



UNIVERSIDADE FEDERAL DO MARANHÃO
Programa de Pós-Graduação em Engenharia Elétrica

Fredson Costa Rodrigues

***APRENDIZADO PROFUNDO PARA A
DETECÇÃO DE PUPILAS DILATADAS OU
CONTRAÍDAS***

**São Luís - MA
2022**

Fredson Costa Rodrigues

**APRENDIZADO PROFUNDO PARA A DETECÇÃO DE
PUPILAS DILATADAS OU CONTRAÍDAS**

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Elétrica na Universidade Federal do Maranhão, como requisito para obtenção do título de Mestre em Ciências da Computação.

Orientador: Prof. Dr. Anselmo Cardoso Paiva
Coorientador: Prof. Dr. Aristófanés Corrêa Silva

São Luís - MA

2022

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

Costa Rodrigues, Fredson.

APREDIZADO PROFUNDO PARA A DETECÇÃO DE PUPILAS
DILATADAS OU CONTRAÍDAS / Fredson Costa Rodrigues. - 2022.
89 f.

Coorientador(a): Aristófaes Côrrea Silva.

Orientador(a): Anselmo Cardoso Paiva.

Dissertação (Mestrado) - Programa de Pós-graduação em
Engenharia Elétrica/ccet, Universidade Federal do
Maranhão, São Luís, 2022.

1. At-Unet. 2. Construção Pupilar. 3. Dilatação
pupilar. 4. Rede Neural de Multitarefa. 5. Segmentação.
I. Cardoso Paiva, Anselmo. II. Côrrea Silva,
Aristófaes. III. Título.

Fredson Costa Rodrigues

**APRENDIZADO PROFUNDO PARA A DETECÇÃO DE
PUPILAS DILATADAS OU CONTRAÍDAS**

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Elétrica na Universidade Federal do Maranhão, como requisito para obtenção do título de Mestre em Ciências da Computação.

Trabalho Aprovado. São Luís - MA, data mês :

Prof. Dr. Anselmo Cardoso Paiva
Orientador

Prof. Dr. Aristófanês Corrêa Silva
Coorientador

**Prof. Dr. João Dallyson Sousa de
Almeida**
Examinador da Banca

Prof. Dr. Andre Castelo Branco
Examinador da Banca

São Luís - MA
2022

*Aos meus pais, irmãos e os verdadeiros amigos, na qual em todos os momentos me
incentivaram a continuar nesta jornada*

Agradecimentos

A Deus, por me dar forças em todos os momentos bons ou ruins.

À minha mãe Ivete Costa e ao meu pai Valber Macedo, por todo amor, afeto e conselhos, e pelo incentivo dado em todos os momentos da minha vida.

Aos meus irmãos Haniel Costa e Josiane Macedo, por todo o auxílio em todos os momentos de dificuldades que enfrentei.

Ao meu orientador, professor Anselmo Paiva, obrigado por todo apoio, aprendizado e cuidado durante os mais de dois anos de orientação. Obrigado pelo voto de confiança, paciência, e incentivo que foram imensuráveis durante todo esse tempo.

Da mesma maneira expresso meus sentimentos de gratidão ao professor João Dallyson, pelas acertadas sugestões durante a realização deste trabalho, e pela paciência e disponibilidade em ajudar os estudantes do núcleo de computação aplicada.

Aos meus queridos amigos Carolina Lima e Alan Moura, que não mediram esforços para me auxiliar a superar os desafios enfrentados, sou grato pelo companheirismo durante a pós-graduação.

Agradeço também a todos os colegas do Núcleo de Computação Aplicada que se disponibilizaram em todos os momentos de dúvidas. Expresso gratidão aos professores Aristófanis, Geraldo e aos demais professores do programa, por serem uma fonte de inspiração por sua dedicação ao trabalho.

Aos meus gestores do trabalho, Léa Cristina, Rafael Adam e Lourdes Mayre, agradeço pela compreensão, encorajamento e flexibilidade que foram essências para a conclusão do mestrado e a execução deste trabalho.

Por fim agradeço a todos que contribuíram diretamente ou indiretamente ao longo do percurso do mestrado e durante a na construção deste trabalho.

“O conhecimento nos faz responsáveis.”

