. Em seguida, foram aplicadas a equalização do histograma, a equalização adaptativa do histograma, CLAHE e a equalização do histograma da subimagem baseada na exposição (ESIHE). Finalmente, os resultados foram avaliados por SNR, PSNR, entropia e medidas de erro de brilho médio absoluto. Os melhores resultados foram alcançados com uma combinação de filtragem BM3D e ESIHE. (SONALI et al., 2019) usou uma abordagem semelhante em que a imagem do fundo foi aprimorada usando filtros de denoising e realce de contraste. Nesse processo, o ruído foi reduzido nos canais R, G e B com filtros de média, mediana, wiener, gaussiano e mediana ponderada. Em seguida, o CLAHE foi aplicado a cada canal. Esses processos melhoraram o PSNR em 7.84%, o SSIM em 1.19% e o Índice de Preservação de Borda (EPI) em 1.28%, em relação aos métodos mais modernos. (LIDONG et al., 2015) propôs um algoritmo híbrido para aprimorar 300 imagens de teste (de diferentes sujeitos humanos). Este método usou a transformada wavelet discreta (DWT) para decompor a imagem original em componentes de baixa e alta frequência e, em seguida, melhorou os componentes de baixa frequência usando CLAHE. A eficácia desses métodos foi avaliada usando parâmetros como PSNR, NE e o incremento de entropia local (LEI).

(ELBALAOUI et al., 2016) propuseram um método para detecção de vasos sanguíneos usando extração de canal verde, extração de máscara e dilatação de borda, com CLAHE para pré-processamento e um filtro de vasos sanguíneos para detecção de vasos sanguíneos. Além disso, uma matriz multiescala Hessian foi projetada a partir do limite adaptável para aprimorar o filtro do vaso. Os resultados foram avaliados por medidas de sensibilidade, especificidade e precisão. (SRENG et al., 2017) realizou o aprimoramento de imagens da retina usando a detecção automática de microaneurismas. Durante o pré-processamento, a análise de componente principal (PCA) foi usada para reduzir o tamanho da imagem, enquanto a transformada wavelet 2D discreta inversa e o filtro de mediana foram usados para reduzir o ruído. Em seguida, a detecção de bordas do Canny e o limiar de entropia foram usados para segmentar a imagem. Por fim, o método foi avaliado no banco de dados DIARETDB1 e apresentou 90% de sucesso na detecção de

microaneurismas.

([FRAZ et al., 2012a](#)) combinaram várias técnicas para automatizar o processo de segmentação dos vasos sanguíneos. Inicialmente, o primeiro derivado de um filtro gaussiano foi usado para determinar as linhas centrais dos vasos sanguíneos. Em seguida, a forma e a orientação do vaso foram extraídas usando um operador morfológico de cartola com corte plano de bits. Finalmente, as linhas de centro e o mapa de orientação foram combinados para identificar os vasos sanguíneos.

([HASSANIEN et al., 2015](#)) propôs um procedimento automatizado de segmentação de vasos sanguíneos, baseado em estratégias de otimização que combinavam o algoritmo de enxame de colônias de abelhas e a função de capacidade fuzzy c-means para detectar vasos mais grossos durante um primeiro nível de otimização e vasos mais finos durante um segundo nível. Da mesma forma, ([LISKOWSKI; KRAWIEC, 2016](#)) usaram uma arquitetura de rede neural profunda para segmentar os vasos sanguíneos. Nesse processo, um extenso conjunto de 400.000 amostras foi obtido pela adoção de estratégias de pré-processamento, normalizações globais de contraste, transformações geométricas e correções gama no banco de dados DRIVE. ([MAJI et al., 2016](#)) apresentou uma rede neural convolucional profunda que segmentou vasos sanguíneos em imagens retiniais. A segmentação dos vasos sanguíneos foi alcançada treinando um conjunto de Rede Neural Convolucional 12 usando imagens brutas. O principal objetivo deste procedimento foi distinguir pixels pertencentes a vasos daqueles que não pertencem. Outra abordagem de segmentação de vasos sanguíneos foi proposta por ([ORLANDO; BLASCHKO, 2014](#)), que treinaram um modelo de campo randomizado condicional totalmente conectado para segmentar a estrutura vascular da retina.

([SOOMRO et al., 2017](#)) usaram um método de aprendizado profundo com pré e pós-processamento para mitigar os problemas de iluminação desigual em imagens do fundo de olho. ([VEGA et al., 2015](#)) propôs a segmentação dos vasos sanguíneos usando uma Rede Neural Treliça com Processamento Dendrítico (LNNDP) sem supervisão dos parâmetros da rede. Na abordagem proposta por ([MORANO et al., 2021](#)) as tarefas de segmentação e classificação das artérias e veias da imagens de fundo de olho são implementadas de forma simultâneas usando uma redes neurais completamente convolucionais FCNNs. O problema original de segmentação é decomposto em outros três problemas a saber: a segmentação das artérias, a segmentação das veias e a segmentação dos vasos sanguíneos. Uma métrica

denominada entropia cruzada binária por três (BCE3) combina as perdas independentes. Em outra abordagem, o trabalho de (DHARMAWAN et al., 2019) apresenta um algoritmo híbrido para segmentar os vasos sanguíneos. Primeiro, um algoritmo destaca os vasos sanguíneos utilizando um método de realce antes de enviar as imagens para uma rede neural convolucional com arquitetura derivada da U-Net.

4.2 Métodos

A diferenciação entre os vasos sanguíneos e fundo da imagem pode ser uma tarefa difícil de ser concebida. Assim, esta seção apresenta os materiais e a abordagem para melhorar as imagens de fundo do olho.

4.2.1 Banco de dados

A segmentação precisa da estrutura vascular das imagens de fundo de olho colorida podem fornecer informações relevantes para diagnosticar um grande número de enfermidades como: retinopatia diabética, degeneração macular, glaucoma e hipertensão (FRAZ et al., 2012c). Um grande número de algoritmos ou técnicas voltadas para a tarefa de segmentação são elaboradas utilizando dois bancos de dados extremamente citados na literatura científica DRIVE e STARE (HOOVER, 2000), (STAAL et al., 2004). Ambas as bases são aceitas pela comunidade científica, em especial para as tarefas de segmentação dos vasos da retina e também para o processo de classificação.

As imagens utilizadas no processo de validação com ANN foram obtidas no banco de dados DRIVE, que está disponível publicamente para estudos de extração de vasos sanguíneos. Este banco de dados contém 40 imagens coloridas do fundo do olho (584×565) que são divididas em um conjunto de teste (20 imagens) e um conjunto de treinamento (20 imagens) (NIEMEIJER et al., 2004). Cada imagem é acompanhada por uma máscara para delimitar o campo de visão (FOV) e uma segmentação manual (*ground truth*) para o conjunto de treinamento. Duas segmentações manuais estão disponíveis para compor o conjunto de teste, onde a segmentação do primeiro especialista é usada como *ground truth*. As informações dos bancos dados STARE e DRIVE podem ser resumidas na Tabela 4.1, em que destacamos que os bancos de dados são compostos tanto por imagem de fundo de olho saudáveis, mas também por imagem com alguma patologia relacionada a visão.

Tabela 4.1: Resumo das informações DRIVE e STARE

Dataset	Imagens	Dimensão	Classificação	câmera
STARE 2000	20	605x700	10 normais 10 patológicas	TopconTRV-50 35°a FOV
DRIVE 2004	40	768x584	33 normais 7 patológicas	Canon CR5 non-mydratic 3-CCD 35° FOV

4.2.2 Filtro de Wiener

O filtro Wiener é um filtro passa-baixa que pode ser aplicado em diferentes contextos para restaurar sinais degradados por ruído. O filtro é baseado em uma abordagem estatística que assume que o sinal e ruído são processos estocásticos lineares estacionários com características espectrais conhecidas. O desempenho desse processo é superior aos outros porque atinge um equilíbrio ideal na compensação de viés-variância (BHATTACHARJEE; DAS, 2013), (GHAEL et al., 1997). Em outras palavras, o filtro *Wiener* é um filtro adaptativo que calcula a média e a variância de um vizinhança e então aplica suavização mínima quando a variação é alta e suavização mais forte quando a variação é baixa.

O filtro minimiza o erro entre o sinal estimado e o sinal original. Por exemplo, dada uma imagem f não corrompida e uma image estimada $\hat{f}$, a medida de erro é calculada na Equação 4.1 da seguinte forma:

$$e^2 = E(f - \hat{f})^2 \quad (4.1)$$

onde $E(\cdot)$ é o valor esperado do argumento. Portanto, a obtenção de uma imagem estimada é reduzida ao desafio de encontrar o mínimo da função de erro quadrático. Para isso, utiliza-se o domínio da frequência e as seguintes suposições são feitas: o ruído e a imagem não estão correlacionados, o ruído ou a imagem tem média zero e os níveis de intensidade na imagem estimada são degradados por uma função linear. Com base nessas condições, o mínimo da função de erro é dada na Equação 4.2 quando (GONZALEZ; WOODS, 2009):

$$\hat{F}(u, v) = \left[\frac{H^*(u, v) S_f(u, v)}{S_f(u, v) |H(u, v)|^2 + S_f(u, v)} \right] G(u, v) \quad (4.2)$$

onde $\hat{F}(u, v)$ é a imagem estimada no domínio da frequência, $H(u, v)$ é a transformação da função de degradação, $G(u, v)$ é a transformação da imagem degradada, $H^*(u, v)$ é o conjugado complexo de $H(u, v)$, e $S_f(u, v) = |F(u, v)|^2$ é o espectro de potência da imagem não degradada. O princípio geral do filtro é que o produto de um valor complexo e seu

conjugado é a magnitude do valor complexo ao quadrado. Portanto, a Equação 4.3:

$$\hat{F}(u, v) = \left[\frac{1}{H(u, v)} \frac{|H(u, v)|^2}{|H(u, v)|^2 + S_n(u, v)/S_f(u, v)} \right] G(u, v) \quad (4.3)$$

onde $S_\eta(u, v) = |N(u, v)|^2$ é o espectro do ruído. O termo $S_\eta(u, v)/S_f(u, v)$ é substituído por um K constante, porque o espectro de potência da imagem não degradada geralmente não é conhecido. Assim podemos resumir a Equação 4.4 como:

$$\hat{F}(u, v) = \left[\frac{1}{H(u, v)} \frac{|H(u, v)|^2}{|H(u, v)|^2 + k} \right] G(u, v) \quad (4.4)$$

No processamento de imagem digital, o filtro de *Wiener* pode ser usado em imagens que foram degradadas por um ruído aditivo de energia constante (BHATTACHARJEE; DAS, 2013). Assim, os parâmetros do filtro de *Wiener* são o tamanho da vizinhança e a potência do ruído.

4.2.3 CLAHE

O histograma representa uma distribuição discreta de probabilidade de intensidade e pode ser definido como uma função $h(r_k) = n_k$ onde r_k é a k -th nível de intensidade k -ésimo, n_k é o número de pixel com intensidade r_k , e $k = 0, 1, \dots, L$. Os picos no histograma representam os níveis de intensidade mais comuns. Portanto, imagens de baixo contraste têm altas concentrações em torno de alguns níveis de intensidade. A equalização do histograma resolve esse problema e melhora o contraste da imagem ao expandir a distribuição dos níveis de intensidade L entre os níveis de intensidade (LIDONG et al., 2015). No entanto, a expansão simultânea de todos os níveis de intensidade pode adicionar intensidades nas extremidades superior e inferior do histograma. Isso, por sua vez, pode fazer a imagem parecer super-exposta ou destacar aspectos indesejados, pois a equalização pode modificar as características básicas do histograma (GONZALEZ; WOODS, 2009). Essas situações podem ser evitadas usando equalização de histograma adaptativa.

Na equalização adaptativa do histograma, cada histograma de uma subimagem redistribui os valores de brilho da imagem, o que melhora o contraste local e traz mais detalhes. No entanto, esse processo pode aumentar o ruído quando usado em imagens com ruído, como imagens médicas. Na prática, o algoritmo CLAHE distribui o brilho da imagem e limita a amplificação do ruído. Para conseguir isso, CLAHE divide uma imagem em pequenos blocos sobrepostos. Em seguida, o aumento de contraste é executado

em cada bloco por equalização do histograma. Um limite de corte é usado para superar problemas de amplificação de ruído cortando a altura do histograma antes de calcular a função de distribuição cumulativa (MA et al., 2017). A forma do histograma de cada bloco é reorganizada de acordo com uma distribuição (uniforme, Rayleigh ou exponencial). O método CLAHE compreende várias etapas (MA et al., 2017).

1. Divide a imagem original em $M \times N$ tiles.
2. Calcula o histograma para cada região baseada no nível de cinza da imagem.
3. Calcula um histograma de contraste limitado para cada região:
 - 3.1 Use o número de níveis de cinza N_{gray} na região e o número de pixels dentro das dimensões X e Y para calcular o número médio de pixels.

$$N_{avg} = (N_X \times N_Y) / N_{gray} \quad (4.5)$$

- 3.2 Dado que N_{CL} é o limite real do clip limit CL , e N_{clip} é o CL normalizado em um intervalo de $[0, 1]$. Recorte os pixels se o número de pixels for maior que N_{CL} .

- 3.3 Calcule o número de pixels recortados $N_{\sum(clip)}$ e a média de pixels restantes.

$$N_{avggray} = N_{\sum(clip)} / N_{gray} \quad (4.6)$$

- 3.4 Dado que $H_{region}(i)$ é o histograma original e $H_{region-clip}(i)$ é o histograma recortado de cada região no i_{th} nível de cinza, a regra de recorte do histograma é dada pelo seguinte:

If $(H_{region}(i) > N_{CL})$ ou $(H_{region}(i) > N_{CL})$ então $H_{region-clip}(i) = N_{CL}$

Else $H_{region-clip}(i) = H_{region}(i) - N_{CL}$

4. Dado que N_{remain} é o número restante de pixels cortados, Step é um número inteiro positivo maior ou igual a 1, redistribua os pixels restantes usando a etapa:

$$Step = N_{gray} / N_{remain} \quad (4.7)$$

- 4.1 Progredir desde o mínimo até o nível de cinza máximo através da etapa anterior. Se o número de pixels no nível de cinza for menor que NCL , distribua um pixel para o nível de cinza.
- 4.2 Se todos os pixels não foram distribuídos na pesquisa, calcule uma nova etapa e inicie uma nova rodada de pesquisa. Pare quando todos os pixels restantes forem distribuídos.
5. Aumente os valores de intensidade em cada região usando uma transformação (por exemplo, rayleigh).
 - 5.1 Crie uma função de transferência transformando o histograma recortado em uma probabilidade cumulativa $P_{input}(i)$. Dado que y_{min} é o limite inferior do valor do pixel, e α é o parâmetro de escala de Rayleigh, a transformada de Rayleigh é dada por:

$$y(i) = y_{min} + \sqrt{2\alpha^2 \ln \left(\frac{1}{1 - P_{input}(i)} \right)} \quad (4.8)$$

$$p(y(i)) = \frac{(y(i) - y_{min})}{\alpha^2} \cdot \exp \left(-\frac{(y(i) - y_{min})}{2\alpha^2} \right) \quad (4.9)$$

para

$$y(i) \geq y_{min} \quad (4.10)$$

6. Reduza os efeitos que mudam abruptamente usando a extensão de contraste linear.

$$y(i) = \frac{x(i) - x_{min}}{x_{max} - x_{min}} \quad (4.11)$$

7. Elimine artefatos de limite calculando a atribuição de nível de cinza de pixels por meio de interpolação bilinear.

4.2.4 Experimentos realizados

De forma geral muitos pesquisadores realizam as suas configurações e ajuste de hiperparâmetros baseados em experimentos empíricos. Neste trabalho a preparação foi dividida em duas fases distintas. Na primeira fase, a revisão dos trabalhos relacionados a segmentação dos vasos sanguíneos da retina, atribuiu pistas dos algoritmos e as bases de dados a serem trabalhadas seção 4.1. Na segunda fase, a configuração da arquitetura ANN

(definição do número de neurônios em cada camada) e na sequência os hiperparâmetros dos algoritmos de pré-processamento e pós-processamento são definidos nas seções 4.2.4 e 4.3.

A Tabela 4.2 é o resultado refinado de 1051 experimentos. Nestes o objetivo foi determinar o número ideal de neurônios para compor a ANN. Assim, cada experimento tem um custo médio de cinco minutos. Portanto, o tempo médio de teste e ajuste nos hiperparâmetros foram de 87h08m. Nos experimentos realizadas diferentes escolhas do número de neurônios que configuram a ANN foram testadas e com diferentes tamanhos de blocos. Assim cada bloco se transforma em um vetor de tamanho 4, 16, 64 e 144 respectivamente que representa o número de neurônios na camada de entrada e saída, visto que a escolha para ANN é simétrica com relação a entrada e saída do dados.

Contudo, no processo de análise da ANN apenas os resultados com *accuracy* superior à taxa de 92% são considerados. Na sequência esses resultados são apresentados na Tabela 4.2. As ANNs são submetidas ao treinamento com dez, cem e duzentas épocas e os tempos de processamento são anotados. O aumento estrutural na ANN influencia diretamente no tempo de processamento e é um detalhe importante. Em contra-partida, o quanto mais robusta é a ANN melhor é a *accuracy* média. Assim, a arquitetura considerada para a próxima fase é a $MLP_{(64-128-64)}$ com a maior desempenho de *accuracy* média. Em seguida, os hiperparâmetros da Tabela 4.3 são combinados com a rede MLP em 3240 experimentos com custo computacional aproximado de 135h. Os detalhes são apresentados na seção 4.3.

Tabela 4.2: ANNs selecionadas na fase experimental. O termo $MLP_{(16-32-16)}$ significa 16 neurônios na camada de entrada, 32 neurônios na camada oculta, e 16 neurônios na camada de saída.

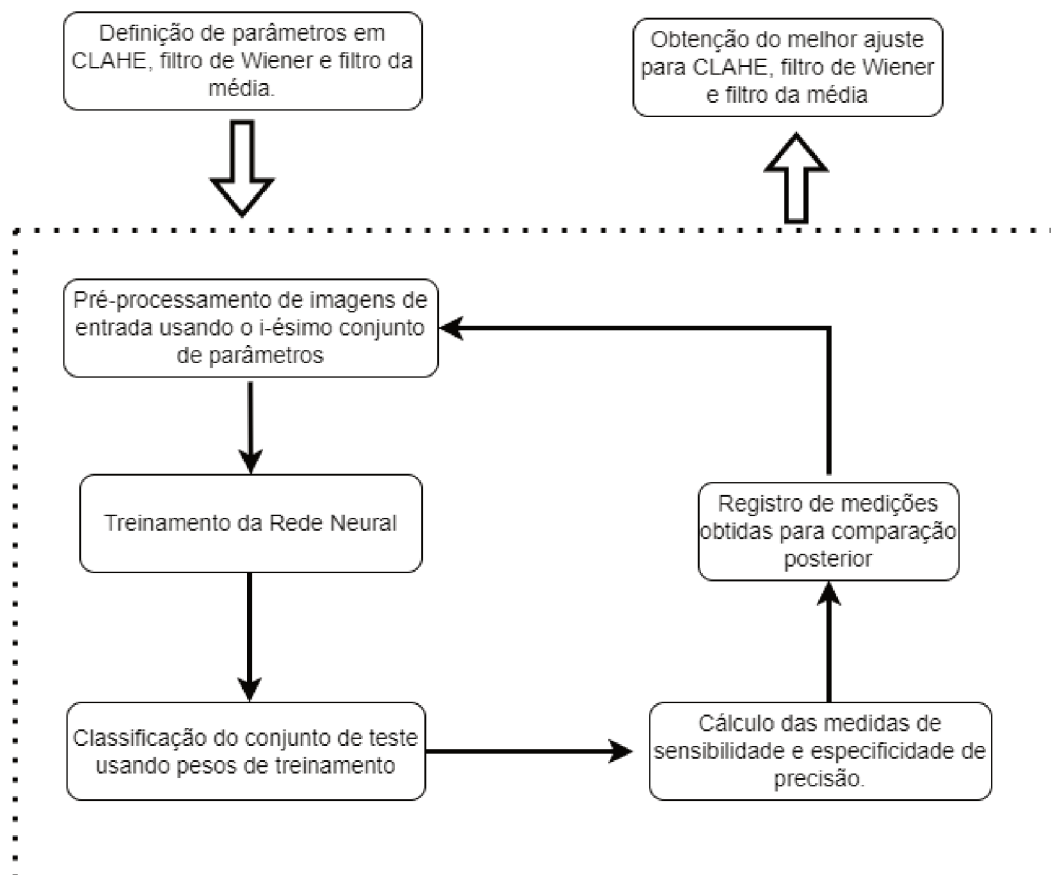
Arquiteturas	10 épocas	100 épocas	200 épocas	Tempo Total	<i>Accuracy</i> média
$MLP_{(16-32-16)}$	280s	3168s	6278s	9726s	0.9108
$MLP_{(16-64-16)}$	361s	3699s	7758	11818s	0.9407
$MLP_{(16-80-16)}$	434s	3941s	8301	12676s	0.9388
$MLP_{(16-100-16)}$	495s	4335s	8598	13428s	0.9490
$MLP_{(64-32-64)}$	115s	1128s	2339s	3582s	0.9502
$MLP_{(64-64-64)}$	167s	1663s	3598s	5428s	0.9515
$MLP_{(64-100-64)}$	200s	1939s	4683s	6822s	0.9536
$MLP_{(64-128-64)}$	253s	3364s	5634s	9251s	0.9540

Fonte: Elaborado pelo autor

4.3 Aprimoramento de imagem para detecção dos vasos sanguíneos usando uma rede neural baseada em CLAHE e o filtro Wiener

Na Figura 4.4, primeiro definimos um conjunto de valores aceitáveis (Tabela 4.3) para os vários parâmetros em CLAHE, o filtro Wiener e o filtro médio. Em seguida, realiza-se duzentas execuções para avaliar o desempenho do ajuste de hiperparâmetro. Para cada conjunto de hiperparâmetros, a rede foi treinada com o conjunto de treinamento do banco de dados DRIVE (rede de configuração fixa definida na Tabela 4.2). Em seguida, as imagens do conjunto de teste foram classificadas usando pesos de treinamento e precisão, sensibilidade e especificidade foram calculadas e registradas. Por fim, as métricas armazenadas foram utilizadas para comparar quantitativamente os resultados e melhorar a imagem de fundo do olho para detecção de vasos sanguíneos. O melhor conjunto de hiperparâmetros foi escolhido após análise estatística completa.