Che Guevara

Resumo

A pupilometria é a medição do diâmetro pupilar, usada em alguns procedimentos médicos a fim de avaliar as reações de dilatação e constrição da pupila. Essas reações podem ser involuntárias causadas pelo Sistema Nervoso Central (SNC) ou provocadas, por meio dos reflexos de luz. A íris e pupila fazem parte do globo ocular responsável por formar a visão humana. A íris é a circunferência maior que contém texturas pigmentadas, permitindo a formação da cor dos olhos. A pupila é a circunferência menor, e se encontra na região interna da íris, e é caracterizada por permitir a entrada da luz para formar a visão. Essa região apresenta duas respostas (estados) comuns mediante estímulos externos (Reflexos de luz) e internos (Sistema Nervoso Central). O estado de dilatação ocorre quando há o aumento do tamanho da pupila, enquanto que no estado de constrição seu tamanho diminui. Entender essas reações pupilares, tem se tornado comum entre os pesquisadores da neurociência e psicologia cognitiva, pois permite a identificação de distúrbios neurológicos por meio dessas reações pupilares. Dessa forma a pupilometria tem se tornado uma estratégia comumente usada em processos clínicos pré-operatórios, e na identificação de distúrbios neurológicos em indivíduos. Neste contexto propõe-se um método computacional, para detectar e medir o tamanho da pupila, baseado em técnicas de processamento de imagens digitais e aprendizado de máquina, a fim de auxiliar psicólogos cognitivos e pesquisadores da neurociência a entender e identificar doenças neurológicas por meio das reações pupilares. O método propõe uma arquitetura de rede neural de multitarefas com a inclusão de mecanismos de atenção, denominado de At-Unet, para segmentar a região da íris e pupila com a intenção de obter o diâmetro da pupila, e calcular o fator de dilatação que define o estado da pupila em dilatada ou contraída. Esse método conseguiu 97,17% de coeficiente de Dice, a partir do experimento de validação cruzada, dessa forma o tipo de estado pupilar estimado apresenta um erro médio do fator de dilatação de 0.0167.

Palavras-chave: At-Unet, Constrição Pupilar, Dilatação pupilar, Rede Neural de Multitarefa, Segmentação.

Abstract

Pupillometry is the measurement of pupillary diameter used in some medical procedures to evaluate reactions of pupillary dilatation and constriction. Those reactions can be involuntary when it is caused by the Central Nervous System (CNS), or provoked by the light reflections. The iris and the pupil compose the ocular globe that is responsible for forming human vision. The iris is the biggest circumference with pigmented textures, which allows the formation of the color of the eyes. The pupil is the smallest circumference and it is in the internal region of the iris; the pupil is characterized by which allows the entry of light to form the vision. This region presents two common responses (states) through externals (reflections of light) and internals stimulus (Central Nervous System). The dilatation state occurs when there is an extension of the pupil size, in a constriction state. On the other hand, its size reduces. Understanding these pupillary reactions has been coming common among the researchers of neuroscience and cognitive psychology since it allows the identification of neurologic disorders. Therefore, pupillometry has been becoming a popular strategy in clinic pre-operative processes, and in the identification of neurologic disorders in some individuals. Considering this context, this work aims to propose a computational method able to detect and measure pupil size, based on processing digital image techniques and machine learning to assist cognitive psychologists and neuroscience researchers to understand and identify neurologic diseases through pupillary reactions. The method of planning a multitasking neural network architecture with the inclusion of attention mechanisms, called At-Unet, to segment the iris and pupil region with the intention of obtaining the pupil diameter, and calculate the dilation factor that defines the state of the pupil in dilated or contracted. This method achieved 97.17% of Dice coefficient, from the cross experiment, so the type of pupil state estimated has an average error of the dilation factor of 0.0167.

Keywords: At-Unet, Constriction pupil, Dilatation pupil, Neural Network of Multitasking, Segmentation.