Figura 4.4: Metodologia proposta para ajustar os parâmetros e aprimorar as imagens de fundo do olho



Fonte: Elaborado pelo autor

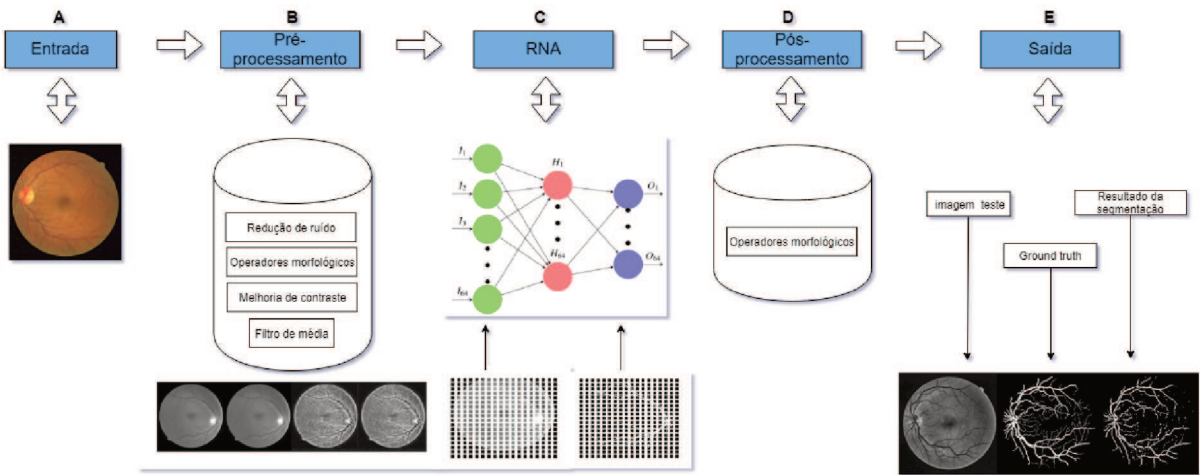
Inspeções visuais foram realizadas para definir os parâmetros iniciais do teste (Tabela 4.3). Os ajustes baseados no julgamento humano carecem de precisão e mesmo pequenos desvios podem afetar a qualidade do treinamento da rede. Portanto, os valores foram escolhidos para cobrir o máximo possível dos domínios da técnica. Todos os parâmetros foram combinados, resultando em 4320 combinações. Este número foi reduzido para 3240 combinações, uma vez que as distribuições uniformes não possuem parâmetros alfa.

Tabela 4.3: Parâmetros e valores para os filtros Wiener, CLAHE e Média

Parameter	Values
Wiener power noise	0.01, 0.015, 0.02
Wiener block size	2, 3
CLAHE distribution	uniform, Rayleigh, exponential
CLAHE Tiles	8, 14, 18, 25, 31
CLAHE Clip limit	0.02, 0.03, 0.04
CLAHE Alpha (Rayleigh and exponential distributions)	0.2, 0.3, 0.4, 0.5
CLAHE bins histogram	200, 256
Mean filter block size	3, 4

Foi utilizada uma ANN com múltiplas camadas para realizar a segmentação dos vasos sanguíneos da retina e garantir que o metodologia de realce da imagem de fundo do olho fornecesse a melhor solução para um problema específico (GONZALEZ; WOODS, 2009). A (Figura 4.5) ilustra a metodologia de segmentação.

Figura 4.5: Metodologia proposta para segmentação dos vasos sanguíneos



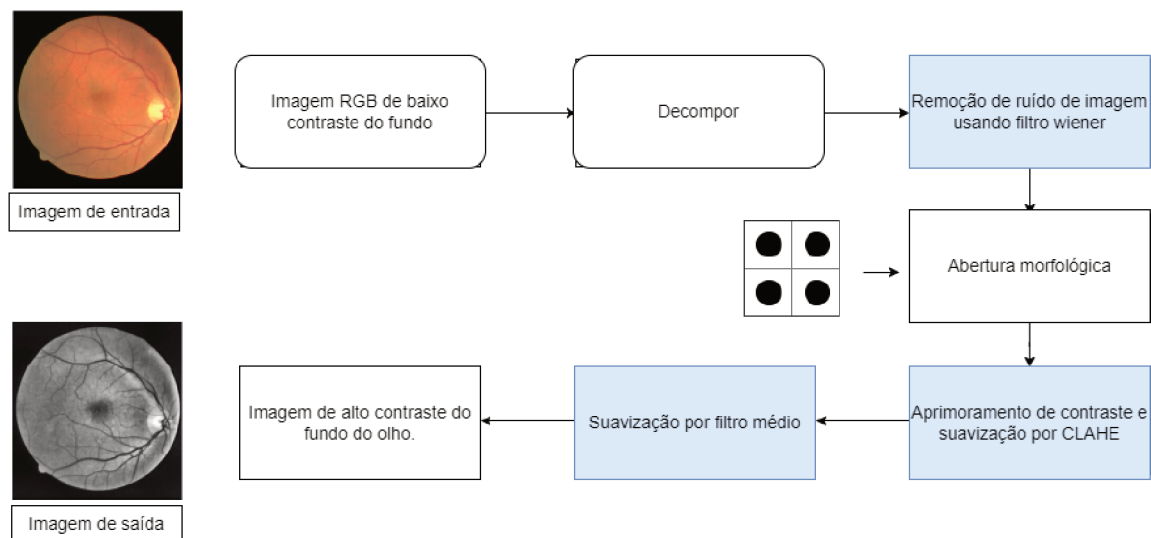
Fonte: Elaborado pelo autor

4.3.1 Pré-processamento

A etapa de pré-processamento (Fig. 4.5) foi dividida em cinco sub-etapas: decomposição, redução de ruído, operações morfológicas, aumento de contraste e suavização por meio de filtragem média.

(SONALI et al., 2019) mostraram que o canal verde da imagem RGB fornece a imagem com mais detalhes dos vasos sanguíneos e ruídos reduzidos. Esta estratégia é utilizada por muitos pesquisadores como (TAMIM et al., 2020) (WANG et al., 2021). Assim, a imagem colorida foi separada em canais RGB; no entanto, apenas o canal verde foi considerado. Em seguida, o ruído do processo de aquisição da imagem foi minimizado por meio da filtragem de *Wiener* (os valores dos parâmetros foram alterados a cada rodada de treinamento da rede). Em seguida a morfologia foi aplicada para eliminar pequenos artefatos de aquisição de imagem ou pigmentação no fundo do olho. Um elemento estruturante (2×2 quadrados) percorreu todas as imagens executando uma operação de abertura morfológica que consistia em erosão em segmentos finos de luz, desaparecimento de pequenos pontos de luz isolados e dilatação que devolveu os segmentos de luz aos seus estado de pré-erosão. Em seguida, o contraste dos vasos sanguíneos foi melhorado pelo **CLAHE** (os parâmetros foram alterados em cada rodada de treinamento da rede) (Fig. 4.6). Por último, o filtro médio foi usado para suavizar os efeitos do pré-processamento anterior (por exemplo, limites artificiais em junções de contornos). Uma janela de tamanho 3×3 é usada para homogeneizar esses efeitos e detalhes de textura.

Figura 4.6: Metodologia proposta para aprimoramento da imagem



Fonte: Elaborado pelo autor

4.3.2 Treinamento MLP

O treinamento é representado no diagrama do estágio C da Fig. 4.4. Os hiperparâmetros da rede (4.2) foram definidos como: 128 neurônios na camada oculta, um bloco 8×8 (64 neurônios na camada de entrada / saída), $\alpha = 0,000282$ e $\eta = 0,009125$ valores definidos na seção 4.2.4. Duzentas épocas foram usadas na fase de treinamento, uma vez que o termo *momentum* garantiu uma convergência acelerada.

Para tornar o treinamento mais robusto, foi realizado o aumento dos dados em relação as imagens. Assim, as imagens foram giradas 180 graus e adicionadas ao conjunto de treinamento para um total de 40 imagens (20 imagens originais e 20 giradas). Cada imagem foi então dividida em blocos de 8×8 , resultando em 5183 blocos e um conjunto de treinamento de $5183 \times 40 = 207.320$ amostras. Os blocos de imagem foram vetorizados, de modo que cada pixel em um bloco vetorizado representasse uma entrada de neurônio. A imagem do *ground truth* e a imagem de entrada também foram separadas em blocos.

A rede utilizou o algoritmo de inicialização Nguyen-Widrow (FAUSETT, 1994) e o algoritmo de retropropagação. A estratégia do método de inicialização *Nguyen-Widrow* gera pesos iniciais e valores de *bias* para uma camada, de forma que as regiões ativas dos neurônios da camada sejam distribuídas de maneira aproximadamente uniforme no espaço de entrada. Desta forma, é possível assumir que existe uma aceleração do processo de treinamento, definindo os pesos iniciais da primeira camada, de modo que cada neurônio receba seu próprio intervalo no início do treinamento. Contudo, a maioria dos ajustes a serem realizados nos pesos da rede será pequena em razão das escolhas iniciais do algoritmo (NGUYEN; WIDROW, 1990).

O método proposto ocorre em duas etapas: para frente e para trás. Na etapa de avanço, cada bloco é apresentado à camada de entrada da rede e o resultado flui pela rede, camada por camada, até que uma resposta seja produzida pela camada de saída. Na etapa anterior, a saída é comparada com a saída desejada para o bloco de entrada, um erro de classificação é calculado e propagado da camada de saída para a camada de entrada. Em seguida, os pesos sinápticos foram alterados entre duas interações com relação ao erro de propagação reversa. Os erros na classificação de *pixels* foram minimizados usando um gradiente descendente estocástico e um termo de *momentum*. A matriz atualizada com o termo de *momentum* é expressa pela Equação 4.12:

$$W_{ji}^L(t+1) = W_{ji}^L(t) + \alpha(W_{ij}^L(t) - W_{ij}^L(t-1)) + \eta\delta_j^L Y_j^{L-1} \quad (4.12)$$

onde η é a taxa de aprendizagem do algoritmo de retropropagação e α é a taxa de *momentum*. $(W)_{ij}^L(t)$ são os pesos da L -ésima camada no tempo t , e $W_{ij}^L(t-1)$ são os pesos da L -ésima camada no tempo $t-1$. Este último termo reflete o quanto as matrizes sinápticas mudaram de uma interação para outra. A matriz de peso W_{ji} deve ser ajustada na direção oposta à do gradiente para minimizar o erro. Portanto, a Eq. 4.12 atualiza os pesos dos neurônios nas camadas envolvidas de acordo com as diferenças entre as respostas de saída e os valores desejados. Além disso, o número de neurônios na camada de entrada parece estar relacionado a α e η , uma vez que esses parâmetros tiveram que ser ajustados sempre que a contagem de neurônios mudou, detalhes na secção 4.2.4.

4.3.3 Pós-processamento

A operação morfológica aberta foi usada no pós-processamento para minimizar pontos isolados. Um elemento de estruturação cruzada 3×3 foi utilizado para eliminar pontos isolados na erosão e retornar os vasos detectados na dilatação ao estado original. Uma operação de pesquisa também foi usada para conectar os vasos sanguíneos com interrupções curtas. Este processo envolveu a busca de uma matriz de conexões possíveis dentro de uma vizinhança, a fim de substituir um padrão que é evidenciado por links entre os *pixels* em questão. Essa matriz exigiu a construção de pequenas pontes entre os *pixels* da imagem binária que a rede classificou como pertencentes a um navio ou não.

4.3.4 Saída

A rede neural mostrou a probabilidade de um determinado neurônio pertencer a um vaso sanguíneo. A função de transferência sigmóide bipolar garantiu que cada pixel de saída foi definido como '0' para um neurônio não-vaso e '1', para um neurônio do vaso sanguíneo. A rede classificou os *pixels* de cada época e calculou os erros relativos à verdade fundamental. Esses erros foram propagados para a camada de entrada. Após 200 épocas, a rede salvou os pesos selecionados durante o treinamento e começou a classificar as imagens no conjunto de teste. Durante a classificação, os *pixels* rotulados da rede de saída foram remontados em blocos que formaram uma imagem reconstituída. A imagem resultante retratou os vasos sanguíneos em branco e o restante da imagem em preto. Para facilitar

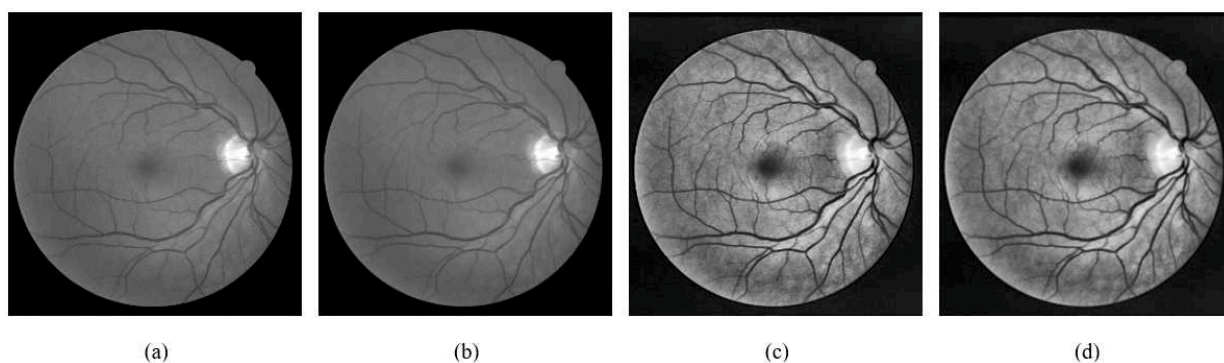
a inspeção, a imagem reconstruída foi plotada ao lado da imagem original e da verdade básica.

4.4 Resultados e Discussões

Esta seção apresenta os resultados obtidos a partir dos métodos descritos na seção anterior. As imagens são processadas com MatLab em um computador com Windows 10 de 64 bits (CPU Intel(R) Core™ i7-8750H @2,20 GHz, 6 core(s), 16 GB de RAM). Os experimentos de segmentação dos vasos sanguíneos são aprestados a seguir.

Na Figura 4.7, uma imagem do banco de dados DRIVE foi aprimorada usando os seguintes parâmetros: ruído de energia *Wiener* 0.015, tamanho do bloco *Wiener* 2, distribuição uniforme **CLAHE**, **CLAHE** Tiles 8, **CLAHE** Clip limite 0.02, **CLAHE** bins histograma 200 e tamanho médio do bloco de filtro 3. O processo de aprimoramento fez com que os vasos sanguíneos parecessem mais escuros e melhor delineados, mantendo a aparência geral da imagem. No entanto, a segmentação dos vasos sanguíneos alcançou o valor de (0.9418 Acc) porque a fóvea e os pigmentos também foram realçados, criando imprecisões no processo de classificação.

Figura 4.7: Aprimoramento na imagem "test". (a) Canal verde original da imagem. (b) Imagem após o filtro de *Wiener* e filtro de abertura. (c) Imagem depois do CLAHE (d) Imagem depois do filtro da média



Fonte: Elaborado pelo autor

Os melhores valores, em relação as métricas de avaliação, encontrados nos 3240 experimentos são submetidos a duzentas rodadas de treinamento e classificação. Assim, a potência ótima de ruído do filtro de *Wiener* e o tamanho do bloco de *Wiener* foram determinados como 0,015 e 2, respectivamente. Esses valores dos parâmetros produziram o melhor máximo, média e mínimo, com valores de desvio padrão semelhantes Tab. 4.4 e Tab. 4.5. Os melhores valores são classificados de forma estatística com: a melhor média,

o menor desvio padrão e o maior valor de máximo. Desta forma, os resultados obtidos a partir dos parâmetros escolhidos refletem visualmente na Fig. 4.7, em que é possível notar as melhoras significativas na etapa de pré-processamento da imagem.

Tabela 4.4: Análise de potência de ruído do filtro *Wiener* por métrica Acc

	0.01	0.015	0.02
Máximo	0.95587	0.95614	0.95588
Média	0.94396	0.94401	0.94391
Mínimo	0.93121	0.93121	0.93099
Desvio Padrão	0.00637	0.00639	0.00639

Tabela 4.5: Análise de tamanho do bloco *Wiener* por medida Acc

Tamanho do bloco <i>Wiener</i>	2x2	3x3
Média	0,971215	0,9713
Desvio padrão	0,000181	0,0002248
Mediana	0,9712	0,97135
Valor máximo	0,9716	0,9716
Valor mínimo	0,9709	0,9707

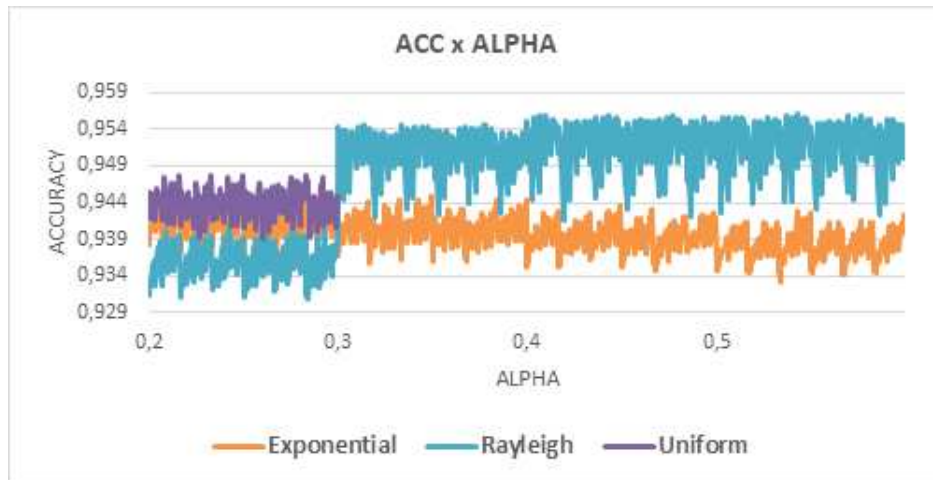
As estatísticas na Tab. 4.6 e Figura 4.8 para as distribuições exponencial, Rayleigh e uniforme, mostram o bom desempenho da distribuição Rayleigh. Como a distribuição uniforme não usa o parâmetro alfa, ela não foi plotada para todos os valores alfa na Fig. 4.8. Tanto a distribuição de Rayleigh quanto a exponencial contemplam o parâmetro alfa; no entanto, a distribuição exponencial foi superior apenas quando alfa é igual a 0.2, enquanto a distribuição de Rayleigh teve melhor desempenho em valores alfa mais altos.

Tabela 4.6: Análise de distribuição CLAHE por medida Acc

	Exponencial	rayleigh	uniform
Maximum	0.94669	0.95614	0.95588
Average	0.93996	0.94543	0.94648
Minimum	0.93315	0.93121	0.93099
STD	0.00236	0.00716	0.00640

Entre os valores testados, 25 *tiles* se mostraram ótimos para CLAHE (Tabela 4.4). Embora esse valor não tenha gerado o melhor desvio padrão mínimo ou o menor desvio padrão, ele gerou valores de Acc médio e máximo mais elevados (Tabela 4.8).

A otimização do cliplimit é semelhante, pois os valores testados produziram valores de Acc alternados (Tabela 4.9). Enquanto 0.04 se destacou para o mínimo e o desvio padrão, não tendeu a gerar as melhores medidas de Acc. Para resolver este problema, foi

Figura 4.8: Resultados usando o parâmetro α 

Fonte: Elaborado pelo autor

Tabela 4.7: Análise de bloco CLAHE por medida Acc

	8	14	18	25	31
Máximo	0.95507	0.95577	0.95587	0.95614	0.95567
Média	0.94209	0.94352	0.94450	0.94455	0.94515
Mínimo	0.93099	0.93260	0.93247	0.93369	0.93393
Desvio Padrão	0.00643	0.00593	0.00640	0.00652	0.00620

Tabela 4.8: Análise de Cliplimit CLAHE por medida Acc

	0.02	0.03	0.04
Máximo	0.95614	0.95544	0.95417
Média	0.94369	0.94453	0.94366
Mínimo	0.93099	0.93122	0.93235
Desvio Padrão	0.00787	0.00609	0.00477

contabilizado o número de testes com Acc maior que 0.95. Esse procedimento mostrou que 0.02 gerou a maior contagem, 321, em comparação com 0.03 e 0.04, que resultou em 287 e 129, respectivamente. Assim, 0.02 e 0.03 foram os valores dos parâmetros mais interessantes. No CLAHE alfa, 0.5 foi ótimo (Tabela 4.9).

Tabela 4.9: Análise do parâmetro α CLAHE por medida Acc

	0.2	0.3	0.4	0.5
Máximo	0.94669	0.94449	0.95461	0.95614
Média	0.94125	0.93892	0.94349	0.95207
Mínimo	0.93572	0.93315	0.93099	0.94170
Desvio Padrão	0.00197	0.00205	0.00800	0.00327

Em relação ao número de bins no CLAHE (Tabela 4.10), os valores testados produziram resultados semelhantes, sendo 200 o melhor máximo e 256 o melhor Acc médio.

O desvio padrão foi menor em 256. O filtro médio foi ligeiramente melhor com blocos de tamanho 3×3 (Tabela 4.11).

Tabela 4.10: Análise CLAHE nBins por medida Acc

	200	256
Máximo	0.95614	0.95550
Média	0.94373	0.94419
Mínimo	0.93121	0.93099
Desvio Padrão	0.00709	0.00558

Tabela 4.11: Análise do tamanho médio do bloco de filtro por medida Acc

	3	4
Máximo	0.95614	0.95558
Média	0.94438	0.94428
Mínimo	0.93121	0.93052
Desvio Padrão	0.00654	0.00665

A Tabela 4.12 compara os resultados do estudo atual com os de outros estudos de segmentação de vasos sanguíneos. A Tabela 4.13 mostra as métricas para cada imagem no conjunto de teste DRIVE após o treinamento e classificação da rede: (I) sem qualquer melhoria; (II) com a melhoria proposta, com base na verdade fundamental do primeiro observador; (III) com a melhoria proposta, utilizando a verdade fundamental do segundo observador.

As imagens 3 e 15 mostram os piores resultados, enquanto as imagens 9, 18 e 20 mostram os melhores resultados. A Figura 4.9 mostra os resultados da segmentação dessas imagens. Imagens com resultados inferiores têm fundo de olho menos uniforme devido às lesões e pigmentação. A imagem “03-test” também possui muitos vasos finos que dificultam a tomada de decisão da rede.