Lista de ilustrações

Figura 1 – (a) Pupilômetro e (b) Imagem obtida pelo pupilômetro.	17
Figura 2 – Anatomia do Olho	28
Figura 3 – Anatomia da íris	29
Figura 4 – Músculos da Íris responsáveis pela Dilatação e Constrição	30
Figura 5 – Constrição e Dilatação Pupilar	32
Figura 6 – Vias de Contrição e Dilatação Pupilar	33
Figura 7 – Representação de uma Imagem Digital: (a) Imagem 2D e (b) Imagem 3D	34
Figura 8 – Modelo de cor RGB	36
Figura 9 – Ilustração da operação de Dilatação	38
Figura 10 – Ilustração da Operação de Erosão	39
Figura 11 – Exemplo do processo de segmentação da pupila usando limiar em imagens infravermelhas dos olhos.	40
Figura 12 – Estrutura de um Neurônio Artificial	41
Figura 13 – Rede Neural <i>Multilayer Perceptron</i>	43
Figura 14 – Exemplo típico de uma CNN	46
Figura 15 – Ilustração da Camada de Convolução	47
Figura 16 – Ilustração do Pooling após a camada de convolução	49
Figura 17 – Arquitetura VGG16	50
Figura 18 – Arquitetura da Rede Neural U-Net	52
Figura 19 – Mecanismo de atenção	53
Figura 20 – Ilustração do <i>Smallest Circle</i>	54
Figura 21 – Metodologia Proposta	57
Figura 22 – UTIRIS Dataset	58
Figura 23 – Arquitetura de Rede Neural Proposta(At-UNet)	60
Figura 24 – Codificador da rede At-UNet	61
Figura 25 – Mecanismo de Atenção da At-UNet	62
Figura 26 – Exemplo do processo de segmentação com a rede neural At-UNet . . .	62
Figura 27 – Imagens segmentadas com áreas que não fazem parte da região da íris.	63
Figura 28 – Ilustração das subetapas do pós-processamento.	63
Figura 29 – Regiões Segmentadas da íris e pupila aplicando o algoritmo <i>Smallest Circle</i> .(a) íris segmentada. (b) aplicação do algoritmo na imagem segmentada. (c) Imagem resultante da íris após a aplicação do algoritmo. (d) Pupila segmentada. (e) Aplicação do algoritmo na imagem segmentada. (f) Resultado Final da pupila segmentada após a aplicação do algoritmo .	65
Figura 30 – Intervalos do fator de dilatação no conjunto de dados	66
Figura 31 – Imagens do conjunto de dados com o fator de dilatação	66

Figura 32 – Processo de Segmentação da Íris e Pupila. (a) Imagem de entrada para segmentação. (b) Resultado da segmentação da íris. (c) Resultado da segmentação da pupila. (d) Pós-processamento do resultado de segmentação da íris. (e) Pós-processamento do resultado da segmentação da pupila. (f) Imagem segmentada da íris pós-processada. (g) Imagem segmentada da pupila pós-processada. (h) Íris e pupila detectada . . . 69

Figura 33 – Imagens de teste classificadas incorretamente o tipo de reação pupilar de acordo com o fator de dilatação calculado. (a) Imagem nº 6 de teste com pupila dilatada. (b) Imagem nº 7 de teste com pupila dilatada. (c) Imagem nº 10 com pupila contraída. (d) Imagem nº 12 de teste com pupila dilatada. 72

Lista de tabelas

Tabela 1 – Principais Trabalhos Relacionados	27
Tabela 2 – Arquitetura VGG16	51
Tabela 3 – Características da base UTIRIS	59
Tabela 4 – Dilatação Pupilar	65
Tabela 5 – Resultados da segmentação usando a Rede Neural de Multitarefa At-Unet	70
Tabela 6 – Resultados do cálculo do fator de dilatação	71
Tabela 7 – Resultados da Validação Cruzada da rede At-Unet de multitarefa . . .	72
Tabela 8 – Comparação de diferentes abordagens para a segmentação da região periocular.	74
Tabela 9 – Artigos publicados que possuem relação com o método proposto. . . .	77