Como foi dito anteriormente, os parâmetros ideais foram definidos em: número de neurônios da camada oculta 128, tamanho de bloco 8×8 (64 neurônios na camada de entrada/saída), α como 0.000282 e η definido como 0.009125. O pré-processamento é baseado principalmente em CLAHE e filtro de Wiener, e a quantidade de épocas é 200. A Tabela 4.14 mostra algumas rodadas de treinamento variando apenas número de neurônios na camada oculta, uma vez que este tem um grande impacto no desempenho da rede, verifica-se que a partir de 128 neurônios não há garantias de que a Acc será maior, assim esse é o número adotado. Observou-se também que o tamanho de bloco 8×8 permite que vasos grossos, médios e finos sejam obtidos com grau de precisão aceitável.

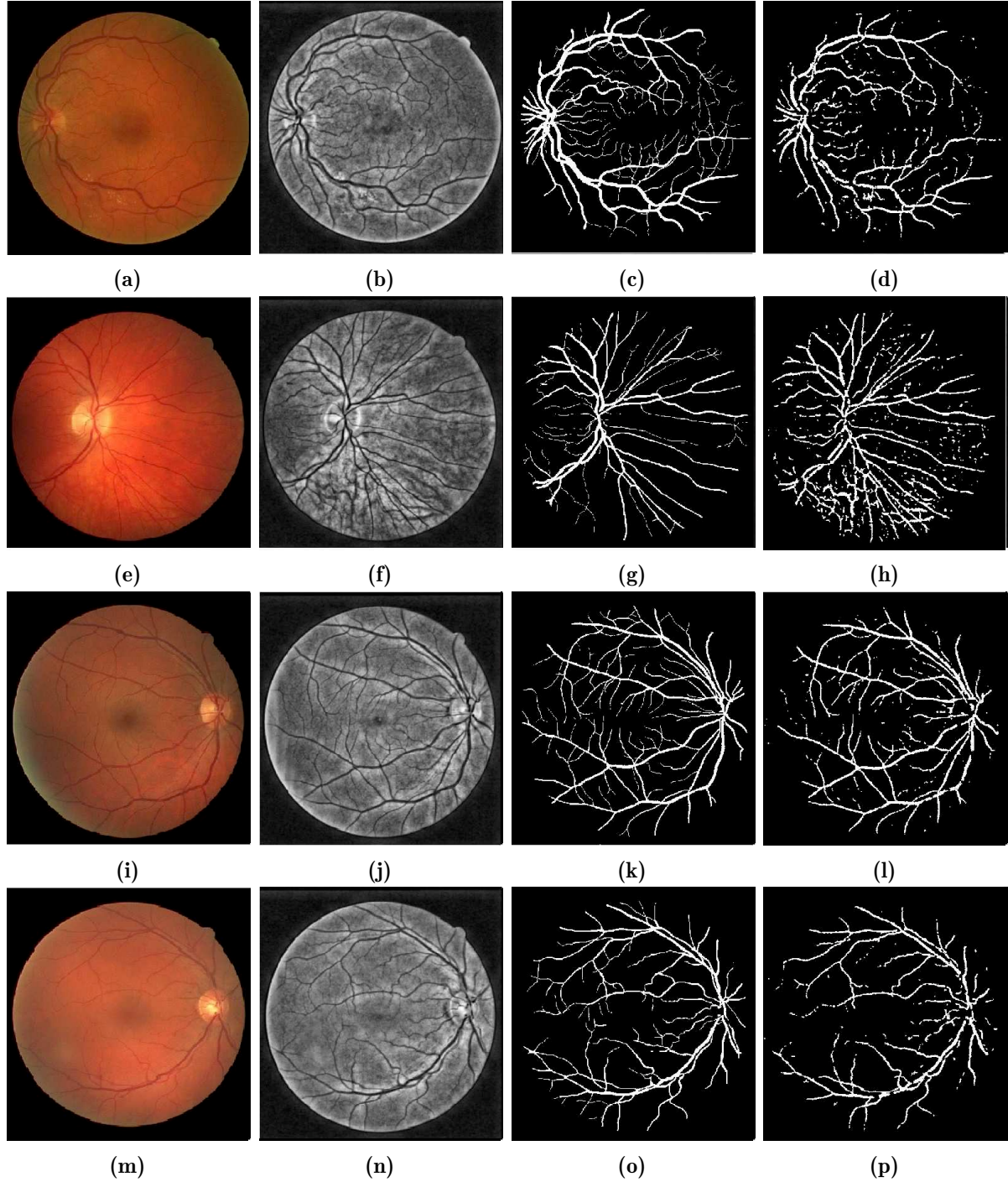
Tabela 4.12: Comparação do método proposto na base de dados DRIVE com métodos do estado da arte.

Métodos	Se	Sp	Acc
Método proposto (primeiro observador humano)	0,7564	0,9696	0,9505
Método proposto (2º observador humano)	0,7841	0,9695	0,9534
(LISKOWSKI; KRAWIEC, 2016)	0,7819	0,9748	0,9472
(MAJI et al., 2016)	—	—	0,9470
(AZZOPARDI et al., 2015)	0,7655	0,9704	0,9442
(FRAZ et al., 2012c)	0,7302	0,9742	0,9422
(VEGA et al., 2015)	0,7444	0,9600	0,9412
(ORLANDO; BLASCHKO, 2014)	0,7897	0,9684	—
(HASSAN et al., 2015)	0,7210	0,9710	0,9388
(ELBALAOUI et al., 2016)	0,7630	0,9764	0,9433

Tabela 4.13: Tabela com as métricas de avaliação

Image	Unenhanced image			First human observer			Second human observer		
	Se	Sp	Acc	Se	Sp	Acc	Se	Sp	Acc
1	0.56495	0.98698	0.93185	0.80224	0.97138	0.95612	0.82096	0.97166	0.95833
2	0.58869	0.99270	0.93214	0.79967	0.97618	0.95790	0.82682	0.97781	0.96242
3	0.63296	0.97631	0.92625	0.77886	0.95754	0.93966	0.84342	0.95518	0.94520
4	0.62371	0.98790	0.93877	0.76912	0.97399	0.95499	0.79749	0.97313	0.95762
5	0.60090	0.98934	0.93602	0.75476	0.97865	0.95740	0.83360	0.97568	0.96399
6	0.46562	0.99495	0.91951	0.67185	0.98455	0.95368	0.68719	0.98307	0.95512
7	0.57174	0.98643	0.93090	0.75982	0.96633	0.94721	0.85146	0.95782	0.95010
8	0.49466	0.98524	0.92341	0.70415	0.97017	0.94696	0.79359	0.96241	0.95100
9	0.27243	0.99860	0.91243	0.60461	0.99025	0.95861	0.60524	0.99013	0.95865
10	0.65061	0.98248	0.94242	0.77707	0.97010	0.95400	0.84099	0.96649	0.95737
11	0.58387	0.98861	0.93554	0.76069	0.96635	0.94777	0.81428	0.96619	0.95344
12	0.52468	0.98923	0.93049	0.76525	0.97380	0.95554	0.80306	0.97223	0.95846
13	0.40187	0.99664	0.91144	0.68561	0.98259	0.95314	0.68047	0.98460	0.95338
14	0.62619	0.98192	0.93990	0.82003	0.96113	0.94966	0.85616	0.95884	0.95112
15	0.79017	0.88597	0.87589	0.88414	0.89204	0.89147	0.87797	0.89412	0.89290
16	0.46726	0.99422	0.92457	0.71518	0.98159	0.95723	0.75942	0.98177	0.96260
17	0.24709	0.99771	0.90497	0.64717	0.98394	0.95515	0.72207	0.98275	0.96324
18	0.41589	0.99361	0.92658	0.74419	0.97706	0.95837	0.72745	0.98554	0.96152
19	0.78923	0.96803	0.94629	0.87896	0.96006	0.95324	0.82473	0.96940	0.95484
20	0.55665	0.98705	0.94066	0.80441	0.97386	0.96125	0.71651	0.98134	0.95652
Average	0.54346	0.98320	0.92650	0.75639	0.96958	0.95047	0.78414	0.96951	0.95339

A Tabela 4.14 mostra os resultados das métricas *sensitivity* (*SN*), *specificity* (*SP*) e *accuracy* (*Acc*) para cada imagem do conjunto de testes da base **DRIVE**. Os resultados são apresentados para o primeiro e segundo observador. Levando em consideração que a base apresenta a segmentação manual dos vasos sanguíneos realizada por dois especialistas, as imagens foram classificadas frente a segmentação de cada um deles. É interessante observar que a rede mostrou maior sintonia com a classificação do segundo observador, uma vez

Figura 4.9: Resultado da segmentação dos vasos sanguíneos

A primeira coluna contém imagens originais: (a) "03_test", (e) "15_test", (i) "19_test", (m) "20_test"; na segunda coluna, os respectivos pré-processamentos, na terceira coluna os respectivos *ground truths* e na última coluna os respectivos resultados obtidos pelo método.

que o *Acc* médio foi melhor. Em relação ao primeiro observador, as imagens "03_test" e "15_test" mostram os piores resultados, enquanto as imagens "19_test" e "20_test" mostram os melhores resultados. A imagem "19_test" apresentou a melhor *sensitivity*, detonando a boa capacidade da rede em encontrar os vasos sanguíneos na referida imagem,

no entanto, esta métrica deve ser analisada em conjunto com a *specificity*, para a qual se observa o valor 0.9785, o que resulta em uma *Acc* de 0.96, o melhor resultado obtido dentre as imagens da base DRIVE. Já a imagem “15_test” apresenta uma *sensitivity* de 0,7793 com uma *specificity* de 0.9449, o que significa que a rede encontrou muitos vasos, mas cometeu muitos erros ao considerar outras estruturas como sendo vasos, isso resulta em uma *Acc* de 0,93, bem abaixo da média geral que é 0.9487. Já para o segundo observador os melhores resultados foram nas imagens “04_test” e “10_test”, e o pior resultado também foi na imagem “15_test”: *Acc* de 0.9316, bem abaixo do restante.

Tabela 4.14: *Accuracy* em função do número de neurônios na camada oculta.

Bloco	α	η	NCO	<i>Acc</i>
8×8	0,000282	0,009125	32	0,9450
			64	0,9468
			80	0,9480
			100	0,9483
			128	0,9487
			140	0,9484
			200	0,9488

A Tabela 4.13 sintetiza um comparativo entre as medidas obtidas de diversos trabalhos, considerando a sensibilidade, especificidade, *accuracy* e área sob a curva ROC. Observa-se que os resultados obtidos pelo método proposto se sobrepõem aos resultados de estudos baseados em aprendizado de máquina, como é o caso de (LISKOWSKI; KRAWIEC, 2016), (MAJI et al., 2016), (VEGA et al., 2015), e outros métodos ((AZZOPARDI et al., 2015), (FRAZ et al., 2012c), (MARÍN et al., 2011), (HASSANIEN et al., 2015), (NGUYEN et al., 2013), etc). Mas não consegue superar os trabalhos de (AL-RAWI; KARAJEH, 2007), (RICCI; PERFETTI, 2007), (LUPAŞCU et al., 2010), (FATHI; NAGHSH-NILCHI, 2013) e (WANG et al., 2015). (AL-RAWI; KARAJEH, 2007) unem dois excelentes métodos que são o *matched filter* e aprendizado por meio de métodos bioinspirados, já (RICCI; PERFETTI, 2007) usam operadores de linha e SVM; (LUPAŞCU et al., 2010) usam informações sobre a estrutura de intensidade local, propriedades espaciais e geometria em múltiplas escalas e um classificador AdaBoost; (FATHI; NAGHSH-NILCHI, 2013) usam wavelet e operador de estrutura circular e (WANG et al., 2015) utilizam CNN e *Random Forest*. Em outras palavras, o método proposto é superado apenas por combinações de métodos reconhecidamente eficientes, ressaltando se dessa forma a importância do pré-processamento e ajuste apresentados.

O filtro *Wiener* é eficiente para imagens de fundo do olho porque usa estimativa adaptativa em *pixels* de uma vizinhança local, o que permite o controle do filtro em resposta à espessura dos vasos sanguíneos. Da mesma forma, **CLAHE** também normaliza o histograma dentro de uma vizinhança limitada. Assim, o maior desafio do presente estudo envolveu a determinação de parâmetros ótimos. Se os parâmetros apropriados forem escolhidos, um modelo de segmentação de vasos sanguíneos supervisionado pode aumentar significativamente a capacidade de generalização da rede. No entanto, se forem escolhidos parâmetros inadequados, o processo de convergência pode ser lento e os resultados não satisfatórios. O alto número de parâmetros nos processos de aprimoramento significava que a faixa de ajustes também era alta e que pequenas variações poderiam produzir variações significativas nos resultados. Portanto, foi testado o ruído de potência do filtro *Wiener* e o tamanho do bloco; a distribuição **CLAHE**, *tiles*, clip limit, alpha, (distribuições Rayleigh e exponencial), número de bins e o filtro da média e tamanho do bloco. Além disso, a rede foi treinada e o Acc foi avaliado para cada conjunto de hiperparâmetros, resultando em 3240 rodadas de treinamento e classificação. Descobriu-se que o método supervisionado produziu os melhores aprimoramentos na qualidade da imagem de fundo do olho para a segmentação dos vasos sanguíneos.

Com relação à configuração da rede, descobriu-se vasos grossos e médios e muitos dos vasos finos nas imagens de entrada puderam ser descobertos quando blocos de 8×8 foram usados. Os vasos finos não foram detectados quando blocos muito grandes foram usados, enquanto os blocos muito pequenos produziram a sobrecarga da rede. Em geral, os métodos baseados na aprendizagem têm dificuldade em identificar vasos finos devido às suas semelhanças com outras estruturas. Uma proporção ideal de neurônios na camada oculta pode ser calculada a partir de um determinado tamanho de bloco e um número conhecido de neurônios na camada de entrada. Por meio desse processo, descobriu-se que o número de 128 neurônios na camada oculta gerou excelentes resultados. O método de Kolmogorov apresentou resultados semelhantes, para escolha do número de neurônios na camada oculta, com $NHL = 2n + 1$, onde n é o número de neurônios na camada de entrada (FAUSETT, 1994). Valores mais altos não produziram ganhos significativos no treinamento da rede.

4.5 Considerações Finais

Observadores humanos e métodos de classificação requerem diferentes definições do que é uma imagem aprimorada. Para um ser humano, as estruturas em uma imagem precisam apenas se tornar mais nítidas, enquanto uma rede neural requer maior diferenciação entre os elementos da imagem. A principal vantagem do método proposto é que ele considera as imagens do fundo de olho dentro de um contexto aplicado e emprega métricas de classificação de vasos sanguíneos adequadas para garantir a qualidade dos realces. Métricas como PSNR e SSIM, bem como algumas outras não se traduzem efetivamente em melhor segmentação, pois desconsideram os principais problemas com as imagens do fundo de olho, ou seja, baixo contraste e equilíbrio de ruído.

O desempenho da segmentação dos vasos sanguíneos do método proposto foi equivalente ao dos métodos mais modernos. Nosso método produziu medidas de especificidade e precisão melhores do que aquelas da arquitetura de rede de aprendizado profundo proposta por (LISKOWSKI; KRAWIEC, 2016) e de quase todos os outros estudos. Assim, este trabalho apresenta um modelo consistente para realce da imagem do fundo de olho que pode ser usado para segmentar vasos sanguíneos em imagens do fundo de olho por um método supervisionado. Estudos futuros podem examinar ajustes de pré-processamento com uma CNN.

CAPÍTULO 5

SEGMENTAÇÃO AUTOMÁTICA DOS VASOS SANGUÍNEOS COMBINANDO FILTRAGEM DE DOMÍNIO DUPLO E UMA ARQUITETURA CNN UNET REDUZIDA.

As imagens biomédicas ocupam um lugar especial na maioria dos procedimentos médicos, são utilizadas no diagnóstico, na triagem, no tratamento e na avaliação da evolução de muitas doenças, inclusive em patologias oftalmológicas (IKRAM *et al.*, 2013). Neste contexto, as imagens de fundo do olho são analisadas e interpretadas por um médico especialista e necessariamente tem uma influência direta na tomada de decisão (ABRÀMOFF *et al.*, 2010). Nem sempre a análise é simples e rápida, pode requerer um grande esforço e concentração por parte do especialista, isso acontece por exemplo, na avaliação do grau de retinopatia diabética que contempla a contagem cuidadosa de pequenos pontos chamados microaneurismas (NARAYANAN *et al.*, 2020).

As diversas patologias retinianas geram sinais que evidenciam sua presença, no entanto, algo comum na identificação de muitas destas patologias é a separação dos sinais das estruturas anatômicas presente no fundo de olho (MOOKIAH *et al.*, 2013). Os principais biomarcadores são os vasos sanguíneos, a fóvea, a mácula e o disco óptico, e os principais sinais encontrados em imagens do fundo de olho são as hemorragias, os microaneurismas, a tortuosidade dos vasos sanguíneos e os exsudatos (ABRÀMOFF *et al.*, 2010). A segmentação automática de estruturas fisiológicas ou biomarcadores pode contribuir para um diagnóstico rápido evitando o desperdício de recursos financeiros e de tempo (GOEL *et al.*, 2017).

A estrutura vascular é um dos principais marcadores pois mudanças nessa estrutura

são indicadores de enfermidades como retinopatia diabética, degeneração macular relacionada a idade, doença cardiovascular, dentre outras, e muitas destas são silenciosas e podem comprometer a visão de modo irreversível (AL-BANDER et al., 2018). Dessa forma, os vasos sanguíneos são um passo chave para compor sistema de segmentação automáticos (GOPINATH et al., 2014).

Um dos aspectos importantes do processamento de imagens digitais é diferenciar os aspectos relevantes da imagem (em imagens do fundo de olho são sinais e biomarcadores) dos irrelevantes (em geral o fundo), além de ruído (ZHOU et al., 2018). A partir desta distinção da imagem e do ruído é possível melhorar o resultado do processamento da imagem, seja ele apenas para segmentação ou com o objetivo de classificar (GOEL et al., 2017), (ZHOU et al., 2018), (DASH; SENAPATI, 2020). Assim, uma segmentação precisa dos vasos sanguíneos é primordial para o estudo e acompanhamento das enfermidades oculares (HASHEMZADEH; Adlpour Azar, 2019). O mapa extraído dos vasos sanguíneos pode ser analisado para verificar mudanças na estrutura ou ainda ser removido da imagem retiniana evidenciando outros marcadores e sinais (UPADHYAY et al., 2020a).

Há alguns desafios importantes em atividades que envolvem a segmentação de imagens do fundo de olho. Na própria aquisição observa-se o baixo contraste, problemas de iluminação e ruído, isso ocorre em função da própria natureza dos tecidos no fundo de olho, do formato esférico da retina e do ruído advindo do equipamento de captura e sensores (ABRÀMOFF et al., 2010). Já na tarefa de segmentar observa-se dificuldade em relação a quantidade de dados disponíveis e rotulados com seus respectivos *ground truth* (HOOVER, 2000), (STAAL et al., 2004).

Neste sentido, os sistemas CAD estão em crescente desenvolvimento devido a dois fatores relevantes. O primeiro é o avanço gerado na tecnologia atual, e o segundo é a evolução no desenvolvimento de algoritmos baseados em aprendizagem de máquina (SENGUPTA et al., 2020). Uma revisão profunda sobre os algoritmos de aprendizado profundo e o impacto das metodologias que utilizam técnicas de pré-processamento e de pós-processamento dos dados antes de submeter as imagens a arquitetura de rede neural profunda foi apresentada por (SALVI et al., 2021). Na mesma linha de investigação, o trabalho de (SENGUPTA et al., 2020) realiza uma revisão dos métodos baseados em aprendizado profundo cujo objetivo é desenvolver ferramentas que possam auxiliar os profissionais em retina nas tarefas de segmentação, classificação ou detecção.

Em especial as abordagens que utilizam os modelos de **CNN** profundas têm alcançado bons resultados em quase todas as diferentes áreas e tarefas do conhecimento científico (TAGHANAKI et al., 2019). Por exemplo, o desempenho na tarefa de detecção e distinção da gravidade da retinopatia diabética é apresentado no trabalho (NARAYANAN et al., 2020). Assim é proposto uma arquitetura híbrida, com a finalidade de detectar e classificar o nível da gravidade da retinopatia diabética. A metodologia é composta por mais de uma **CNN**, e a abordagem com análise de componentes principais é usada para reduzir a dimensão das imagens. Na sequência um **SVM** é utilizado na construção de uma nova estratégia. O desempenho é avaliado em diferentes condições de pré-processamento e os resultados superam o estado da arte na tarefa de detecção e na tarefa de distinguir a gravidade da retinopatia diabética.

Portanto, o presente capítulo apresenta uma metodologia para segmentar os vasos sanguíneos da retina, com objetivo de lidar com os desafios e problemas relacionados a imagens do fundo de olho. O objetivo do método proposto é investigar se a modificação da arquitetura **CNN** Unet, com redução do número de hiperparâmetros, em conjunto com estratégias de aprimoramento dual (tempo/frequência), podem aumentar o desempenho da segmentação dos vasos sanguíneos. Outro ponto de investigação é trabalhar com o modelo de cores da imagem, e assim usar todas as informações com eficiência. As contribuições foram resumidas da seguinte forma:

- Uma nova estratégia para aprimorar imagens, que envolve trabalhar com algoritmos no domínio do tempo e no domínio da frequência combinando as melhores informações.
- Adaptação e o aprimoramento das informações de cores no modelo RGB explorando todas as informações presentes em imagens de retina.
- Configuração e personalização da arquitetura **CNN** Unet original para um modelo reduzido, mas capaz de manter a precisão da segmentação dos vasos sanguíneos.

A estrutura do capítulo segue da seguinte forma: a seção "Trabalhos relacionados" analisa os métodos mais recentes baseados em métodos supervisionados e não supervisionados para segmentar os vasos sanguíneos da retina. Na seção "Métodos" é descrita o método proposto e as estratégias definidas para o pré-processamento e o processamento

principal. Nas seções "Discussões e Conclusões" os resultados e conclusões são apresentados, e por fim, a seção de "Considerações Finais".

5.1 Trabalhos Relacionados

Muitos métodos e técnicas foram propostos para segmentação dos vasos sanguíneos em imagens de fundo do olho (TAGHANAKI et al., 2019), (AGURTO et al., 2015), (WANG et al., 2021), (Saha Tchinda et al., 2021), (TAMIM et al., 2020). De modo geral estas técnicas podem ser resumidas ou categorizadas em dois tipos, métodos supervisionados, e métodos não supervisionado (IMRAN et al., 2019).

As abordagens que utilizam os métodos supervisionados como ponto de partida necessitam que as amostras dos *pixels* dos vasos sanguíneos e os não vasos sanguíneos sejam classificados por um oftalmologista (ALMOTIRI et al., 2018). A partir desta marcação é possível construir modelos com diferentes finalidades, como por exemplo, segmentação, detecção ou classificação (MOOKIAH et al., 2013).

A grande maioria dos trabalhos atuais estão focados em algoritmos de aprendizado de máquinas (ALMOTIRI et al., 2018). Uma das grandes vantagens em utilizar estes algoritmos é a possibilidade de construir algoritmos de segmentação poderosos e que podem ser ajustados mediante a presença de novas informações (RAZZAK et al., 2018). No trabalho de (ZHANG; CHUNG, 2018) o método desenvolvido com CNN é proposto para segmentar os vasos sanguíneos em imagens da retina. A arquitetura Unet é adaptada para focar sua atenção na fronteira dos vasos sanguíneos. Durante a fase de treinamento é inserido um mecanismo com reconhecimento da borda para caracterizar as informações da rede. Em outra perspectiva, uma CNN profunda e múltipla *Multiple Deep Convolutional Neural Network* (MDCNN) é construída e treinada para realizar uma tarefa de detecção nas imagens do fundo de olho (GUO et al., 2018). O problema da detecção é formulado como um problema de classificação múltipla, em que uma estratégia de aprendizagem por meio de incremento melhora o desempenho da rede. Em uma outra estratégia, o trabalho (TIAN et al., 2020) usa uma CNN de múltiplos caminhos para segmentar os vasos sanguíneos. A rede proposta utiliza duas estratégias para realizar a tarefa de segmentação. Na primeira, um filtro gaussiano de baixa frequência é usado para determinar as características gerais dos vasos sanguíneos. Segundo um filtro gaussiano de alta frequência é usado para determinar as características locais. No trabalho de (Saha Tchinda et al., 2021) um conjunto de filtros

clássicos focam suas delimitações na detecção das bordas. No passo inicial, o vetor de características é obtido utilizando os filtros de detecção de bordas. No passo seguinte, os recursos obtidos são destinados ao treinamento da ANN, com objetivo de discriminar *pixels* que pertence aos vasos sanguíneos e *pixels* que não pertencem aos vasos sanguíneos. Um novo modelo para segmentar os vasos sanguíneos da retina é apresentado por (WANG et al., 2021). O modelo combina *Nest Unet* e *patch-learning*, e para etapa de treinamento uma estratégia de extração das amostras é implementada visando a segmentação dos vasos de calibre menor.

Na imagem da retina os padrões de textura semelhante aos vasos sanguíneos são considerados um problema para a tarefa de segmentação. Na abordagem proposta por (UPADHYAY et al., 2021) um novo algoritmo foi proposto para segmentar os vasos sanguíneos, composto por dois métodos conhecidos na literatura, a transformada *wavelet gabor* 2-D e o popular *gaussian matched filtering*. Para cada nível da *wavelet*, os padrões de textura são suprimidos usando o *Local Laplacian filter* e assim o aprimoramento da imagem tem objetivo de destacar os vasos sanguíneos. Na sequência as imagens aprimoradas são combinadas e servem de entrada para a etapa de treinamento da arquitetura Unet. A morfologia dos vasos sanguíneos grossos e finos é tratada no trabalho de (UPADHYAY et al., 2020b). Uma abordagem multiescala utilizou transformada *wavelet* direcional para gerar imagens em três canais. A extração dos vasos finos é obtida extraindo uma imagem composta apenas por coeficientes reais da transformada *wavelet*. Assim uma abordagem em múltiplas escalas permite destacar em diferentes escalas os vasos finos. Na sequência as imagens aprimoradas são submetidas a Unet para classificação.

No trabalho de (LI et al., 2016) uma CNN profunda com uma poderosa capacidade de indução, somada a uma estratégia de treinamento eficiente, reduz a incerteza causada pelo ruído da imagem, e da patologia. A segmentação supervisionada dos vasos sanguíneos foi proposta no trabalho (GANESAN et al., 2020). A abordagem proposta usa os descritores de forma baseados em momento de *Zernike*. Um vetor de recurso com onze dimensões composto por recursos estatísticos e recursos de forma é calculado. Os coeficientes ideais do momento de *Zernike* são derivados com relação a sua diferenciabilidade máxima entre os vasos sanguíneos e os detalhes do fundo da imagem. Na abordagem proposta por (KHAN et al., 2020), a segmentação automática dos vasos retinianos é obtida por uma estratégia que combina as etapas de pré-processamento, segmentação, aprimoramento baseado na

estrutura do vaso e uma etapa de pós-processamento. Neste sentido, o pré-processamento elimina o ruído e melhora o contraste da imagem. Em seguida, a segmentação usa uma análise de dados auto-organizados iterativa modificada (*MISO-DATA*) e dessa forma obtém uma imagem binária que é fundida com a imagem segmentada utilizando o filtro *B-COSFIRE* (*Combination Of Shifted Filter Responses*). No trabalho de (HASHEMZADEH; Adlpour Azar, 2019), a segmentação dos vasos sanguíneos é resolvida combinando dois métodos derivados do aprendizado de máquinas. O método não supervisionado é utilizado para detectar os vasos mais grossos e o método supervisionado é utilizado para detectar os vasos mais finos. O principal objetivo desta abordagem é atenuar o problema de alta variância intra classe dos recursos da imagem.

Na literatura atual existem muitas técnicas que visam o aprimoramento da imagem, das quais o algoritmo **CLAHE** é utilizado com objetivo de melhorar o contraste das imagens. Além disso o algoritmo **CLAHE**, com seus diferentes ajustes empíricos tem sido usado com sucesso em muitas abordagens de segmentação. Técnicas de realce de contraste podem ser listadas em: realce de contraste, normalização local, equalização de histograma limitado por contraste, mascaramento linear não agudo e transformações de contorno. No trabalho de (WANG et al., 2021) é proposto um método que aprimora a imagem do fundo de olho atuando de forma simultânea na correção de iluminação, no aprimoramento de detalhes e na supressão de ruído. Inicialmente uma estimativa global do ruído na imagem permite dividir a imagem em três camadas: a camada base, o detalhe e o ruído. Na camada de base os problemas de iluminação são corrigidos, na camada de detalhe a supressão de artefatos é realizado no canal azul, e na camada de ruído a supressão é alcançada descartando esta camada antes de recompor a imagem. O método proposto apresenta um bom desempenho na correção dos problemas de iluminação, aumento de detalhes e supressão de ruído ou artefato. No trabalho de (BANDARA; GIRAGAMA, 2018) desenvolveram uma técnica de aprimoramento integrada com algoritmo Tyler Coye e um algoritmo de reconstrução dos vasos sanguíneos baseados na transformada de linha de *Hough*. A Tabela 5.1 a seguir resume os trabalhos baseados em métodos supervisionados e seus respectivos resultados com base nas métricas mais utilizadas.

Contudo, as abordagens não supervisionadas são capazes de classificar os *pixels* de vasos sanguíneos ou não vasos sanguíneos sem nenhum dado de treinamento rotulado. As abordagens baseadas em métodos não supervisionados são capazes de extrair informações

Tabela 5.1: Tabela de Métodos Supervisionados

Methods	Database	Se	Sp	Acc
(ZHANG; CHUNG, 2018)	DRIVE	0.8723	0.9618	0.9504
(ZHANG; CHUNG, 2018)	STARE	0.7673	0.9901	0.9712
(GUO et al., 2019)	DRIVE	0.7891	0.9804	0.9561
(GUO et al., 2019)	STARE	0.8212	0.9843	0.9674
(TIAN et al., 2020)	DRIVE	0.8639	0.9690	0.9580
(Saha Tchinda et al., 2021)	DRIVE	0.7552	0.9775	0.9480
(Saha Tchinda et al., 2021)	STARE	0.7265	0.9759	0.9548
(WANG et al., 2021)	DRIVE	0.8060	0.9869	0.9512
(WANG et al., 2021)	STARE	0.8230	0.9945	0.9641
(UPADHYAY et al., 2020b)	DRIVE	0.789	0.972	0.956
(UPADHYAY et al., 2020b)	STARE	0.736	0.981	0.961
(UPADHYAY et al., 2020b)	HRF	0.750	0.972	0.952
(LI et al., 2016)	DRIVE	0.7569	0.9816	0.9527
(GANESAN et al., 2020)	DRIVE			0.945
(GANESAN et al., 2020)	STARE			0.9486
(KHAN et al., 2020)	DRIVE	0.766	0.972	0.954
(HASHEMZADEH; Adlpour Azar, 2019)	DRIVE			0.9531
(HASHEMZADEH; Adlpour Azar, 2019)	STARE			0.9752
(BANDARA; GIRAGAMA, 2018)	DRIVE			0.9411
(BANDARA; GIRAGAMA, 2018)	DRIVE			0.9489
(SOOMRO et al., 2019)	DRIVE	0.87		0.956
(SOOMRO et al., 2019)	STARE	0.808		0.954
(SOOMRO et al., 2019)	CHASE_DB1	0.886		0.976
(SOOMRO et al., 2019)	HRF	0.829		0.962

das imagens por meio de resposta de filtro ou outras técnicas baseadas em modelo. Portanto, no contexto de retinografias digitais de extrair as características como a estrutura dos vasos sanguíneos e posteriormente classificar a nível de pixel.

No trabalho de (WANG et al., 2019) é apresentado um método não supervisionado para realizar automaticamente a extração dos vasos sanguíneos. O método propõe um aprimoramento na região dos vasos denominada *revised top-bottom-hat transformation*. Esta abordagem no aprimoramento funciona como um recurso de achatamento da elipse circunscrita mínima para reconhecer o vaso sanguíneo. Um novo método não supervisionado integrando transformada *wavelet* discreta (DWT) com algoritmo de *Tyler Coye* é proposto por (DASH; SENAPATI, 2020) para segmentar de forma automática os vasos sanguíneos. A metodologia proposta é dividida em quatro fases. Na primeira fase, a transformada *wavelet* discreta (DWT) é aplicada nas imagens da retina. Na segunda fase, as imagens decompostas pela DWT são aprimoradas aplicando a correção gama. Na terceira fase,

a conversão das imagens aprimoradas para o espaço de cores YIQ. Por fim, o algoritmo *Tyler Coye* é usado para detectar os vasos sanguíneos e medir seus parâmetros.

Em outra abordagem (GOUR; KHANNA, 2017) apresentou uma proposta para segmentar os vasos sanguíneos em quatro etapas. A primeira etapa é composta por técnicas de pré-processamento, a segunda pela transformada morfológica *top-hat*, a terceira, uma filtragem mediana híbrida, e por fim, a etapa de pós-processamento. Em uma abordagem diferenciada, o trabalho apresentado por (SOLÍS-PÉREZ et al., 2019) propôs um novo método baseado em *kernel* gaussiano fracionário envolvendo a derivada fracionária de *Atangana-Baleanu* para detectar os vasos sanguíneos em imagens do fundo de olho.

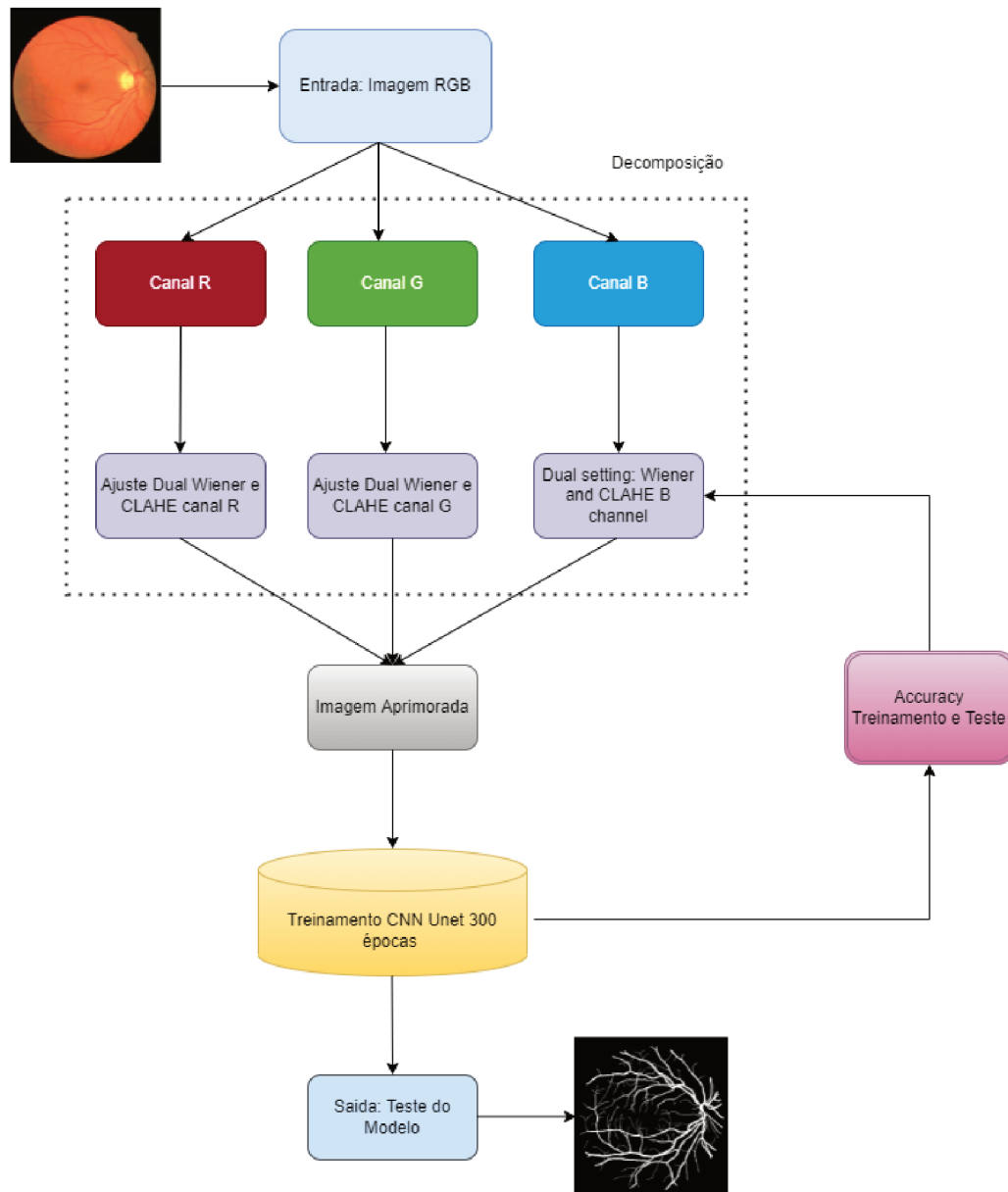
A Tabela 5.2 a seguir resume os resultados dos trabalhos apresentados, baseados em métodos não supervisionados.

Tabela 5.2: Tabela de Métodos não supervisionados

Méthods		Se	Sp	Acc
(WANG et al., 2019)	DRIVE			0.9446
(WANG et al., 2019)	STARE			0.9503
(DASH; SENAPATI, 2020)	DRIVE	0.7403	0.9905	0.9661
(GOUR; KHANNA, 2017)	DRIVE	0,7215	0,9730	0,9409
(GOUR; KHANNA, 2017)	STARE	0,7405	0,9649	0,9369
(SOLÍS-PÉREZ et al., 2019)	DRIVE	0.8270		0.9558

5.2 Métodos

A segmentação automática dos vasos sanguíneos da retina pode ser utilizada para compor sistemas de diagnóstico auxiliados por computador e assim acelera o processo de triagem de várias enfermidades que afetam o globo ocular. Entretanto, muitas pesquisas foram realizadas para segmentar os vasos sanguíneos, mas os problemas que destacamos na seção de introdução, como ruído e baixo contraste interferem de forma direta no desempenho dos algoritmos de segmentação. Neste sentido, o método proposto combina filtragem de domínio duplo com uma arquitetura CNN e é avaliado nos bancos de dados DRIVE e STARE. Assim, na Figura 5.1, as etapas do método proposto são divididas em quatro estágios: a preparação do banco de dados, o pré-processamento, a adaptação da arquitetura CNN Unet para imagens RGB, e por fim, as métricas de avaliação.

Figura 5.1: Metodologia geral para segmentação dos vasos sanguíneos

Fonte: Elaborado pelo autor

5.2.1 Banco de Dados

A descrição inicial da base é formada por 40 imagens do fundo de olho (584×565), sendo que 20 imagens são destinadas para o conjunto de treinamento e 20 imagens para o conjunto de teste e validação. Além disso, o conjunto com 20 imagens é rotulado por dois especialistas e a estrutura detalhada dos vasos sanguíneos é delineada e serve como um padrão ouro para construção de novos algoritmos. A segmentação do primeiro especialista é usada como padrão ouro (NIEMEIJER et al., 2004). Para este trabalho a base foi aumentada e redistribuída, conforme a seguir:

- Na primeira parte: as imagens da pasta de treinamento (20 imagens), e seus respectivos *ground truth* foram espelhadas (ou seja, rotacionadas em 180^0), gerando 20 novas imagens, totalizando 40 imagens.
- Na segunda etapa: como o tamanho inicial da imagem era 584 linhas e 565 colunas, foi necessário recortar o fundo da imagem. Essa operação ocorreu nas linhas, aproveitando o intervalo de linhas 9 : 574, gerando assim uma imagem de tamanho 565 linhas por 565 colunas.
- Na terceira etapa: após o espelhamento, um total de 60 imagens foram obtidas. Destas imagens, aproximadamente 76%, 15% e 9% foram para as fases treinamento, teste e validação respectivamente. conforme mostrado na Tabela 5.3.

Tabela 5.3: Aumento do bando de dados

Banco de dados	Total	Treinamento	Teste	Validação
Drive	60	46	9	5
Ground truth	60	46	9	5

- Na quarta etapa, para aumentar o número de imagens para as fases de treinamento, teste e validação, foi necessário o janelamento das imagens, o qual será explicado na fase de pré-processamento, gerando assim, a Tabela 5.4.

Tabela 5.4: Número de paths para treinamento e teste

Banco de dados	Total	Treinamento	Teste	Validação
Drive	540	414	81	45
Ground truth	540	414	81	45

5.2.2 Pré-Processamento

A fase de pré-processamento é considerada uma etapa importante no processamento de imagens, pois é nesta fase que se pode atenuar ou destacar características de interesse para o processamento principal (ZHOU et al., 2018). A combinação de filtros pode destacar as imagens do fundo de olho, atenuando os ruídos de alta e baixa frequência e regulando de forma eficiente a luminosidade e o contraste das imagens do fundo de olho. Assim, foram adaptados os filtros *Wiener* e *CLAHE* para aprimorar as imagens do fundo de olho. A próxima seção traz uma breve descrição dos filtros e seus parâmetros.

5.2.3 Filtros

De modo geral, a modelagem dos filtros no domínio da frequência utilizam como ponto de partida uma modificação da transformada de *Fourier* e o teorema da convolução (GONZALEZ; WOODS, 2009). Essa modificação é realizada em função de dois parâmetros: a magnitude do espectro e o ângulo de fase (GONZALEZ; WOODS, 2009). O filtro de *Wiener* pertence a uma classe importante denominada de sistemas lineares invariantes com o tempo. Estes sistemas exercem um papel fundamental no processamento de sinais digitais. Em especial podem ser aplicados na reconstrução de imagens bidimensionais corrompidas por ruído. Os parâmetros do filtro *Wiener*, como tamanho do *kernel* e estimativa da potência do ruído podem ser analisadas empiricamente por meio de biblioteca presentes no *matlab* e no *python* (GONZALEZ; WOODS, 2009).

Uma versão muito conhecida na literatura para trabalhar com ajuste de intensidade local e global é denominada de **CLAHE**. Neste algoritmo o brilho da imagem é distribuído e a amplificação do ruído é limitada. No processo de aplicação do algoritmo **CLAHE** a imagem é dividida em pequenos blocos sobrepostos. Em seguida, para cada bloco da imagem é aplicado a equalização do histograma aumentando o contraste local. Na sequência um limite de corte na altura do histograma é utilizado para conter a amplificação do ruído antes de calcular a função de distribuição acumulada da imagem (MA et al., 2017). Por fim, a forma do histograma de cada bloco é reorganizada de acordo com uma distribuição (uniforme, Rayleigh ou exponencial). Neste trabalho, o número de experimentos é combinado em razão dos parâmetros avaliados para cada etapa de filtragem. Assim, o espaço de busca é composto por 90 possíveis combinações. Os valores dos hiperparâmetros são baseados nos trabalhos de (HASHEMZADEH; Adlpour Azar, 2019), (ZHOU et al., 2018) e (SANTOS et al., 2020), e resumidos na Tabela 5.5. O objetivo principal é aprimorar as imagens com eficiência na etapa de treinamento da **CNN** Unet proposta.