Lista de abreviaturas e siglas

AWB	<i>Automatic White Balance</i>
BSOM	<i>Batch-SOM</i>
CNN	<i>Convolutional Neural network</i>
ENS	Sistema Nervoso Entérico
FCNN	Fully Convolutional Neural Network
IoU	<i>Intersection Over Union</i>
LT	Transformação Logaritma
LMS	<i>Least Mean Square</i>
LP	<i>Linear-Programming-Like</i>
LSFE	<i>Squares Fitting of Ellipse</i>
MLP	<i>Multilayer Perceptron</i>
NTM	<i>Normalized Template Matching</i>
PDIF	<i>Pupil Detection After Isolation and Fitting</i>
PLT	Transformação da Lei de Potência
PRL	<i>Pupil Light Response</i>
PNR	<i>Pupil Near Response</i>
PPR	<i>Psychosensory Pupil Response</i>
PP	<i>Pulling e Pushing</i>
RNN	<i>Redes Neurais Recorrentes</i>
RPN	<i>Region Proposal Network</i>
SNA	Sistema Nervoso Autônomo
SNC	Sistema Nervoso Central
SNS	Sistema Nervoso Simpático
SNP	Sistema Nervoso Parassimpático
TEA	Transtorno do Espectro Autismo

Sumário

1	INTRODUÇÃO	16
1.1	Objetivo Geral	19
1.2	Objetivos Específicos	19
1.3	Contribuições	19
1.4	Organização do Trabalho	19
2	TRABALHOS RELACIONADOS	21
3	FUNDAMENTAÇÃO TEÓRICA	28
3.1	Olho Humano	28
3.2	Pupilometria	31
3.3	A dilatação e constrição pupilar	31
3.4	Processamento de Imagens	34
3.4.1	Modelos de Representação da Cores	35
3.4.2	Operações Morfológicas	36
3.4.2.1	Operação Morfológica de Dilatação	37
3.4.2.2	Operação Morfológica de Erosão	37
3.4.3	Segmentação	39
3.5	Redes Neurais	41
3.6	Redes <i>Multilayer Perceptron</i>	42
3.7	Aprendizado Profundo	44
3.8	Redes Neurais Convolucionais	45
3.8.1	Camada de Convolução	47
3.8.2	Camada de <i>Pooling</i>	48
3.8.3	VGG16	48
3.8.4	U-Net	50
3.8.5	Mecanismo de Atenção Redes Neurais	53
3.9	Smallest Circle	54
3.10	Métricas de Avaliação	55
3.11	Considerações Finais	56
4	METODOLOGIA	57
4.1	Aquisição da Base de Imagens	58
4.2	Segmentação da Íris e Pupila	59
4.3	Pós-Processamento da Segmentação	63
4.4	Fator de dilatação Pupilar	64

5	RESULTADOS	68
5.1	Segmentação da Iris e Pupila	68
5.2	Cálculo do Fator de Dilatação	70
5.3	Validação Cruzada	72
5.4	Comparação com outros Trabalhos	73
5.5	Hardware e Softwares Utilizados	74
6	CONCLUSÃO	75
6.1	Trabalhos Futuros	76
6.2	Produções Científicas	76
	REFERÊNCIAS	78

1 Introdução

A pupila é uma abertura localizada no centro da íris dos olhos. Ela é responsável por controlar a quantidade de luz que na retina, possibilitando a formação da imagem por meio dos estímulos nervoso que são enviados ao cérebro. A pupila é preta devido não haver iluminação no fundo do olho, ou seja, isso significa que a pupila não é uma superfície opaca. Outra característica das pupilas é que geralmente elas parecem perfeitamente redondas e seus tamanhos são idênticos, além disso o diâmetro normal de uma pupila humana, varia entre cerca de 2 a 8 mm (MATHÔT, 2018). Mudanças no diâmetro da pupila respondem a três tipos distintos de estímulos: elas se contraem em resposta à luz (*pupil light response* ou *PRL*), se contraem em resposta à fixação próxima de um objeto (*pupil near response* ou *PNR*) e se dilatam em resposta a aumentos na excitação e esforço mental, desencadeados por um estímulo externo ou espontaneamente (*psychosensory pupil response* ou *PPR*) (MATHÔT, 2018). Dessa forma as pupilas podem variar entre 2 a 4 mm quando estão contraídas e 4 a 8 mm quando estão dilatadas.

A pupila é considerada uma janela para o sistema nervoso autônomo (*SNA*) como índice do funcionamento do sistema nervoso simpático e parassimpático. O sistema nervoso simpático é responsável pelas reações do corpo em perigo, esforço intenso e estresse físico. O sistema nervoso parassimpático está relacionada a resposta do organismo em situações de calma, ansiedade, repouso e digestão. Sendo assim a pupila fornece informações sobre o sistema nervoso de um indivíduo, tornando-se um objeto de estudo em várias pesquisas (de Vries et al., 2021; HALL; CHILCOTT, 2018; LYNCH, 2018).