Os algoritmos **CLAHE** e filtro *Wiener* foram utilizados em muitos cenários diferentes, mas com um objetivo em comum, aprimorar no sentido de processamento digital de sinais, os diferentes sinais adquiridos, seja eles um sinal unidimensional ou bidimensional (GHAEL et al., 1997), (ESCORCIA-GUTIERREZ et al., 2020). Embora seja evidente o sucesso obtido pelos métodos mencionados anteriormente, na tarefa de segmentação dos vasos sanguíneos, ainda existe problemas que envolvem a estruturas grossas e finas no sentido da

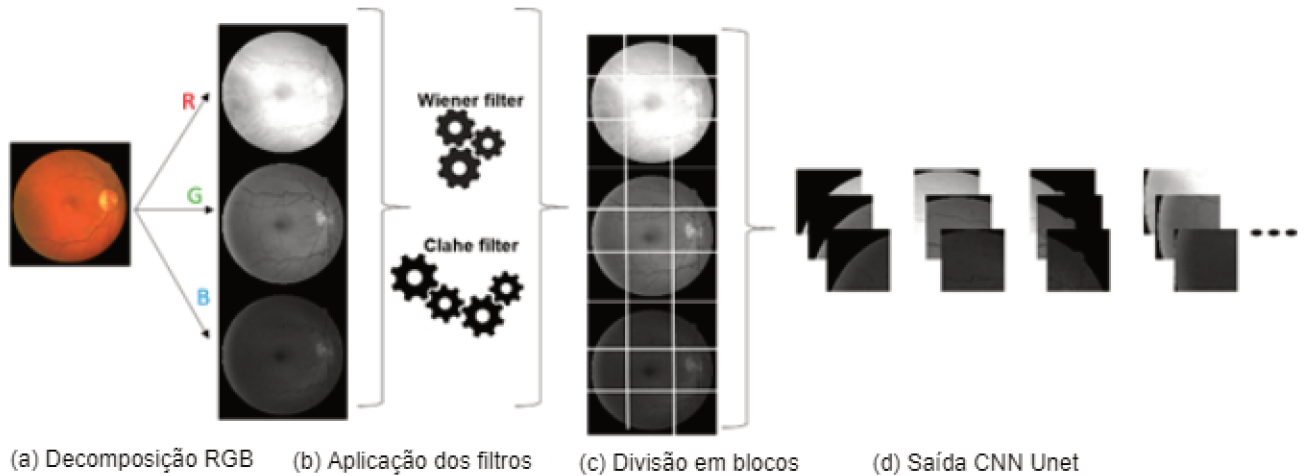
Tabela 5.5: Parâmetros para os filtros *Wiener* e **CLAHE**

Parâmetros	Valores
Wiener power noise	0.01, 0.015, 0.02
CLAHE clip limit	0.02, 0.03, 0.04
Parâmetros	Valores
Wiener block size	2, 3
CLAHE tiles	8, 14, 18, 25, 31

conectividade, o contraste pobre e a variação de iluminação não estão totalmente superadas (MOCCIA et al., 2018), (KHAN et al., 2019).

5.2.4 Esquema do pré-processamento

A etapa de pré-processamento (Figura 5.2) foi dividida em quatro sub-etapas: Decomposição do canal, Aplicação dos filtros, janelamento da imagem e image *patch*:

Figura 5.2: Diagrama da fase de pré-processamento

Fonte: Elaborado pelo autor

Decomposição do canal

Nesta sub-etapa, a imagem do fundo de olho é decomposta nos canais RGB gerando assim uma imagem para cada canal conforme a Figura 5.2 (a). Assim cada informação da imagem está descrita no plano de cores e quando combinados digitalmente representam as matrizes de intensidades ou cores observados na imagem.

Aplicação dos filtros

Na sub-etapa de aplicação dos filtros, as três imagens geradas a partir dos canais RGB são submetidas à dois processos de filtragem (Figura 5.2 (b)). O primeiro processo é

uma filtragem no domínio da frequência utilizando o filtro *Wiener*, e no segundo processo é uma filtragem baseada no algoritmo **CLAHE**. Ambos os processos precisam de um ajuste empírico para o bom desempenho do processo de filtragem, Tabela 4.3. O resultado final da combinação ótima de ambos os filtros é dado pela Equação 5.1:

$$I_{filter}(x,y) = I_{redwienerCLAHE}(x,y) + I_{greenwienerCLAHE}(x,y) + I_{bluwienerCLAHE}(x,y) \quad (5.1)$$

Resumidamente, nesta etapa o ruído do processo de aquisição da imagem é minimizado por meio da filtragem *Wiener* com os parâmetros da Tabela 4.3. Em seguida, o contraste dos vasos sanguíneos é aprimorado utilizando o **CLAHE** com os parâmetros da Tabela 4.3.

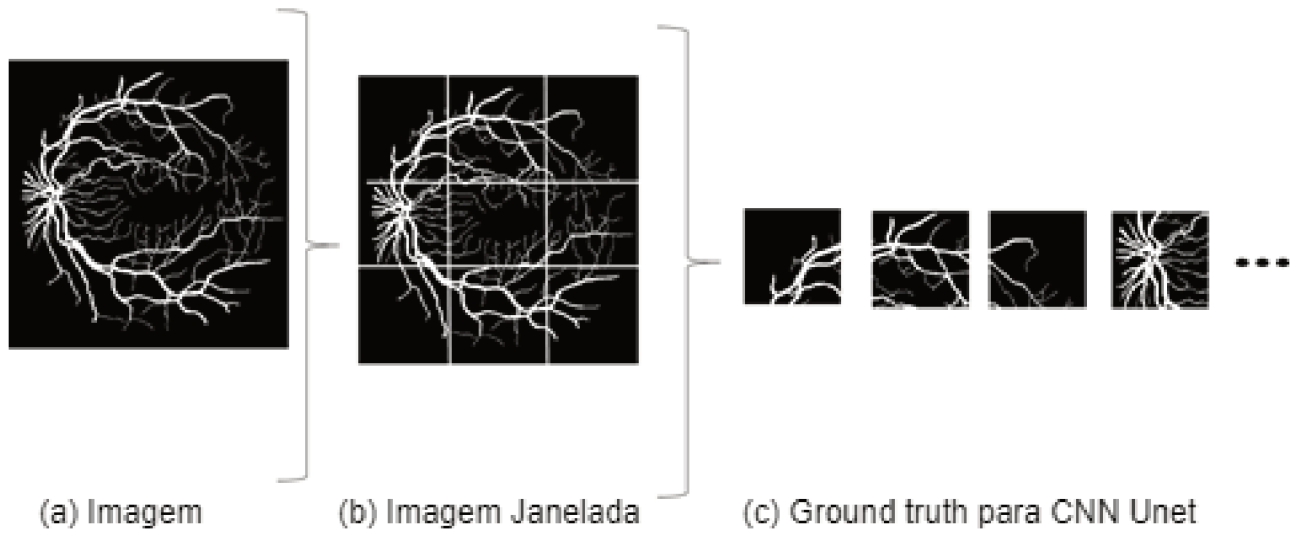
Divisão em blocos

Com a subetapa divisão em blocos, devido a quantidade limitada de imagens foi necessário fazer o aumento de dados para tornar o treinamento mais robusto. Assim, cada imagem foi então subdividida em 3 blocos por linhas e 3 blocos por colunas, resultando em 9 blocos (cada bloco tem o tamanho 188×188), gerando uma base de imagens com um total de 540 imagens. Conforme é apresentado na Tabela 5.4 e na Figura 4.2(B). O *ground truth* também foi dividido de forma análoga a imagem de entrada. O processo de janelamento do *ground truth* é apresentado na Figura 5.3(C) e na Tabela 5.4. Os blocos de imagem são submetidos a entrada da **CNN Unet**.

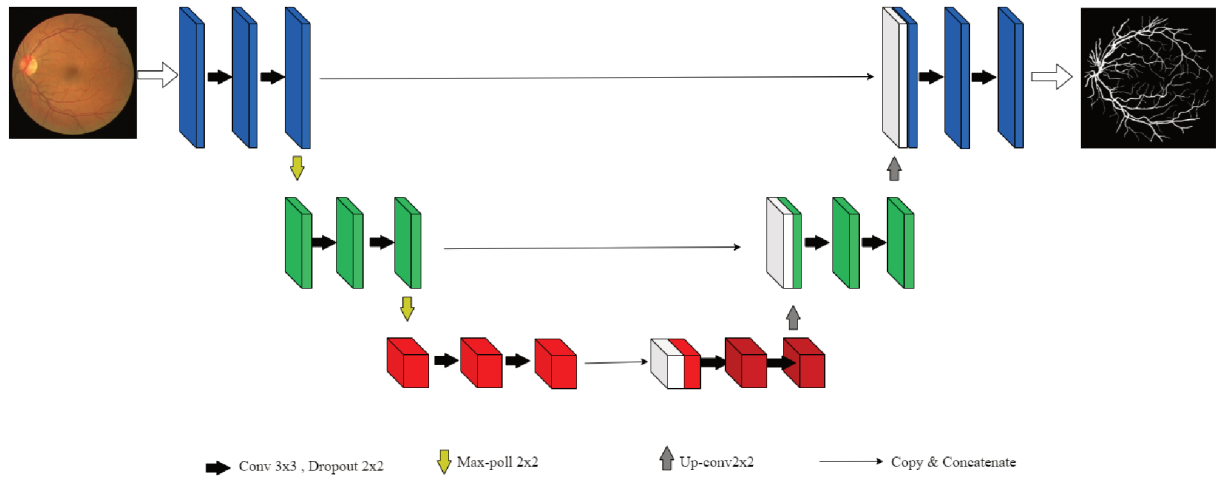
5.2.5 CNN Unet

As camadas presentes neste modelo de **CNN Unet** consistem em camadas de dupla convolução com função de ativação Relu, seguida de *Dropout*, e aplicação de *max-pooling*, isto no processo de contração da rede. No processo de expansão temos novamente a dupla convolução seguida de Relu e a aplicação *up-sampling* e a concatenação entre camadas de contração e expansão. Nesta fase, o treinamento da **CNN Unet** é realizado com as imagens aprimoradas da etapa de pré-processamento. A primeira parte da arquitetura Unet, Figura 5.4, pode ser vista como uma estrutura convencional de **CNN**, em que a informação de entrada é submetida a convolução com diferentes *kernels* e assim as informações são extraídas formando vários mapas de características.

Na fase de validação de acordo com a Tabela 5.4 tem-se cinco imagens da base

Figura 5.3: Divisão do *ground truth* em blocos

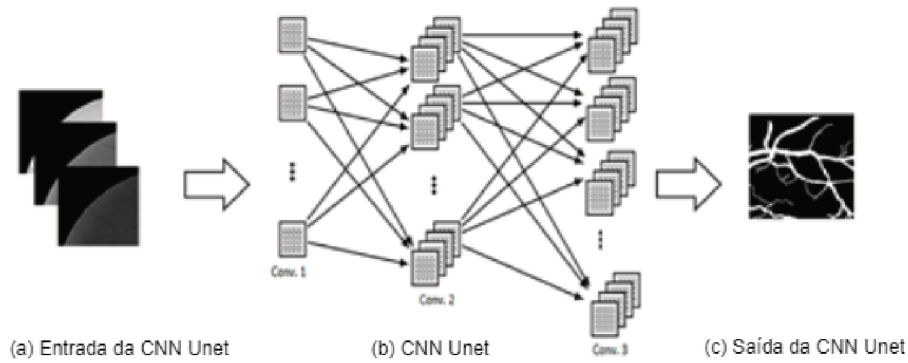
Fonte: Elaborado pelo autor

Figura 5.4: Arquitetura da **CNN** Unet utilizada

Fonte: Elaborado pelo autor

DRIVE e cinco no *ground truth*. Essas imagens foram decompostas em 45 subimagens na fase de pré-processamento. As Figuras 5.5 e 5.6 apresentam um exemplo de segmentação de uma imagem da base drive na fase de validação. Assim, uma imagem da base drive gera 9 sub-imagens decompostas em três canais, formando assim a entrada da **CNN** Unet Figura 5.5 (a).

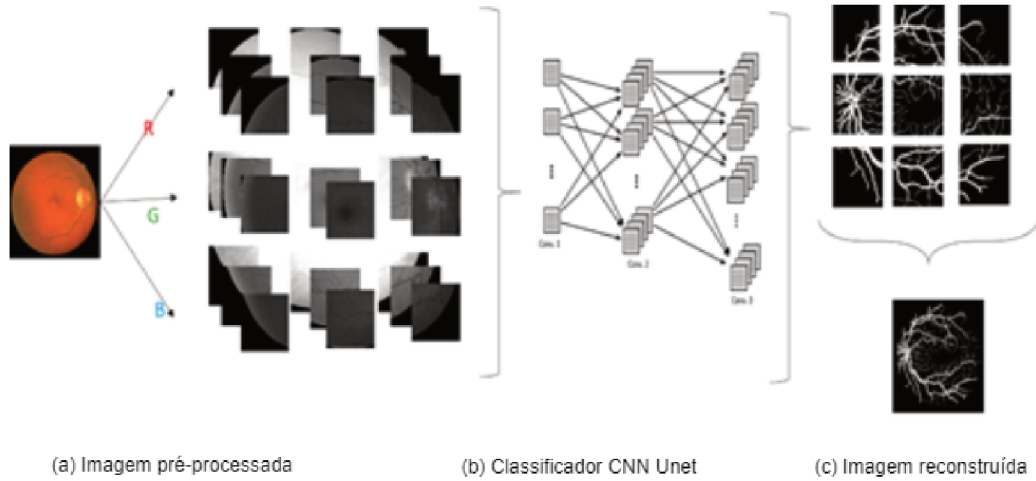
Cada entrada gerada da **CNN** Unet é processada de forma individual, formando uma subimagem segmentada, conforme mostra a Figura 5.6. No primeiro treinamento os pesos são inicializados de forma aleatória. A rede foi treinada usando o método de otimização de gradiente Adam (**KINGMA; BA, 2015**). Os parâmetros $\beta_1 = 0,9$ e $\beta_2 = 0,999$, são um

Figura 5.5: Esquema de segmentação da CNN Unet

Fonte: Elaborado pelo autor

padrão e a taxa de aprendizado inicial $\alpha = 0,001$, decaída por um fator de 10 (KINGMA; BA, 2015) e são utilizados no método.

Assim, após a segmentação das nove subimagens da Figura 5.5 (a), tem-se nove subimagens segmentadas conforme a Figura 5.5 (c). Na sequência é possível reconstruir a imagem segmentada de forma completa. A imagem reconstruída e o *ground truth* do primeiro especialista são usadas no cálculo das métricas.

Figura 5.6: Esquemas de entrada e saída da rede Unet proposta

Fonte: Elaborado pelo autor

5.3 Experimentos

O treinamento inicial das imagens foram executado em Python em um computador com Windows 10 de 64 bits (CPU Intel(R) Core™ i7-8750H @2,20 GHz, 6 core(s), 16

GB de RAM e uma placa gráfica GeForce GTX 1050Ti) e posteriormente adaptado para ambiente "Colab" que é um serviço para processamento de algoritmos de *machine learning* disponível na nuvem.

A configuração da arquitetura **CNN** e a definição dos hiperparâmetros são determinadas com base no entendimento do problema a ser resolvido. Neste sentido, os experimentos iniciais definiram a profundidade da **CNN** em conjunto com as noventa possibilidades de configuração apresentadas na Tabela 5.5. Assim, o experimento inicial usa 300 épocas para cada treinamento com um tempo médio de 19 minutos. O tempo total de treinamento nesta fase inicial é de $90 \times 19 = 1710/60$, ou seja, aproximadamente *28h30min*.

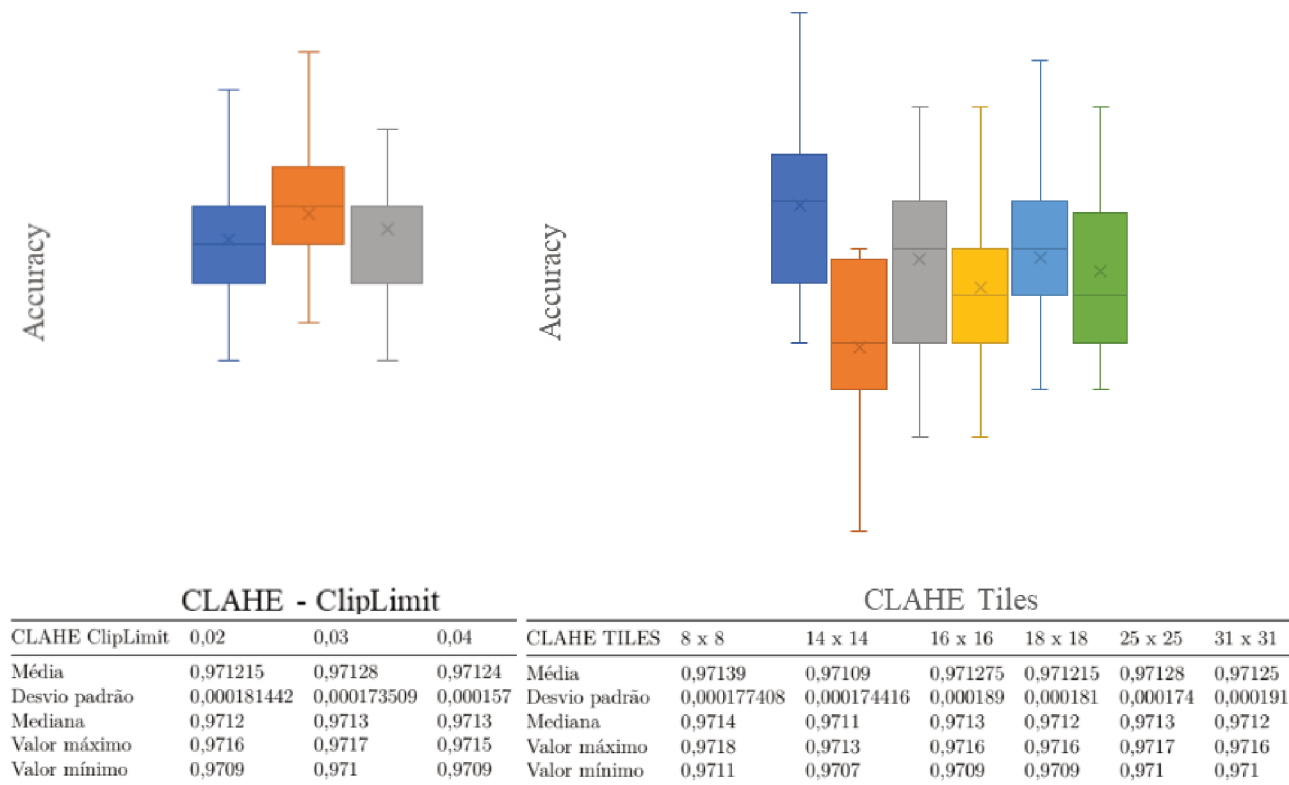
5.4 Resultados e Discussões

Na Figura 5.7 analisamos o desempenho de dois hiperparâmetros *CLAHE ClipLimit* e *Tiles* usando a Tabela 5.5. A consistência dos hiperparâmetros são verificados com 100 execuções. Assim, os tempos médios registrados são de *37h14min* e *18h41min* na fase de treinamento. Os valores mais precisos foram considerados com a melhor média e o menor desvio padrão, os outros valores de mediana e valor máximo e valor mínimo compõem o experimento para checar a estabilidade dos valores analisados na Figura 5.7. Assim, o valor ClipLimit de 0.03 e Tiles de 8×8 apresentam a melhor configuração para os hiperparâmetros. Este valores contrastam com os valores determinados em (HASHEMZADEH; Adlpour Azar, 2019), CLipLimite 0.01 e Tiles 8×8 .

Os vasos são mais espessos na origem e mais finos nos limites da retina, assim é necessário fornecer a **CNN** informações com boa precisão destes elementos para que treinamento seja robusto e o classificador final seja eficiente. Deste modo, os hiperparâmetros *Wiener power noise* e *Wiener block size* são determinados usando a Tabela 5.5 e a consistência dos valores são verificados com 100 repetições. Os tempos médios registrados são de *18h37min* e *12h25min* respectivamente. Na Figura 5.8 os hiperparâmetros mais regulares são 0,02 (estima a potência do ruído) para Wiener power noise e 3×3 (determina o tamanho da janela do filtro).

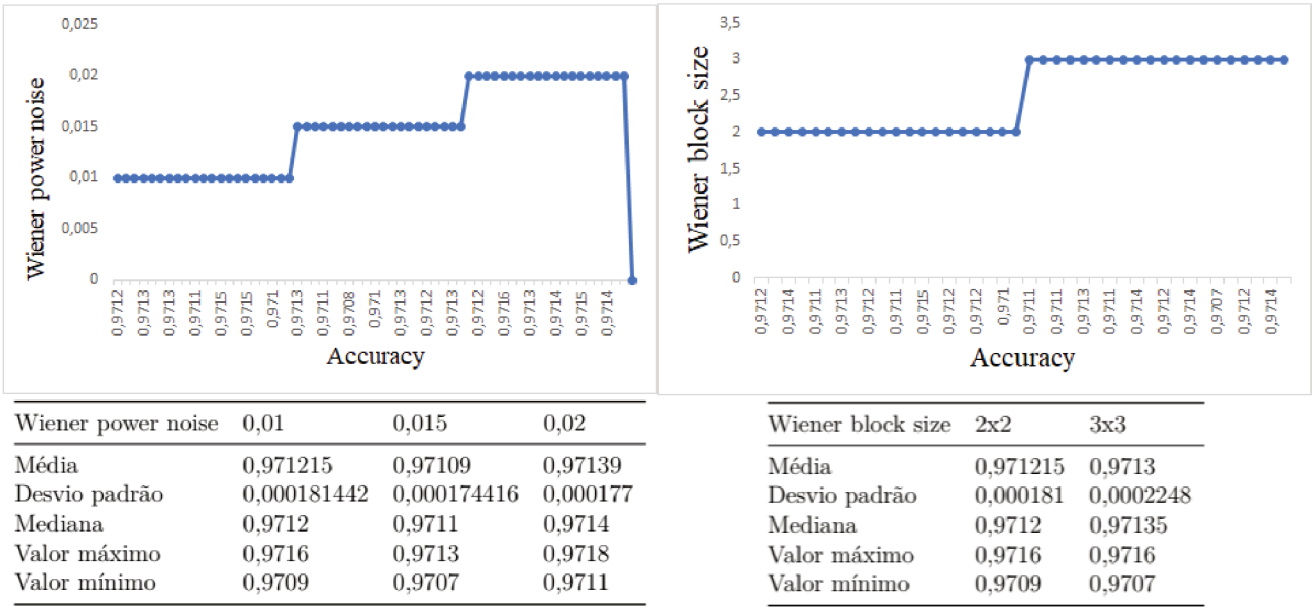
Muitos trabalhos utilizam somente o canal verde da imagem para compor sua estratégia de segmentação (Saha Tchinda et al., 2021), (FAN; MO, 2016), (SONALI et al., 2019). Entretanto, a estratégia proposta combina os algoritmos da fase de pré-

Figura 5.7: Análise dos parâmetros **CLAHE** no processo de melhoria do contraste da imagem da retina



Fonte: Elaborado pelo autor

Figura 5.8: Análise dos parâmetros do filtro Wiener no processo de atenuação dos ruídos presentes na imagem da retina

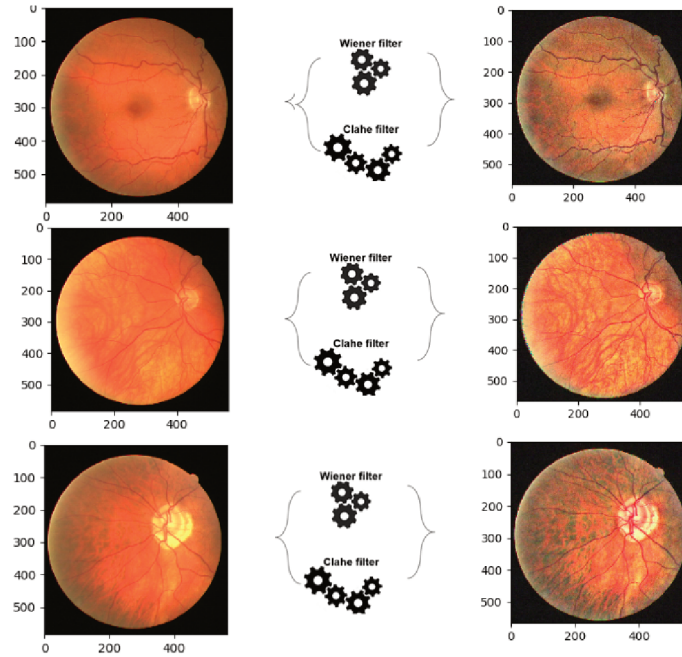


Fonte: Elaborado pelo autor

processamento e os canais RGB. Na Figura 5.9 o destaque dos elementos como vasos sanguíneos, mácula e disco óptico são evidentes. Essas imagens são utilizadas no processo

de treinamento do modelo **CNN** Unet com foco nos vasos sanguíneos.