A dilatação e a constrição pupilar consistem na alteração do tamanho da pupila diante de estímulos externos ou neurológicos. Essa mobilidade da pupila pode ser espontânea ou forçada. O aumento pupilar é induzido por um estímulo de aumento do diâmetro da pupila também conhecido como resposta pupilar. Alguns estudos revelam que a expansão e a constrição pupilar, aumentam com a demanda de tarefas (GOLDINGER; PAPESH, 2012). O tamanho da pupila é afetado pelo nível de iluminação na retina, o estado de acomodação do olho, por condições sensoriais e emocionais, e pelo processamento de informações cognitivos-afetivos (NOWAK et al., 2014).

As respostas pupilares representadas pela dilatação ou constrição da pupila podem indicar sintomas de doenças neurológicas ou visuais. O tamanho da pupila ao estímulo de luz pode ser um indicador sensível para avaliação da atividade do sistema nervoso autônomo, sendo assim, a partir desses valores, é possível identificar os pacientes com lesões cerebrais, além da possibilidade de ajudar em procedimentos pré-operatórios (HALL; CHILCOTT, 2018) e de inferir o estágio da doença de Parkinson (SOOYEOUN HONG

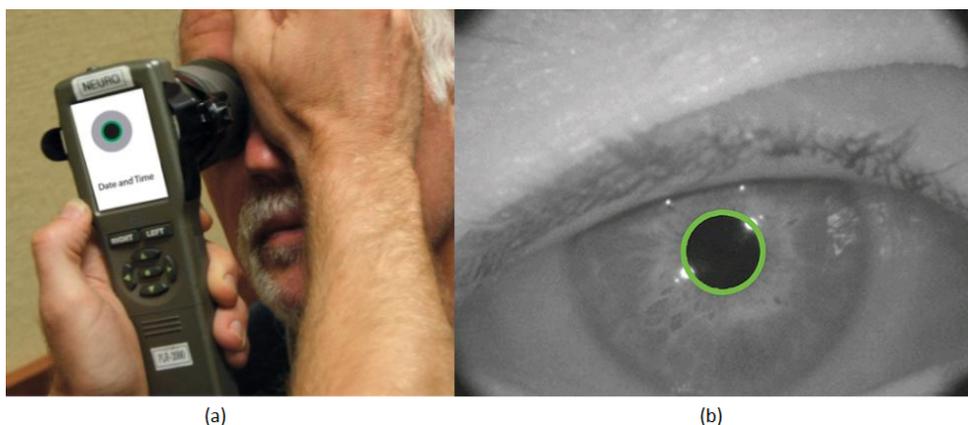
JEONG-HO, 2021).

Há algumas aplicações médicas para a avaliação das respostas pupilares, elas incluem a detecção de trauma cerebral, avaliação de álcool, intoxicação por medicamentos e estudos de doenças oftalmológicas como retinopatia diabética e síndrome de *Horner* (PIAGGIO et al., 2021). As reações pupilares podem variar entre os indivíduos que apresentam problemas neurais ou não, portanto este viés possibilita inclusive a avaliação da resposta pupilar, entre pessoas com o transtorno do espectro autista (*TEA*) (de Vries et al., 2021), e a doença de Parkinson (GIZA et al., 2010), a fim de inferir as principais características dessas doenças.

Para a análise dessas respostas pupilares é necessário realizar a pupilometria. A pupilometria é uma técnica conveniente e viável que é rotineiramente usada para medir o tamanho da pupila em várias condições fisiológicas e patológicas (DU et al., 2005). Trata-se de uma técnica de monitorização objetiva não invasiva, utilizada principalmente para avaliar o sistema nervoso autônomo (BREMNER, 2004).

Vários estudos mostram que existem diferentes maneiras de aplicar a pupilometria por meio de pupilômetros automatizados projetados e estabelecidos para a análise da pupila (CHAGLASIAN; AKBAR; PROBST, 2006). Os primeiros pupilômetros eram lentos e de baixa precisão devido à baixa taxa de quadros lidos, mas atualmente os pupilômetros automatizados evoluíram e apresentam alta precisão e confiabilidade (PADMAVATHI; KUPPUSAMY, 2019). A Figura 1 apresenta um pupilômetro portátil infravermelho, desenvolvido para avaliar reações pupilares e medir o tamanho da pupila. Esse equipamento tem uma precisão de 0,05 mm.

Figura 1 – (a) Pupilômetro e (b) Imagem obtida pelo pupilômetro.



Fonte: (LARSON; BEHREND, 2015)

As avaliações desses equipamentos consistem em duas medidas importantes quanto ao tempo da resposta pupilar. A primeira é a estimativa de quantos segundos a pupila leva para dilatar ou contrair, e a segunda é a estimativa do tamanho total da pupila durante a

aplicação do pupilômetro. A possibilidade de medir a dilatação e a constrição em vários intervalos de tempo para avaliar o esforço cognitivo ao ouvir, falar e até sentir, é um indicador viável e importante da pupilometria (WINN et al., 2018a).

Portanto existem vários equipamentos desenvolvidos para aplicar o a pupilometria a fim de medir o tamanho da pupila e o tempo de reação pupilar. Em Nowak et al. (2014) é proposto um sistema em que os olhos, direito e esquerdo, são iluminados de forma independente para gerar imagens alternadas e capturadas por uma câmera. O sistema é composto por quatro módulos principais, e aplica um algoritmo que aproxima a forma da pupila a uma elipse de melhor ajuste. Além disso ele disponibiliza um painel controlador para definir os parâmetros de precisão para a medição da pupila.

A maioria dos trabalhos são baseados em segmentar a pupila e a íris para estimar o diâmetro da pupila, e realizar a avaliação pupilar. Esse é um processo importante em boa parte dos trabalhos desenvolvidos para realizar a pupilometria. Dessa maneira este trabalho apresenta e discute um método computacional baseado em Redes Neurais Convolucionais (*CNN'S*) para realizar a segmentação e detecção da íris e pupila, a fim de obter e calcular o diâmetro de uma pupila dilatada ou contraída no intuito de estimar o tipo de reação pupilar.

Uma possibilidade de análise das pupilas, baseadas na pupilometria é a classificação do estado pupilar dos olhos. As reações devem acontecer no olho direito e esquerdo quando estimuladas, logo as pupilas devem apresentar as mesmas simetrias (mesmo diâmetro de dilatação ou constrição), durante esse processo de análise da reação pupilar. Os estados pupilares são definidos em Midríase, Miose e Anisocóricas, através deles é possível distinguir e categorizar possíveis doenças e questões psicológicas e neurológicas (JU YANG HEE KYUNG, 2017).

A Midríase é o estado de dilatação e a miose é o estado de constrição pupilar, esses dois estados são considerados normais pois apresentam a mesma simetria entre a pupila do olho direito e esquerdo. O estado anisocórico ocorre quando não existe simetria entre as pupilas, ou seja, elas apresentam diâmetros diferentes, uma deve estar dilatada enquanto a outra está contraída (JU YANG HEE KYUNG, 2017).

Neste aspecto, é necessário buscar recursos e tecnologias para realizar a pupilometria de maneira precisa, com a finalidade de estudar e compreender esses estados e reações pupilares, em várias áreas da medicina. Sendo assim, é necessário melhorar e desenvolver equipamentos, técnicas e métodos, para realizar a pupilometria com alta precisão e de baixo custo.

1.1 Objetivo Geral

Este trabalho objetiva apresentar um método computacional para segmentar a íris e pupila da região dos olhos em imagens, a fim de avaliar e definir o seu estado pupilar em dilatada ou contraída. Dessa forma o método basea-se em uma arquitetura de rede neural de multitarefas integrando mecanismo de atenção durante o processo de segmentação.

1.2 Objetivos Específicos

Dessa forma os principais objetivos específicos para atingir o objetivo geral são:

- Desenvolver um método de segmentação da íris e pupila.
- Investigar a utilização de mecanismo de atenção para melhorar o processo de segmentação.
- Calcular e definir o diâmetro total da pupila.
- Estimar o tipo de reação pupilar em cada imagem de teste.
- Analisar os resultados obtidos utilizando bases de dados públicas como UTIRIS.
- Comparar o desempenho do método proposto com outros trabalhos relacionados.
- Analisar as vantagens e limitações do método proposto.