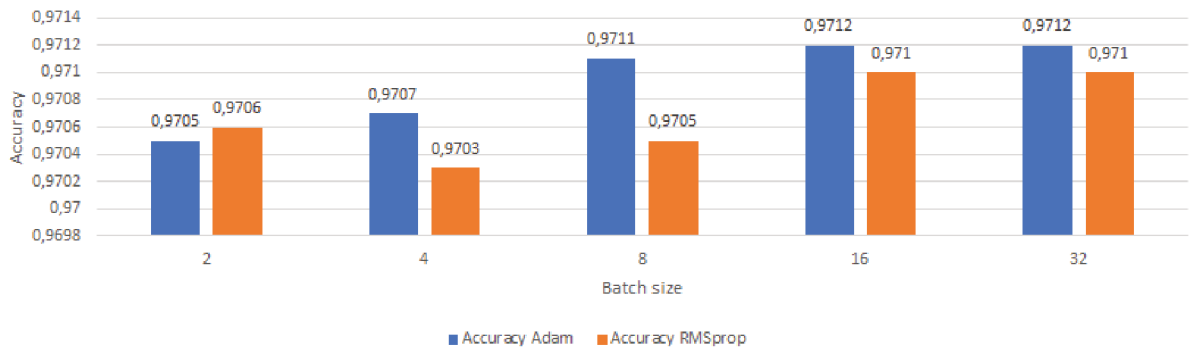
Figura 5.9: Resultado do processo de aprimoramento da imagem usando **CLAHE** e filtro *Wiener*



Fonte: Elaborado pelo autor

A avaliação dos algoritmos de otimização de gradientes *Adam* e *RMSprop* são aplicadas no modelo **CNN** Unet proposto e o tempo de processamento foi de 21h15min com 100 repetições. A Figura 5.10 apresenta o melhor desempenho para o algoritmo Adam, e os valores de *batch size* 16 e 32 com melhores medidas de *accuracy*.

Figura 5.10: Análise dos algoritmos de otimização *Adam* e *RMSProp*

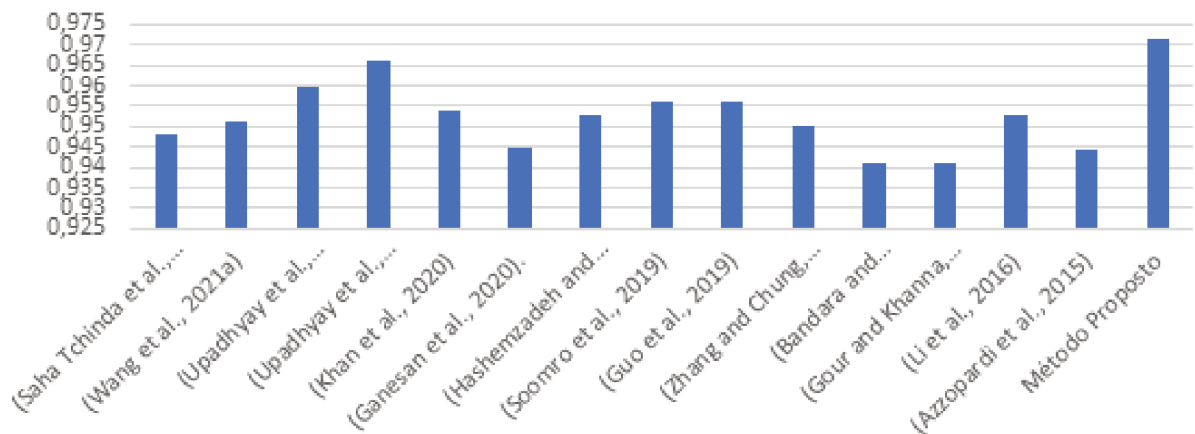


Fonte: Elaborado pelo autor

A arquitetura **CNN** Unet proposta até aqui é uma variação reduzida da arquitetura original com variações na estratégia de extrair os mapas de características. Os valores dos hiperparâmetros investigados e determinados nas seções anteriores consumiram um tempo total de 107h32min. A **CNN** Unet proposta tem como objetivo segmentar novas imagens

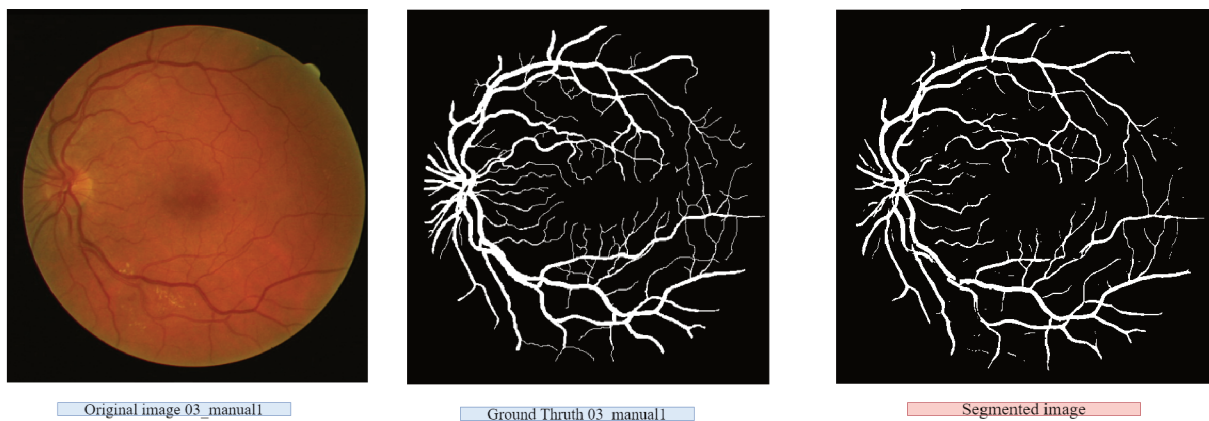
da retina. Assim, a rede foi executada por 100 repetições e o tempo de processamento de 31h20min. Portanto, o tempo total para compor o algoritmo de segmentação é de 136h42min. Os resultados em relação a tarefa de segmentação dos vasos sanguíneos é apresentado na Tabela 5.11 e na Figura 5.12. As primeiras impressões visuais, indicam excelente segmentação da estrutura vascular geral compostas por veias e artérias e *accuracy* média de segmentação de 0.9712.

Figura 5.11: Comparação da segmentação utilizando a CNN Unet proposta neste trabalho



Fonte: Elaborado pelo autor

Figura 5.12: Aplicação da CNN Unet treinada na imagem de test 03-manual



Fonte: Elaborado pelo autor

A Tabela 5.6 apresenta o resumo para os algoritmos baseados em métodos supervisionados e não supervisionados. Todas as estratégias propostas utilizam as estratégias de CNN ou ANN para desenvolver e aprimorar o algoritmo de segmentação dos vasos sanguíneos. Assim, a CNN Unet proposta neste trabalho superar tanto os algoritmos supervisionados quanto os não supervisionados.

Tabela 5.6: Comparação do método proposto com métodos supervisionados e não supervisionados.

Métodos supervisionados	Acc	Métodos não supervisionados	Acc
(ZHANG; CHUNG, 2018)	0.9504	(WANG et al., 2019)	0.9446
(GUO et al., 2019)	0.9561	(DASH; SENAPATI, 2020)	0.9661
(TIAN et al., 2020)	0.9580	(GOUR; KHANNA, 2017)	0.9409
(Saha Tchinda et al., 2021)	0.9480	(SOLÍS-PÉREZ et al., 2019)	0.9558
(WANG et al., 2021)	0.9512	Método proposto	0.9713
(UPADHYAY et al., 2020b)	0.952		
Método proposto	0.9713		

5.5 Considerações Finais

A modelagem proposta neste trabalho realizou a segmentação dos vasos sanguíneos da retina, combinando os melhores parâmetros para aprimorar as imagens do fundo de olho, e na sequência estas imagens constituem a entrada de uma arquitetura **CNN** Unet adaptada para imagens RGB. A **CNN** Unet proposta usa estratégias como convolução e função de ativação *Relu*, seguida de *dropout*, *max pooling*, concatenação e *upsampling* para construir o modelo de segmentação. Contudo, vale destacar que a adaptação do modelo reduziu a carga computacional em relação a arquitetura original. A **CNN** Unet adaptada tem por objetivo combinar as operações de convolução, função Relu, pooling e dropout durante todo o processo de contração da rede. Esta escolha de operações faz com que uma rede neural diferente seja usada para cada pequeno *mini-batch* de exemplos de treinamento. Assim, a **CNN** Unet proposta pode se tornar generalista para a tarefa de segmentação. Outro aspecto importante apresentado é a etapa de pré-processamento das imagens, que podem ser facilmente adaptadas para outros algoritmos. Os resultados experimentais mostram que é possível explorar técnicas de extração de características em sintonia com algoritmos que atuam no tempo e na frequência. Para comunidade de diagnóstico de doenças da retina em que a dificuldade é a disponibilidade de dados relevantes é muito importante a concepção deste tipo de algoritmo. O desempenho da **CNN** Unet é equivalente aos trabalhos mais modernos e pode compor sistema **CAD** e contribuir diagnósticos automatizados de doenças oculares como retinopatia diabética e glaucoma.

O número de enfermidades oculares como retinopatia diabética e glaucoma afetam muitas pessoas em todo o mundo. Neste cenário, a etapa de diagnóstico destas enfermidades é essencial para a tomada de decisão e realização do tratamento para evitar e controlar o processo de degeneração ou cegueira total. A tarefa de segmentação dos vasos sanguíneos é realizada por um médico especialista altamente treinado e pode levar um tempo razoável para a segmentação dos vasos sanguíneos. Assim, o desenvolvimento de algoritmos automáticos é uma tarefa importante para compor sistemas CAD. Nestes o tempo gasto para a tarefa de segmentação é realizado em poucos segundos.

Neste sentido, esta tese investigou se as estratégias baseadas ANN e CNN são capazes de aumentar o desempenho em segmentação de imagens de fundo de olho, com ênfase nos vasos sanguíneos em termos de métricas de avaliação, como sensibilidade, especificidade e precisão, que são utilizadas em imagens médicas. Assim, foram apresentadas duas contribuições para obter uma segmentação precisa dos vasos sanguíneos. Ambas podem ser incorporadas em sistemas de diagnósticos automáticos.

Na primeira, descrevemos uma estratégia para melhorar a qualidade da imagem para a detecção de vasos sanguíneos usando CLAHE, um filtro de Wiener e uma rede MLP. Este processo depende do ajuste cuidadoso do pré-processamento da imagem para que os vasos sanguíneos possam ser classificados pela rede neural artificial quanto à precisão, especificidade e sensibilidade. A rede treinada com os parâmetros ideais indicam um desempenho da tarefa de segmentação dos vasos sanguíneos equivalente a métodos mais modernos.

Na segunda, uma arquitetura discriminativa supervisionada CNN Unet foi adotada e

adaptada como algoritmo de segmentação dos vasos sanguíneos para imagens no formato rgb. Assim, o modelo de aprendizagem proposto é composto de seis camadas com a capacidade de detectar boas características nas imagens e formar hierarquias de características não lineares. Os resultados preliminares, em relação a métrica de precisão, para o banco de dados [DRIVE](#), indica um valor de 0.9713, superior ao todos os valores de precisão dos trabalhos revisados nesta pesquisa. Por fim os métodos de análise de imagem da retina desenvolvidos até este momento têm grande potencial não apenas para fornecer um diagnóstico menos subjetivo, mas também acelerar o processo de diagnóstico para uma variedade de doenças da retina, e conseqüentemente pode ser adaptado em um ambiente clínico.

Assim, as [ANN](#) e [CNN](#) apresentadas também serão ferramentas valiosas a serem usadas para outros problemas médicos desafiadores. Descobertas recentes têm indicado que lesões vasculares na retina com a forma de chama ou micro-hemorragias e infartos da camada de fibras nervosas, podem estar presentes em uma percentagem considerável de pacientes graves com COVID-19 ([LANI-LOUZADA et al., 2020](#)). Assim, como trabalho futuro, podemos agregar ao processo de segmentação, mais uma etapa, com a capacidade de classificar as micro-hemorragias.

REFERÊNCIAS BIBLIOGRÁFICAS

- ABDULLAH, M.; FRAZ, M. M.; BARMAN, S. A. Localization and segmentation of optic disc in retinal images using circular Hough transform and grow-cut algorithm. *PeerJ* 4:e2003, v. 4, n. 1, p. 1–22, 2016. Disponível em: <<https://doi.org/10.7717/peerj.2003>>.
- ABRÀMOFF, M.; GARVIN, M.; SONKA, M. Retinal Imaging and Image Analysis. *IEEE Reviews in Biomedical Engineering*, v. 3, p. 169–208, 2010. ISSN 1937-3333. Disponível em: <<https://doi.org/10.1109/RBME.2010.2084567>>.
- AGGARWAL, C. C. et al. Neural networks and deep learning. *Springer*, Springer, v. 10, p. 978–3, 2018.
- AGURTO, C.; YU, H.; MURRAY, V.; PATTICHIS, M. S.; NEMETH, S.; BARRIGA, S.; SOLIZ, P. A multiscale decomposition approach to detect abnormal vasculature in the optic disc. *Computerized Medical Imaging and Graphics*, Elsevier Ltd, v. 43, p. 137–149, 2015. ISSN 18790771. Disponível em: <<https://doi.org/10.1016/j.compmedimag.2015.01.001>>.
- AL-BANDER, B.; AL-NUAIMY, W.; WILLIAMS, B. M.; ZHENG, Y. Multiscale sequential convolutional neural networks for simultaneous detection of fovea and optic disc. *Biomedical Signal Processing and Control*, Elsevier Ltd, v. 40, p. 91–101, 2018. ISSN 17468108. Disponível em: <<http://dx.doi.org/10.1016/j.bspc.2017.09.008>>.
- AL-RAWI, M.; KARAJEH, H. Genetic algorithm matched filter optimization for automated detection of blood vessels from digital retinal images. *Computer methods and programs in biomedicine*, Elsevier, v. 87, n. 3, p. 248–253, 2007. Disponível em: <<https://doi.org/10.1016/j.cmpb.2007.05.012>>.
- ALMOTIRI, J.; ELLEITHY, K.; ELLEITHY, A. Retinal vessels segmentation techniques and algorithms: a survey. *Applied Sciences*, Multidisciplinary Digital Publishing Institute, v. 8, n. 2, p. 155, 2018. Disponível em: <<https://doi.org/10.3390/app8020155>>.
- ASAD, A. H.; AMRY, E. E.; HASSANIEN, A. E. Retinal vessels segmentation based on water flooding model. *2013 9th International Computer Engineering Conference (ICENCO)*, v. 1, n. 1, p. 43–48, 2013. Disponível em: <<http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htmarnumber=6736474>>.
- AZZOPARDI, G.; STRISCIUGLIO, N.; VENTO, M.; PETKOV, N. Trainable COSFIRE filters for vessel delineation with application to retinal images. *Medical Image Analysis*,

Elsevier B.V., v. 19, n. 1, p. 46–57, 2015. ISSN 13618423. Disponível em: <<http://dx.doi.org/10.1016/j.media.2014.08.002>>.

BALASUBRAMANIAN, K.; ANANTHAMOORTHY, N. P. Robust retinal blood vessel segmentation using convolutional neural network and support vector machine. *Journal of Ambient Intelligence and Humanized Computing*, Springer Berlin Heidelberg, v. 12, n. 3, p. 3559–3569, 2019. ISSN 18685145. Disponível em: <<https://doi.org/10.1007/s12652-019-01559-w>>.

BANDARA, A. M.; GIRAGAMA, P. W. A retinal image enhancement technique for blood vessel segmentation algorithm. *arXiv*, p. 1–5, 2018. ISSN 23318422. Disponível em: <<https://doi.org/10.1109/ICIINFS.2017.8300426>>.

BENGIO, Y.; GOODFELLOW, I.; COURVILLE, A. *Deep Learning*. [S.l.]: MIT press Massachusetts, USA, 2017. v. 1.

BHATTACHARJEE, U.; DAS, P. Performance Evaluation of Wiener Filter and Kalman Filter Combined with Spectral Subtraction in Speaker Verification System. *International Journal of Innovative Technology and Exploring Engineering (IJITEE)*, v. 2, n. 2, p. 108–112, 2013. Disponível em: <<https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.675.728&rep=rep1&type=pdf>>.

BISHOP, C. M. *Pattern Recognition and Machine Learning*. [S.l.: s.n.], 2006. v. 4. 738 p. Doi:<<http://dx.doi.org/10.1117/1.2819119>>. ISSN 10179909. ISBN 9780387310732.

CHAUDHURI, S.; CHATTERJEE, S.; KATZ, N.; NELSON, M.; GOLDBAUM, M. Detection of blood vessels in retinal images using two-dimensional matched filters. *IEEE transactions on medical imaging*, v. 8, n. 3, p. 263–269, 1989. ISSN 0278-0062. Doi:<<http://dx.doi.org/10.1109/42.34715>>.

CHOLLET, F. *Deep learning with Python*. [S.l.]: Simon and Schuster, 2021.

CHUTATAPE, O.; Liu Zheng; KRISHNAN, S. Retinal blood vessel detection and tracking by matched Gaussian and Kalman filters. v. 20, n. 6, p. 3144–3149, 2002. Disponível em: <<https://ieeexplore.ieee.org/abstract/document/746160>>.

CIREŞAN, D. C.; GIUSTI, A.; GAMBARDELLA, L. M.; SCHMIDHUBER, J. Deep neural networks segment neuronal membranes in electron microscopy images. *Advances in Neural Information Processing System (NIPS)*, p. 2852–2860, 2012. ISSN 10495258. Disponível em: <<https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.367.3014&rep=rep1&type=pdf>>.

CIREŞAN, D. C.; MEIER, U.; GAMBARDELLA, L. M.; SCHMIDHUBER, J. Deep, big, simple neural nets for handwritten digit recognition. *Neural Computation*, v. 22, n. 12, p. 3207–3220, 2010. ISSN 08997667. Disponível em: <https://doi.org/10.1162/NECO_a_00052>.

CONDURACHE, A. P.; MERTINS, A. Segmentation of retinal vessels with a hysteresis binary-classification paradigm. *Computerized Medical Imaging and Graphics*, Elsevier Ltd, v. 36, n. 4, p. 325–335, 2012. Disponível em: <<http://dx.doi.org/10.1016/j.compmedimag.2012.02.002>>.

- DASH, S.; SENAPATI, M. R. Enhancing detection of retinal blood vessels by combined approach of DWT, Tyler Coye and Gamma correction. *Biomedical Signal Processing and Control*, Elsevier Ltd, v. 57, p. 101740, 2020. ISSN 17468094. Disponível em: <<https://doi.org/10.1016/j.bspc.2019.101740>>.
- DELIBASIS, K. K.; KECHRINIOTIS, A. I.; TSONOS, C.; ASSIMAKIS, N. Automatic model-based tracing algorithm for vessel segmentation and diameter estimation. *Computer Methods and Programs in Biomedicine*, Elsevier Ireland Ltd, v. 100, n. 2, p. 108–122, 2010. ISSN 01692607. Disponível em: <<http://dx.doi.org/10.1016/j.cmpb.2010.03.004>>.
- DHARMAWAN, D. A.; LI, D.; NG, B. P.; RAHARDJA, S. A New Hybrid Algorithm for Retinal Vessels Segmentation on Fundus Images. *IEEE Access*, v. 7, p. 41885–41896, 2019. ISSN 21693536. Disponível em: <<https://doi.org/10.1109/ACCESS.2019.2906344>>.
- DORION, T. *Manual de Exame do Fundo de Olho*. 1. ed. Barueri - SP: Manole, 2002. ISBN 85-204-1203-3.
- DUCHI, J.; SINGER, Y. Efficient online and batch learning using forward backward splitting. *Journal of Machine Learning Research*, v. 10, p. 2899–2934, 2009. ISSN 15324435. Disponível em: <<https://www.jmlr.org/papers/volume10/duchi09a/duchi09a.pdf>>.
- DUMOULIN, V.; VISIN, F. A guide to convolution arithmetic for deep learning. p. 1–31, 2016. Disponível em: <<http://arxiv.org/abs/1603.07285>>.
- ELBALAOUI, A.; FAKIR, M.; TAIFI, K.; MERBOUHA, A. Automatic Detection of Blood Vessel in Retinal Images. *Proceedings - Computer Graphics, Imaging and Visualization: New Techniques and Trends, CGiV 2016*, IEEE, p. 324–332, 2016. Disponível em: <<https://ieeexplore.ieee.org/document/7467731>>.
- ESCORCIA-GUTIERREZ, J.; TORRENTS-BARRENA, J.; GAMARRA, M.; ROMERO-ARCA, P.; VALLS, A.; PUIG, D. Convexity Shape Constraints for Retinal Blood Vessel Segmentation and Foveal Avascular Zone Detection. *Computers in Biology and Medicine*, v. 127, n. July, p. 104049, 2020. ISSN 00104825. Disponível em: <<https://doi.org/10.1016/j.compbimed.2020.104049>>.
- FADZIL, M. A.; IZHAR, L. I.; VENKATACHALAM, P.; KARUNAKAR, T. Extraction and reconstruction of retinal vasculature. *Journal of medical engineering & technology*, Taylor & Francis, v. 31, n. 6, p. 435–442, 2007. Disponível em: <<https://doi.org/10.1080/03091900601111201>>.
- FAN, Z.; MO, J. J. Automated blood vessel segmentation based on de-noising auto-encoder and neural network. *Proceedings - International Conference on Machine Learning and Cybernetics*, IEEE, v. 2, p. 849–856, 2016. ISSN 21601348. Disponível em: <<https://doi.org/10.1109/ICMLC.2016.7872998>>.
- FARNELL, D. J.; HATFIELD, F.; KNOX, P.; REAKES, M.; SPENCER, S.; PARRY, D.; HARDING, S. P. Enhancement of blood vessels in digital fundus photographs via the application of multiscale line operators. *Journal of the Franklin institute*, Elsevier, v. 345, n. 7, p. 748–765, 2008. Disponível em: <<https://doi.org/10.1016/j.jfranklin.2008.04.009>>.

FATHI, A.; NAGHSH-NILCHI, A. R. Automatic wavelet-based retinal blood vessels segmentation and vessel diameter estimation. *Biomedical Signal Processing and Control*, Elsevier Ltd, v. 8, n. 1, p. 71–80, 2013. ISSN 17468094. Disponível em: <<https://doi.org/10.1016/j.bspc.2012.05.005>>.

FAUSETT, L. *Fundamentals of Neural Networks: Architectures, Algorithms, and Applications*. Upper Saddle River: Prentice-Hall, Inc., 1994. ISBN 0-13-334186-0.

FAZAL, M. I.; PATEL, M. E.; TYE, J.; GUPTA, Y. The past, present and future role of artificial intelligence in imaging. *European Journal of Radiology*, Elsevier, v. 105, n. May, p. 246–250, 2018. ISSN 18727727. Disponível em: <<https://doi.org/10.1016/j.ejrad.2018.06.020>>.

FENG, S.; ZHUO, Z.; PAN, D.; TIAN, Q. CcNet: A cross-connected convolutional network for segmenting retinal vessels using multi-scale features. *Neurocomputing*, Elsevier B.V., v. 392, p. 268–276, 2020. ISSN 18728286. Disponível em: <<https://doi.org/10.1016/j.neucom.2018.10.098>>.

FOROUZANFAR. Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks, 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015. *The Lancet*, v. 388, n. 10053, p. 1659–1724, 2016. ISSN 1474547X. Disponível em: <[https://doi.org/10.1016/S0140-6736\(16\)31679-8](https://doi.org/10.1016/S0140-6736(16)31679-8)>.

FRANGI, A. F.; NIESSEN, W. J.; VINCKEN, K. L.; VIERGEVER, M. A. Multiscale vessel enhancement filtering. In: SPRINGER. *International conference on medical image computing and computer-assisted intervention*. 1998. p. 130–137. Disponível em: <<https://doi.org/10.1007/BFb0056195>>.

FRAZ, M. M.; BARMAN, S. A.; REMAGNINO, P.; HOPPE, A.; BASIT, A.; UYYANONVARA, B.; RUDNICKA, A. R.; OWEN, C. G. An approach to localize the retinal blood vessels using bit planes and centerline detection. *Computer methods and programs in biomedicine*, Elsevier, v. 108, n. 2, p. 600–616, 2012. Disponível em: <<https://doi.org/10.1016/j.cmpb.2011.08.009>>.

FRAZ, M. M.; REMAGNINO, P.; HOPPE, A.; UYYANONVARA, B.; RUDNICKA, A. R.; OWEN, C. G.; BARMAN, S. A. An ensemble classification-based approach applied to retinal blood vessel segmentation. *IEEE Transactions on Biomedical Engineering*, IEEE, v. 59, n. 9, p. 2538–2548, 2012. ISSN 00189294. Disponível em: <<https://doi.org/10.1109/TBME.2012.2205687>>.

FRAZ, M. M.; REMAGNINO, P.; HOPPE, A.; UYYANONVARA, B.; RUDNICKA, A. R.; OWEN, C. G.; BARMAN, S. A. Blood vessel segmentation methodologies in retinal images—a survey. *Computer methods and programs in biomedicine*, v. 108, n. 1, p. 407–33, 2012. ISSN 1872-7565. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0169260712000843>>.

GANESAN, K.; NAIK, G.; ADAPA, D.; RAJ, A. N. J.; ALISETTI, S. N.; ZHUANG, Z. A supervised blood vessel segmentation technique for digital Fundus images using Zernike Moment based features. *PLoS ONE*, v. 15, n. 3, p. 1–23, 2020. ISSN 19326203. Disponível em: <<http://dx.doi.org/10.1371/journal.pone.0229831>>.

GEETHARAMANI, R.; BALASUBRAMANIAN, L. Retinal blood vessel segmentation employing image processing and data mining techniques for computerized retinal image analysis. *Biocybernetics and Biomedical Engineering*, Korea Institute of Oriental Medicine, v. 36, n. 1, p. 102–118, 2016. ISSN 0208-5216. Disponível em: <<http://dx.doi.org/10.1016/j.bbe.2015.06.004>>.

GEETHARAMANI, R.; BALASUBRAMANIAN, L. Retinal blood vessel segmentation employing image processing and data mining techniques for computerized retinal image analysis. *Biocybernetics and Biomedical Engineering*, Elsevier, v. 36, n. 1, p. 102–118, 2016. Disponível em: <<https://doi.org/10.1016/j.bbe.2015.06.004>>.

GHAEL, S. P.; SAYEED, A. M.; BARANIUK, R. G. Improved wavelet denoising via empirical Wiener filtering. *Proceedings of SPIE*, p. 389–399, 1997. ISSN 0277786X. Disponível em: <<https://doi.org/10.1117/12.292799>>.

GOEL, N.; YADAV, A.; SINGH, B. M. Medical image processing: A review. *2nd IEEE International Conference on Innovative Applications of Computational Intelligence on Power, Energy and Controls with their Impact on Humanity, CIPECH 2016*, IEEE, p. 57–62, 2017. Disponível em: <<https://ieeexplore.ieee.org/abstract/document/7918737>>.

GONZALEZ, R. C.; WOODS, R. C. *Processamento digital de imagens*. [S.l.]: Pearson Educação, 2009.

GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A.; BENGIO, Y. *Deep learning*. [S.l.]: MIT press Cambridge, 2016. v. 1.

GOPINATH, B.; CHIHA, J.; PLANT, A. J. H.; THIAGALINGAM, A.; BURLUTSKY, G.; KOVOOR, P.; LIEW, G.; MITCHELL, P. Associations between retinal microvascular structure and the severity and extent of coronary artery disease. *Atherosclerosis*, Elsevier Ltd, v. 236, n. 1, p. 25–30, 2014. ISSN 18791484. Disponível em: <<http://dx.doi.org/10.1016/j.atherosclerosis.2014.06.018>>.

GOUR, N.; KHANNA, P. Blood Vessel Segmentation using Hybrid Median Filtering and Morphological Transformation. *2017 13th International Conference on Signal-Image Technology and Internet-Based Systems (SITIS)*, IEEE, n. c, p. 151–157, dec 2017. Disponível em: <<http://ieeexplore.ieee.org/document/8334739>>.

GRAVES, A. Connectionist temporal classification. In: *Supervised Sequence Labelling with Recurrent Neural Networks*. Springer, 2012. p. 61–93. Disponível em: <<https://doi.org/10.1145/1143844.1143891>>.

GUO, S.; WANG, K.; KANG, H.; ZHANG, Y.; GAO, Y.; LI, T. BTS-DSN: Deeply supervised neural network with short connections for retinal vessel segmentation. *International Journal of Medical Informatics*, Elsevier, v. 126, n. March, p. 105–113, 2019. ISSN 18728243. Disponível em: <<https://doi.org/10.1016/j.ijmedinf.2019.03.015>>.

GUO, Y.; BUDAK, Ü.; ŞENGÜR, A. A novel retinal vessel detection approach based on multiple deep convolution neural networks. *Computer Methods and Programs in Biomedicine*, v. 167, p. 43–48, 2018. ISSN 18727565. Disponível em: <<https://doi.org/10.1016/j.cmpb.2018.10.021>>.

- HASHEMZADEH, M.; Adlpour Azar, B. Retinal blood vessel extraction employing effective image features and combination of supervised and unsupervised machine learning methods. *Artificial Intelligence in Medicine*, Elsevier, v. 95, n. March, p. 1–15, 2019. ISSN 18732860. Disponível em: <<https://doi.org/10.1016/j.artmed.2019.03.001>>.
- HASSAN, G.; EL-BENDARY, N.; HASSANIEN, A. E.; FAHMY, A.; M., S. A.; SNASEL, V. Retinal Blood Vessel Segmentation Approach Based on Mathematical Morphology. *Procedia Computer Science*, v. 65, p. 612–622, 2015. ISSN 1877-0509. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S1877050915028355>>.
- HASSANIEN, A. E.; EMARY, E.; ZAWBAA, H. M. Retinal blood vessel localization approach based on bee colony swarm optimization, fuzzy c-means and pattern search. *Journal of Visual Communication and Image Representation*, Elsevier Inc., v. 31, p. 186–196, aug 2015. ISSN 10959076. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S1047320315001182>>.
- HAYKIN, S. *Neural networks and learning machines*, 3/E. [S.l.]: Pearson Education India, 2010.
- HEBB, D. O. *The organisation of behaviour: a neuropsychological theory*. Science Editions New York, 1949. Disponível em: <<https://doi.org/10.4324/9781410612403>>.
- HOOVER, A. Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response. *IEEE Transactions on Medical Imaging*, v. 19, n. 3, p. 203–210, 2000. ISSN 02780062. Disponível em: <<https://doi.org/10.1109/42.845178>>.
- IKRAM, M. K.; ONG, Y. T.; CHEUNG, C. Y.; WONG, T. Y. Retinal vascular caliber measurements: Clinical significance, current knowledge and future perspectives. *Ophthalmologica*, v. 229, n. 3, p. 125–136, 2013. ISSN 00303755. Disponível em: <<https://doi.org/10.1159/000342158>>.
- IMRAN, A.; LI, J.; PEI, Y.; YANG, J. J.; WANG, Q. Comparative Analysis of Vessel Segmentation Techniques in Retinal Images. *IEEE Access*, IEEE, v. 7, p. 114862–114887, 2019. ISSN 21693536. Disponível em: <<https://doi.org/10.1109/ACCESS.2019.2935912>>.
- IOFFE, S.; SZEGEDY, C. Australian Literary Journalism and “Missing Voices”. *Journalism Practice*, v. 10, n. 6, p. 730–743, aug 2016. ISSN 1751-2786. Disponível em: <<http://www.tandfonline.com/doi/full/10.1080/17512786.2015.1058180>>.
- JEBASEELI, T. J.; DURAI, C. A. D.; PETER, J. D. Segmentation of retinal blood vessels from ophthalmologic Diabetic Retinopathy images. *Computers and Electrical Engineering*, Elsevier Ltd, v. 73, p. 245–258, 2019. ISSN 00457906. Disponível em: <<https://doi.org/10.1016/j.compeleceng.2018.11.024>>.
- JIA, D.; ZHUANG, X. Learning-based algorithms for vessel tracking: A review. *Computerized Medical Imaging and Graphics*, Elsevier Ltd, v. 89, n. January, p. 101840, 2021. ISSN 18790771. Disponível em: <<https://doi.org/10.1016/j.compmedimag.2020.101840>>.
- JIN, Q.; CHEN, Q.; MENG, Z.; WANG, B.; SU, R. Construction of Retinal Vessel Segmentation Models Based on Convolutional Neural Network. *Neural Processing Letters*, Springer US, v. 52, n. 2, p. 1005–1022, 2020. ISSN 1573773X. Disponível em: <<https://doi.org/10.1007/s11063-019-10011-1>>.

JONES, M. T. Deep learning architectures and the rise of artificial intelligence. *IBM DeveloperWorks*, September, v. 8, 2017.

JONES, S.; EDWARDS, R. T. Diabetic retinopathy screening: A systematic review of the economic evidence. *Diabetic Medicine*, v. 27, n. 3, p. 249–256, 2010. ISSN 14645491. Disponível em: <<https://doi.org/10.1111/j.1464-5491.2009.02870.x>>.

KARAALI, A.; DAHYOT, R.; SEXTON, D. J. DR-VNet: Retinal Vessel Segmentation via Dense Residual UNet. nov 2021. Disponível em: <<http://arxiv.org/abs/2111.04739>>.

KESKAR, N. S.; SOCHER, R. Improving generalization performance by switching from adam to sgd. *arXiv preprint arXiv:1712.07628*, 2017. Disponível em: <<https://doi.org/10.48550/arXiv.1712.07628>>.

KHAN, K. B.; KHALIQ, A. A.; JALIL, A.; IFTIKHAR, M. A.; ULLAH, N.; AZIZ, M. W.; ULLAH, K.; SHAHID, M. A review of retinal blood vessels extraction techniques: challenges, taxonomy, and future trends. Springer London, v. 22, n. 3, p. 767–802, 2019. ISSN 14337541. Disponível em: <<https://doi.org/10.1007/s10044-018-0754-8>>.

KHAN, K. B.; SIDDIQUE, M. S.; AHMAD, M.; MAZZARA, M. A Hybrid Unsupervised Approach for Retinal Vessel Segmentation. *BioMed Research International*, v. 2020, n. December, p. 1–20, dec 2020. ISSN 23146141. Disponível em: <<https://www.hindawi.com/journals/bmri/2020/8365783/>>.

KINGMA, D. P.; BA, J. L. Adam: A method for stochastic optimization. *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, p. 1–15, 2015. Disponível em: <<https://arxiv.org/pdf/1412.6980v9.pdf>>.

KOVÁCS, G.; HAJDU, A. A self-calibrating approach for the segmentation of retinal vessels by template matching and contour reconstruction. *Medical Image Analysis*, v. 29, p. 24–46, 2016. ISSN 13618423. Disponível em: <<https://doi.org/10.1016/j.media.2015.12.003>>.

KRIZHEVSKY, A.; SUTSKEVER, I.; Geoffrey E., H. ImageNet Classification with Deep Convolutional Neural Networks. *Advances in Neural Information Processing Systems 25 (NIPS2012)*, p. 1–9, 2012. ISSN 10495258. Disponível em: <<https://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf>>.

KUMAR, K.; RAVI, T.; RANI, B.; HABIBULLAH, K. Detection of retinal diseases by tracing vessel trees and accurately segmenting vessels. *Int J Eng, Citeseer*, v. 4, p. 182–7, 2012. Disponível em: <<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.677.9695&rep=rep1&type=pdf>>.

KUMAR, S.; CHOUDHARY, S.; GUPTA, R.; KUMAR, B. Performance Evaluation of Joint Filtering and Histogram Equalization Techniques for Retinal Fundus Image Enhancement. *2018 5th IEEE Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering, UPCON 2018*, IEEE, p. 18–22, 2018. Disponível em: <<https://ieeexplore.ieee.org/document/8597050>>.

LAM, B. S. Y.; YAN, H. A novel vessel segmentation algorithm for pathological retina images based on the divergence of vector fields. *IEEE Transactions on Medical Imaging*, v. 27, n. 2, p. 237–246, 2008. ISSN 02780062. Disponível em: <<https://doi.org/10.1109/TMI.2007.909827>>.

- LANI-LOUZADA, R.; do Val Ferreira Ramos, C.; CORDEIRO, R. M.; SADUN, A. A. Retinal changes in COVID-19 hospitalized cases. *PLoS ONE*, v. 15, n. 12 December, 2020. ISSN 19326203. Disponível em: <<http://dx.doi.org/10.1371/journal.pone.0243346>>.
- LECUN, Y.; BOTTOU, L.; BENGIO, Y.; HAFFNER, P. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, v. 86, n. 11, p. 2278–2323, 1998. ISSN 00189219. Disponível em: <<https://doi.org/10.1109/5.726791>>.
- LI, H.; ZHANG, J.; NIE, Q.; CHENG, L. A retinal vessel tracking method based on Bayesian theory. *Proceedings of the 2013 IEEE 8th Conference on Industrial Electronics and Applications, ICIEA 2013*, IEEE, n. 81271650, p. 232–235, 2013. Disponível em: <<https://ieeexplore.ieee.org/abstract/document/6566372>>.
- LI, Q.; FENG, B.; XIE, L.; LIANG, P.; ZHANG, H.; WANG, T. A cross-modality learning approach for vessel segmentation in retinal images. *IEEE Transactions on Medical Imaging*, IEEE, v. 35, n. 1, p. 109–118, 2016. ISSN 1558254X. Disponível em: <<https://doi.org/10.1109/TMI.2015.2457891>>.
- LI, Q.; YOU, J.; ZHANG, D. Vessel segmentation and width estimation in retinal images using multiscale production of matched filter responses. *Expert Systems with Applications*, Elsevier Ltd, v. 39, n. 9, p. 7600–7610, 2012. ISSN 09574174. Disponível em: <<http://dx.doi.org/10.1016/j.eswa.2011.12.046>>.
- LIDONG, H.; WEI, Z.; JUN, W.; ZEBIN, S. Combination of contrast limited adaptive histogram equalisation and discrete wavelet transform for image enhancement. *IET Image Processing*, IET, v. 9, n. 10, p. 908–915, 2015. Disponível em: <<https://doi.org/10.1049/iet-ipr.2015.0150>>.
- LIM, L. S.; MITCHELL, P.; SEDDON, J. M.; HOLZ, F. G.; WONG, T. Y. Age-related macular degeneration. *The Lancet*, Elsevier Ltd, v. 379, n. 9827, p. 1728–1738, 2012. ISSN 1474547X. Disponível em: <[http://dx.doi.org/10.1016/S0140-6736\(12\)60282-7](http://dx.doi.org/10.1016/S0140-6736(12)60282-7)>.
- LISKOWSKI, P.; KRAWIEC, K. Segmenting Retinal Blood Vessels with Deep Neural Networks. *IEEE Transactions on Medical Imaging*, v. 35, n. 11, p. 2369–2380, 2016. ISSN 1558254X. Disponível em: <<http://dx.doi.org/10.1109/TMI.2016.2546227>>.
- LIU, I.; SUN, Y. Recursive tracking of vascular networks in angiograms based on the detection—deletion scheme. *IEEE Transactions on Medical Imaging*, v. 12, n. 2, p. 334–341, 1993. ISSN 1558254X. Disponível em: <<https://ieeexplore.ieee.org/document/232264>>.
- LU, Y.; ZHOU, Y.; QIN, J. A Convolutional Encoder-Decoder Architecture for Retinal Blood Vessel Segmentation in Fundus Images. *2018 5th International Conference on Systems and Informatics, ICSAI 2018*, IEEE, n. Icsai, p. 1071–1075, 2019. Disponível em: <<https://doi.org/10.1109/ICSAI.2018.8599380>>.
- LUPAŞCU, C. A.; TEGOLO, D.; TRUCCO, E. FABC: Retinal vessel segmentation using AdaBoost. *IEEE Transactions on Information Technology in Biomedicine*, v. 14, n. 5, p. 1267–1274, 2010. ISSN 10897771. Disponível em: <<https://doi.org/10.1109/TITB.2010.2052282>>.

MA, J.; FAN, X.; YANG, S. X.; ZHANG, X.; ZHU, X. Contrast Limited Adaptive Histogram Equalization Based Fusion for Underwater Image Enhancement. *Preprints*, n. March, p. 1–27, 2017. Disponível em: <<http://dx.doi.org/10.20944/preprints201703.0086.v1>>.

MA, Z.; LI, H. Retinal vessel profiling based on four piecewise Gaussian model. *International Conference on Digital Signal Processing, DSP*, IEEE, v. 2015-September, n. 81271650, p. 1094–1097, 2015. Disponível em: <<https://ieeexplore.ieee.org/abstract/document/7252048>>.

MAHADEVAN, V.; NARASIMHA-IYER, H.; ROYSAM, B.; TANENBAUM, H. L. Robust model-based vasculature detection in noisy biomedical images. *IEEE Transactions on Information Technology in Biomedicine*, v. 8, n. 3, p. 360–376, 2004. ISSN 10897771. Disponível em: <<https://doi.org/10.1109/TITB.2004.834410>>.

MAJI, D.; SANTARA, A.; MITRA, P.; SHEET, D. Ensemble of Deep Convolutional Neural Networks for Learning to Detect Retinal Vessels in Fundus Images. mar 2016. Disponível em: <<https://arxiv.org/pdf/1603.04833.pdf>>.

MARÍN, D.; AQUINO, A.; GEGÚNDEZ-ARIAS, M. E.; BRAVO, J. M. A new supervised method for blood vessel segmentation in retinal images by using gray-level and moment invariants-based features. *IEEE Transactions on Medical Imaging*, v. 30, n. 1, p. 146–158, 2011. ISSN 02780062. Disponível em: <<https://ieeexplore.ieee.org/abstract/document/5545439>>.

MCCULLOCH, W. S.; PITTS, W. A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, Springer, v. 5, n. 4, p. 115–133, 1943. Disponível em: <<https://doi.org/10.1007/BF02478259>>.

MINSKY, M. *Heuristic aspects of the artificial intelligence problem*. Ed. Services Technical Information agency, 1956. Disponível em: <<https://apps.dtic.mil/sti/citations/AD0236885>>.

MOCCIA, S.; De Momi, E.; El Hadji, S.; MATTOS, L. S. Blood vessel segmentation algorithms — Review of methods, datasets and evaluation metrics. *Computer Methods and Programs in Biomedicine*, Elsevier B.V., v. 158, p. 71–91, 2018. ISSN 18727565. Disponível em: <<https://doi.org/10.1016/j.cmpb.2018.02.001>>.

MOOKIAH, M. R. K.; ACHARYA, U. R.; CHUA, C. K.; LIM, C. M.; NG, E. Y.; LAUDE, A. Computer-aided diagnosis of diabetic retinopathy: A review. *Computers in Biology and Medicine*, v. 43, n. 12, p. 2136–2155, 2013. ISSN 00104825. Disponível em: <<https://doi.org/10.1016/j.combiomed.2013.10.007>>.