1.3 Contribuições

A principal contribuição deste trabalho é apresentar um método capaz de estimar o tipo de reação pupilar baseada em um processo de segmentação da íris e pupila. Dessa forma outras contribuições podem ser citadas:

- A realização da segmentação simultânea da íris e pupila através da rede neural convolucional de multitarefas.
- A utilização do fator de dilatação para estimar e classificar o tipo de reação pupilar.

1.4 Organização do Trabalho

Esta dissertação está organizada em 6 capítulos descritos a seguir.

O **Capítulo 2** descreve os trabalhos relacionados, expondo suas metodologias e resultados relacionados à detecção da íris e pupila dilatadas em imagens.

O **Capítulo 3** apresenta a fundamentação teórica, descrevendo conceitos importantes, além de apresentar o embasamento teórico das principais técnicas e algoritmos utilizados neste trabalho.

O **Capítulo 4** apresenta uma descrição detalhada sobre o desenvolvimento da metodologia proposta para a segmentação da íris e pupila dilatada ou contraída.

No **Capítulo 5**, os resultados são apresentados e discutidos em relação às técnicas aplicadas.

Finalmente, o **Capítulo 6** apresenta as considerações finais sobre o trabalho realizado e os artigos científicos publicados, relacionados ao tema.

2 Trabalhos Relacionados

Neste capítulo são apresentados trabalhos e técnicas relacionadas ao processo de detecção da íris e pupila para a extração de medidas e avaliação da resposta pupilar. Logo os trabalhos citados, usam métodos específicos que variam em técnicas de processamento de imagens e aprendizado profundo. Os pesquisadores da visão computacional buscam melhorias na precisão da detecção de pupilas dilatadas ou contraídas, com a produção de tecnologias de baixo custo, para auxiliar os especialistas na pupilometria.

Bastos et al. (2012) discutem uma abordagem, baseada em aprendizado profundo, desenvolvida para detectar as pupilas em imagens, segmentando essa região de interesse. Eles expõem suas visões sobre o papel da segmentação das pupilas em tarefas de rastreamento ocular e compara com outras possíveis propostas na visão computacional. Para atingir o objetivo da tarefa, é discutido o processo de segmentação da pupila combinando o algoritmo *Pulling e Pushing* (PP) modificado, com a rede neural *Batch-SOM (BSOM)* introduzida por Venkatesh, Raja e Ramya (2006). Durante a detecção nota-se que a pupila na imagem sofre oclusões causadas por reflexos especulares. Então a ideia principal é obter uma estimativa inicial do contorno pupilar usando o algoritmo PP para segmentar a região de interesse, e em seguida aplicar rede neural para ajustar com precisão a estimativa para os limites exatos da pupila. Além disso algumas técnicas de pré-processamento de imagens são usadas para fazer os ajustes necessários nas imagens da região dos olhos que estão principalmente oclusas por cílios/pálpebras e reflexos especulares. Para a avaliação são selecionadas duas bases públicas de imagens infravermelho. As bases *MMU1* e *Casia-IrisV3-Interval* tem 2.655 e 450 imagens, respectivamente, e são comparadas durante os experimentos com outros trabalhos. A proposta atinge 97,95% de precisão na base de dados Casia Iris e 95,33% na MMU1.

O estudo de Bozomitu et al. (2016) demonstraram um algoritmo capaz de realizar a detecção da pupila usando *Least Squares* ou método dos mínimos quadrados. O seu objetivo foi obter as coordenadas da pupila para estimar a sua posição em imagens de infravermelho. É necessário realizar uma sequência de etapas antes da aplicação deste algoritmo, descrito como *squares fitting of ellipse* (LSFE). O autor descreve que na imagem fornecida ao método, é preciso filtrar a imagem em tempo real para que em seguida aconteça a binarização, de acordo com o limiar estabelecido. Após isso, é realizada a detecção da borda pupilar com a finalidade de aplicar o algoritmo LSFE, para estimar as coordenadas da pupila. O LSFE propõe adaptar o melhor ajuste possível de uma elipse em volta da região segmentada. A precisão desta técnica foi testada primeiro em um grupo de imagens em condições específicas, obtendo o erro relativo das coordenadas detectadas.