MORANO, J.; HERVELLA, Á. S.; NOVO, J.; ROUCO, J. Simultaneous segmentation and classification of the retinal arteries and veins from color fundus images. *Artificial Intelligence in Medicine*, v. 118, p. 102116, aug 2021. ISSN 09333657. Disponível em: <<https://doi.org/10.1016/j.artmed.2021.102116>>.

NARAYANAN, B. N.; HARDIE, R. C.; De Silva, M. S.; KUETERMAN, N. K. Hybrid machine learning architecture for automated detection and grading of retinal images for diabetic retinopathy. *Journal of Medical Imaging*, v. 7, n. 03, p. 1, 2020. ISSN 2329-4302. Disponível em: <<https://doi.org/10.1117/1.JMI.7.3.034501>>.

NEKOVEI, R.; SUN, Y. Back-propagation network and its configuration for blood vessel detection in angiograms. *IEEE Transactions on Neural Networks*, IEEE, v. 6, n. 1, p. 64–72, 1995. Disponível em: <<https://doi.org/10.1109/72.363449>>.

NGO, L.; HAN, J.-H. Multi-level deep neural network for efficient segmentation of blood vessels in fundus images. *Electronics Letters*, v. 53, n. 16, p. 1096–1098, aug 2017. ISSN 0013-5194. Disponível em: <<https://ietresearch.onlinelibrary.wiley.com/doi/pdf/10.1049/el.2017.2066>>.

NGUYEN, D.; WIDROW, B. Improving the learning speed of 2-layer neural networks by choosing initial values of the adaptive weights. *IJCNN. International Joint Conference on Neural Networks*, p. 21–26, 1990. Disponível em: <<https://doi.org/10.1109/IJCNN.1990.137819>>.

NGUYEN, U. T.; BHUIYAN, A.; PARK, L. A.; RAMAMOHANARAO, K. An effective retinal blood vessel segmentation method using multi-scale line detection. *Pattern Recognition*, Elsevier, v. 46, n. 3, p. 703–715, 2013. ISSN 00313203. Disponível em: <<http://dx.doi.org/10.1016/j.patcog.2012.08.009>>.

NIEMEIJER, M.; STAAL, J.; GINNEKEN, B. van; LOOG, M.; ABRAMOFF, M. D. Comparative study of retinal vessel segmentation methods on a new publicly available database. In: . [s.n.], 2004. v. 5370, p. 648–656. Disponível em: <<http://dx.doi.org/10.1117/12.535349>>.

NOH, K. J.; PARK, S. J.; LEE, S. Scale-space approximated convolutional neural networks for retinal vessel segmentation. *Computer Methods and Programs in Biomedicine*, Elsevier B.V., v. 178, p. 237–246, 2019. ISSN 18727565. Disponível em: <<https://doi.org/10.1016/j.cmpb.2019.06.030>>.

ORLANDO, J. I.; BLASCHKO, M. Learning fully-connected crfs for blood vessel segmentation in retinal images. In: SPRINGER. *international conference on medical image computing and computer-assisted intervention*. 2014. p. 634–641. Disponível em: <https://doi.org/10.1007/978-3-319-10404-1_79>.

OVREIU, S.; CRISTESCU, I.; BALTA, F.; SULTANA, A.; OVREIU, E. Early Detection of Glaucoma Using Residual Networks. *2020 13th International Conference on Communications, COMM 2020 - Proceedings*, p. 161–164, 2020. Disponível em: <<https://doi.org/10.1109/COMM48946.2020.9141990>>.

PATTERSON, J.; GIBSON, A. *Deep learning: A practitioner's approach*. [S.l.]: "O'Reilly Media, Inc.", 2017.

PEDRINI, H.; SCHWARTZ, W. R. *Análise de imagens digitais: princípios, algoritmos e aplicações*. THOMSON PIONEIRA, 2007. ISBN 9788522105953. Disponível em: <<https://books.google.com.br/books?id=13KAPgAACAAJ>>.

PICCINELLI, M. Clinical cases. *Psicoterapia e Scienze Umane*, v. 8960, n. 3, p. 481–489, 2020. ISSN 03942864. Disponível em: <<https://www.wjgnet.com/2307-8960/full/v8/i18/3971.htm>>.

- RAMANI, R.; VANITHA, N.; VALARMATHY, S. The Pre-Processing Techniques for Breast Cancer Detection in Mammography Images. *International Journal of Image, Graphics and Signal Processing*, v. 5, n. 5, p. 47–54, 2013. ISSN 20749074. Disponível em: <<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.851.1555&rep=rep1&type=pdf>>.
- RAZZAK, M. I.; NAZ, S.; ZAIB, A. Deep learning for medical image processing: Overview, challenges and the future. *Lecture Notes in Computational Vision and Biomechanics*, v. 26, p. 323–350, 2018. ISSN 22129413. Disponível em: <https://doi.org/10.1007/978-3-319-65981-7_12>.
- RIBEIRO, A.; LOPES, A. P.; SILVA, C. A. Ensemble learning approaches for retinal vessel segmentation. *6th IEEE Portuguese Meeting on Bioengineering, ENBENG 2019 - Proceedings*, IEEE, p. 3–6, 2019. Disponível em: <<https://doi.org/10.1109/ENBENG.2019.8692566>>.
- RICCI, E.; PERFETTI, R. Retinal blood vessel segmentation using line operators and support vector classification. *IEEE Transactions on Medical Imaging*, v. 26, n. 10, p. 1357–1365, 2007. ISSN 02780062. Disponível em: <<https://doi.org/10.1109/TMI.2007.898551>>.
- ROBBINS, H.; MONRO, S. A stochastic approximation method. *The annals of mathematical statistics*, JSTOR, p. 400–407, 1951. Disponível em: <<http://www.jstor.org/stable/2236703>>.
- RONNEBERGER, O.; FISCHER, P.; BROX, T. U-net: Convolutional networks for biomedical image segmentation. In: SPRINGER. *International Conference on Medical image computing and computer-assisted intervention*. 2015. p. 234–241. Disponível em: <https://doi.org/10.1007/978-3-319-24574-4_28>.
- ROSENBLATT, F. Perceptron simulation experiments. *Proceedings of the IRE*, IEEE, v. 48, n. 3, p. 301–309, 1960. Disponível em: <<https://ieeexplore.ieee.org/abstract/document/4066017>>.
- ROSSANT, F.; BADELLINO, M.; CHAVILLON, A.; BLOCH, I.; PAQUES, M. A morphological approach for vessel segmentation in eye fundus images, with quantitative evaluation. *Journal of Medical Imaging and Health Informatics*, v. 1, p. 42–49, 2011. ISSN 21567018. Disponível em: <<https://doi.org/10.1166/jmihi.2011.1006>>.
- Saha Tchinda, B.; TCHIOTSOP, D.; NOUBOM, M.; LOUIS-DORR, V.; WOLF, D. Retinal blood vessels segmentation using classical edge detection filters and the neural network. *Informatics in Medicine Unlocked*, Elsevier Ltd, v. 23, p. 100521, 2021. ISSN 23529148. Disponível em: <<https://doi.org/10.1016/j.imu.2021.100521>>.
- SALAZAR-GONZALEZ, A.; KABA, D.; LI, Y.; LIU, X. Segmentation of the blood vessels and optic disk in retinal images. *IEEE Journal of Biomedical and Health Informatics*, IEEE, v. 18, n. 6, p. 1874–1886, 2014. ISSN 21682194.
- SALVI, M.; ACHARYA, U. R.; MOLINARI, F.; MEIBURGER, K. M. *The impact of pre- and post-image processing techniques on deep learning frameworks: A comprehensive review for digital pathology image analysis*. Elsevier Ltd, 2021. 104129 p. Disponível em: <<https://doi.org/10.1016/j.compbimed.2020.104129>>.

SANTOS, J. C. M. dos; CARRIJO, G. A.; de Fátima dos Santos Cardoso, C.; FERREIRA, J. C.; SOUSA, P. M.; PATROCÍNIO, A. C. Fundus image quality enhancement for blood vessel detection via a neural network using CLAHE and Wiener filter. *Research on Biomedical Engineering*, Research on Biomedical Engineering, v. 36, n. 2, p. 107–119, 2020. ISSN 24464740. Disponível em: <<https://doi.org/10.1007/s42600-020-00046-y>>.

SENGUPTA, S.; SINGH, A.; LEOPOLD, H. A.; GULATI, T.; LAKSHMINARAYANAN, V. Ophthalmic diagnosis using deep learning with fundus images – A critical review. *Artificial Intelligence in Medicine*, Elsevier, v. 102, n. November 2019, p. 101758, 2020. ISSN 18732860. Disponível em: <<https://doi.org/10.1016/j.artmed.2019.101758>>.

SINGH, N. P.; SRIVASTAVA, R. Retinal blood vessels segmentation by using Gumbel probability distribution function based matched filter. *Computer Methods and Programs in Biomedicine*, Elsevier Ireland Ltd, v. 129, p. 40–50, 2016. ISSN 0169-2607. Disponível em: <<http://dx.doi.org/10.1016/j.cmpb.2016.03.001>>.

SOLÍS-PÉREZ, J. E.; GÓMEZ-AGUILAR, J. F.; ESCOBAR-JIMÉNEZ, R. F.; REYES-REYES, J. Blood vessel detection based on fractional Hessian matrix with non-singular Mittag-Leffler Gaussian kernel. *Biomedical Signal Processing and Control*, Elsevier Ltd, v. 54, p. 101584, 2019. ISSN 17468108. Disponível em: <<https://doi.org/10.1016/j.bspc.2019.101584>>.

SON, J.; PARK, S. J.; JUNG, K. H. Towards Accurate Segmentation of Retinal Vessels and the Optic Disc in Fundoscopic Images with Generative Adversarial Networks. *Journal of Digital Imaging*, Journal of Digital Imaging, v. 32, n. 3, p. 499–512, 2019. ISSN 1618727X. Disponível em: <<https://doi.org/10.1007/s10278-018-0126-3>>.

SONALI; SAHU, S.; SINGH, A. K.; GHRERA, S. P.; ELHOSENY, M. An approach for de-noising and contrast enhancement of retinal fundus image using CLAHE. *Optics and Laser Technology*, Elsevier Ltd, v. 110, p. 87–98, 2019. ISSN 00303992. Disponível em: <<https://doi.org/10.1016/j.optlastec.2018.06.061>>.

SOOMRO, T. A.; AFIFI, A. J.; GAO, J.; HELLWICH, O.; KHAN, M. A.; PAUL, M.; ZHENG, L. Boosting sensitivity of a retinal vessel segmentation algorithm with convolutional neural network. In: IEEE. *2017 International Conference on Digital Image Computing: Techniques and Applications (DICTA)*. 2017. p. 1–8. Disponível em: <<https://doi.org/10.1109/DICTA.2017.8227413>>.

SOOMRO, T. A.; AFIFI, A. J.; GAO, J.; HELLWICH, O.; ZHENG, L.; PAUL, M. Strided fully convolutional neural network for boosting the sensitivity of retinal blood vessels segmentation. *Expert Systems with Applications*, Elsevier Ltd, v. 134, p. 36–52, 2019. ISSN 09574174. Disponível em: <<https://doi.org/10.1016/j.eswa.2019.05.029>>.

SRENG, S.; MANEERAT, N.; HAMAMOTO, K. Automated microaneurysms detection in fundus images using image segmentation. *2nd Joint International Conference on Digital Arts, Media and Technology 2017: Digital Economy for Sustainable Growth, ICDAMT 2017*, IEEE, p. 19–23, 2017. Disponível em: <<https://doi.org/10.1109/ICDAMT.2017.7904926>>.

STAAL, J.; ABRÀMOFF, M. D.; NIEMEIJER, M.; VIERGEVER, M. A.; Van Ginneken, B. Ridge-based vessel segmentation in color images of the retina. *IEEE Transactions on Medical Imaging*, v. 23, n. 4, p. 501–509, 2004. ISSN 02780062. Disponível em: <<https://doi.org/10.1109/TMI.2004.825627>>.

- SUERO, A.; MARIN, D.; GEGUNDEZ-ARIAS, M. E.; BRAVO, J. M. Locating the Optic Disc in Retinal Images Using Morphological Techniques. *Iwbbio*, p. 18–20, 2013. Disponível em: <<https://pdfs.semanticscholar.org/8227/421bdd762782672a3cdbca394bd59415d3dd.pdf>>.
- TAGHANAKI, S. A.; ABHISHEK, K.; COHEN, J. P.; COHEN-ADAD, J.; HAMARNEH, G. Deep Semantic Segmentation of Natural and Medical Images: A Review. oct 2019. Disponível em: <<http://arxiv.org/abs/1910.07655>>.
- TAMIM, N.; ELSHRKAWEY, M.; AZIM, G. A.; NASSAR, H. Retinal blood vessel segmentation using hybrid features and multi-layer perceptron neural networks. *Symmetry*, v. 12, n. 6, 2020. ISSN 20738994. Disponível em: <<https://doi.org/10.3390/sym12060894>>.
- TIAN, C.; FANG, T.; FAN, Y.; WU, W. Multi-path convolutional neural network in fundus segmentation of blood vessels. *Biocybernetics and Biomedical Engineering*, Nalecz Institute of Biocybernetics and Biomedical Engineering of the Polish Academy of Sciences, v. 40, n. 2, p. 583–595, 2020. ISSN 02085216. Disponível em: <<https://doi.org/10.1016/j.bbe.2020.01.011>>.
- UPADHYAY, K.; AGRAWAL, M.; VASHIST, P. Unsupervised multiscale retinal blood vessel segmentation using fundus images. *IET Image Processing*, v. 14, n. 11, p. 2616–2625, 2020. ISSN 17519659. Disponível em: <<https://ietresearch.onlinelibrary.wiley.com/doi/pdfdirect/10.1049/iet-ipr.2019.0969>>.
- UPADHYAY, K.; AGRAWAL, M.; VASHIST, P. Wavelet based Fine-to-Coarse Retinal Blood Vessel Extraction using U-net Model. *SPCOM 2020 - International Conference on Signal Processing and Communications*, p. 19–23, 2020. Disponível em: <<https://ieeexplore.ieee.org/abstract/document/9179575>>.
- UPADHYAY, K.; AGRAWAL, M.; VASHIST, P. U-net based multi-level texture suppression for vessel segmentation in low contrast regions. *European Signal Processing Conference*, v. 2021-Janua, p. 1304–1308, 2021. ISSN 22195491. Disponível em: <<https://doi.org/10.23919/Eusipco47968.2020.9287387>>.
- VEGA, R.; SANCHEZ-ANTE, G.; FALCON-MORALES, L. E.; SOSSA, H.; GUEVARA, E. Retinal vessel extraction using Lattice Neural Networks with dendritic processing. *Computers in Biology and Medicine*, v. 58, p. 20–30, 2015. ISSN 18790534. Disponível em: <<https://doi.org/10.1016/j.compbiomed.2014.12.016>>.
- VERMEER, K. A.; VOS, F. M.; LEMIJ, H. G.; VOSSEPOEL, A. M. A model based method for retinal blood vessel detection. *Computers in Biology and Medicine*, v. 34, n. 3, p. 209–219, 2004. Disponível em: <[https://doi.org/10.1016/S0010-4825\(03\)00055-6](https://doi.org/10.1016/S0010-4825(03)00055-6)>.
- VILLALOBOS-CASTALDI, F. M.; FELIPE-RIVERÓN, E. M.; SÁNCHEZ-FERNÁNDEZ, L. P. A fast, efficient and automated method to extract vessels from fundus images. *Journal of Visualization*, v. 13, n. 3, p. 263–270, 2010. ISSN 13438875. Disponível em: <<https://doi.org/10.1007/s12650-010-0037-y>>.
- VLACHOS, M.; DERMATAS, E. Multi-scale retinal vessel segmentation using line tracking. *Computerized Medical Imaging and Graphics*, v. 34, n. 3, p. 213 – 227, 2010. ISSN 0895-6111. Disponível em: <<http://dx.doi.org/10.1016/j.compmedimag.2009.09.006>>.

VOSTATEK, P.; CLARIDGE, E.; UUSITALO, H.; HAUTA-KASARI, M.; FÄLT, P.; LENSU, L. Performance comparison of publicly available retinal blood vessel segmentation methods. *Computerized medical imaging and graphics : the official journal of the Computerized Medical Imaging Society*, Elsevier Ltd, v. 55, p. 2–12, jan 2017. ISSN 1879-0771. Disponível em: <<https://doi.org/10.1016/j.compmedimag.2016.07.005>>.

WANG, J.; LI, Y. J.; YANG, K. F. Retinal fundus image enhancement with image decomposition and visual adaptation. *Computers in Biology and Medicine*, Elsevier Ltd, v. 128, n. August 2020, p. 104116, 2021. ISSN 18790534. Disponível em: <<https://doi.org/10.1016/j.combiomed.2020.104116>>.

WANG, S.; YIN, Y.; CAO, G.; WEI, B.; ZHENG, Y.; YANG, G. Hierarchical retinal blood vessel segmentation based on feature and ensemble learning. *Neurocomputing*, v. 149, Part B, p. 708 – 717, 2015. ISSN 0925-2312. Disponível em: <<http://dx.doi.org/10.1016/j.neucom.2014.07.059>>.

WANG, W.; WANG, W.; HU, Z. Segmenting retinal vessels with revised top-bottom-hat transformation and flattening of minimum circumscribed ellipse. *Medical and Biological Engineering and Computing*, Medical & Biological Engineering & Computing, v. 57, n. 7, p. 1481–1496, 2019. ISSN 17410444. Disponível em: <<http://dx.doi.org/10.1007/s11517-019-01967-2>>.

WANG, Y.-B.; ZHU, C.-Z.; YAN, Q.-F.; LIU, L.-Q. A novel vessel segmentation in fundus images based on svm. In: *2016 International Conference on Information System and Artificial Intelligence (ISAI)*. [S.l.: s.n.], 2016. p. 390–394.

WHO. Global Report on Diabetes. *Isbn*, v. 978, p. 6–86, 2016. Disponível em: <<https://apps.who.int/iris/handle/10665/204871>>.

World Health Organization. *World report on vision*. [s.n.], 2019. v. 214. 1–160 p. ISBN 9789241516570. Disponível em: <<https://apps.who.int/iris/bitstream/handle/10665/328717/9789240017184-ara.pdf>>.

WU, C. H.; AGAM, G.; STANCHEV, P. A general framework for vessel segmentation in retinal images. *Proceedings of the 2007 IEEE International Symposium on Computational Intelligence in Robotics and Automation, CIRA 2007*, p. 37–42, 2007. Disponível em: <<https://ieeexplore.ieee.org/abstract/document/4269924>>.

XIE, C. H.; LIU, Y. J.; CHANG, J. Y. Medical image segmentation using rough set and local polynomial regression. *Multimedia Tools and Applications*, v. 74, n. 6, p. 1885–1914, 2015. ISSN 15737721. Disponível em: <<https://doi.org/10.1007/s11042-013-1723-2>>.

XU, X.; NIEMEIJER, M.; SONG, Q.; SONKA, M.; GARVIN, M. K.; REINHARDT, J. M.; ABRÀMOFF, M. D. Vessel boundary delineation on fundus images using graph-based approach. *IEEE Transactions on Medical Imaging*, IEEE, v. 30, n. 6, p. 1184–1191, 2011. ISSN 02780062. Disponível em: <<https://ieeexplore.ieee.org/abstract/document/5680652>>.

YAN, Z.; YANG, X.; CHENG, K. T. A Three-Stage Deep Learning Model for Accurate Retinal Vessel Segmentation. *IEEE Journal of Biomedical and Health Informatics*, IEEE, v. 23, n. 4, p. 1427–1436, 2019. ISSN 21682208. Disponível em: <<https://doi.org/10.1109/JBHI.2018.2872813>>.

- YEDIDYA, T.; HARTLEY, R. Tracking of blood vessels in retinal images using kalman filter. *Proceedings - Digital Image Computing: Techniques and Applications, DICTA 2008*, p. 52–58, 2008. Disponível em: <<https://ieeexplore.ieee.org/abstract/document/4699999>>.
- ZAKI, W. M. D. W.; ZULKIFLEY, M. A.; HUSSAIN, A.; HALIM, W. H. W. A.; MUSTAFA, N. B. A.; TING, L. S. Diabetic retinopathy assessment: Towards an automated system. *Biomedical Signal Processing and Control*, Elsevier Ltd, v. 24, p. 72–82, 2016. ISSN 17468108. Disponível em: <<http://dx.doi.org/10.1016/j.bspc.2015.09.011>>.
- ZANA, F.; KLEIN, J. C. Segmentation of vessel-like patterns using mathematical morphology and curvature evaluation. *IEEE Transactions on Image Processing*, v. 10, n. 7, p. 1010–1019, 2001. ISSN 10577149. Disponível em: <<https://ieeexplore.ieee.org/abstract/document/931095>>.
- ZHANG, Y.; CHUNG, A. C. Deep supervision with additional labels for retinal vessel segmentation task. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, v. 11071 LNCS, p. 83–91, 2018. ISSN 16113349. Disponível em: <https://doi.org/10.1007/978-3-030-00934-2_10>.
- ZHANG, Z.; SRIVASTAVA, R.; LIU, H.; CHEN, X.; DUAN, L.; Kee Wong, D. W.; KWOH, C. K.; WONG, T. Y.; LIU, J. A survey on computer aided diagnosis for ocular diseases. *BMC medical informatics and decision making*, v. 14, n. 1, p. 80, 2014. ISSN 1472-6947. Disponível em: <<http://www.biomedcentral.com/1472-6947/14/80>>.
- ZHANG, Z.; XING, F.; SU, H.; SHI, X.; YANG, L. Recent Advances in the Applications of Convolutional Neural Networks to Medical Image Contour Detection. 2017. Disponível em: <<http://arxiv.org/abs/1708.07281>>.
- ZHOU, L.; RZESZOTARSKI, M. S.; SINGERMAN, L. J.; CHOKREFF, J. M. The Detection and Quantification of Retinopathy Using Digital Angiograms. v. 13, n. 4, 1994. Disponível em: <<https://doi.org/10.1109/42.363106>>.
- ZHOU, M.; JIN, K.; WANG, S.; YE, J.; QIAN, D. Color Retinal Image Enhancement Based on Luminosity and Contrast Adjustment. *IEEE Transactions on Biomedical Engineering*, IEEE, v. 65, n. 3, p. 521–527, 2018. ISSN 15582531. Disponível em: <<https://ieeexplore.ieee.org/abstract/document/7918600>>.