Wangwiwattana, Ding e Larson (2017) apresentam duas abordagens diferentes para aplicar a pupilometria e obter a estimativa do diâmetro da pupila quando dilatada ou contraída. Uma abordagem é baseada em *deep learning* no qual é denominada de PupilNet e a outra é fundamentada no uso do algoritmo *Starburst*. A pupilometria usando as duas técnicas apresenta um bom desempenho. O Mazziotti et al. (2021) expõe uma ferramenta chamada de *MEYE* baseada em Redes Neurais Convolucionais para detectar e medir o tamanho da pupila.

Nugroho, Nasrun e Setianingsih (2017) discutem uma aplicação para a detecção de pupilas dilatadas ou contraídas. Os autores discutiram a detecção de mentiras a partir das respostas pupilares localizadas, usando técnicas de processamento de imagens. Para conseguir atingir o objetivo desta tarefa, fizeram uso da transformada de *Hough* (DUDA; HART, 1972) e Lógica *Fuzzy*. As imagens são adquiridas através de câmeras de infravermelho, com a finalidade de encontrar a região dos olhos e assim conseguir estimar a posição das pupilas por meio da transformada de *Hough*. Em seguida, é calculado o diâmetro pupilar no intuito de distinguir quando as pupilas estão dilatadas ou não. Dessa forma, a proposta consegue estimar a mentira nas informações adquiridas através das reações pupilares detectadas. O detector aplica um limiar para indicar a região da pupila e assim localiza-lá por meio da transformada de *Hough*. A metodologia foi testada por completo, adquirindo imagens em tempo real de 30 indivíduos, para avaliação do desempenho do detector. Segundo os autores, a metodologia apresenta 84% de precisão.

Kiruthiga e Arumuganathan (2017) propõem detectar a pupila no rastreamento ocular, apresentando um meio de suavizar as imagens da íris possibilitando, segmentar a região da pupila de maneira precisa. Para que a tarefa aconteça são realizadas duas etapas principais durante o desenvolvimento do trabalho. Primeiro é aplicado *fractional differential mask* para suavizar a região da íris a fim de remover os ruídos e reflexos especulares nas imagens da região dos olhos, no qual dificulta o processo de detecção da pupila e aumenta o nível de contraste entre a íris e a pupila. E em seguida é segmentada a região da pupila usando a transformada de *Waletet*. Durante o desenvolvimento do trabalho observou-se que o processo de segmentação é mais efetivo quando é realizada a primeira etapa discutida, além disso todas as imagens são convertidas em apenas um canal de cor. As bases de dados UTIRIS e CASIA.v1 são usadas para avaliar o desempenho da abordagem, e incluem 158 e 756 imagens, respectivamente, apenas da região dos olhos. Durante a etapa de segmentação, a transformada de *Wavelet* é usada, principalmente, para reduzir a complexidade computacional durante essa etapa e preservar as informações da imagem. O desempenho deste trabalho é comparado com outras técnicas e apresenta bons resultados na execução da tarefa.

A pesquisa de Ramlee, Ramli e Noh (2017) mostram um método de segmentação da pupila usando transformação logaritmo (LT) e a transformação da lei de potência (PLT).

Já Gowroju, Aarti e Kumar (2021) apresentam um método baseado em *deep learning* para realizar a segmentação da pupila. Estes propuseram uma modificação na arquitetura da rede U-Net, a fim de realizar uma segmentação mais precisa da pupila.

A detecção robusta de pupilas em imagens coloridas usando aprendizado profundo é discutida em Kitazumi e Nakazawa (2018). A proposta não se limita apenas em aplicar a detecção de pupilas em imagens de infravermelho, adquiridas por equipamentos que podem ser de auto custo. Expõe a possibilidade de detectar essa região, em tempo real mediante os reflexos especulares, ou condições específicas que a região dos olhos pode apresentar, e dessa maneira prejudicar a execução da tarefa. O método, primeiro, detecta a região do olhos para segmentar a pupila por meio de uma CNN, e em seguida estima as coordenadas do centro da pupila. A U-Net é usada para cumprir a tarefa de segmentação, sendo assim, são selecionadas três bases de dados para avaliar o processo de segmentação e detecção da pupila. Duas destas bases são de domínio público e apresentam características diferentes. A primeira base a ser citada é a UBIRISv2, que apresenta apenas imagens da região dos olhos ocluídas por pálpebras, cílios e reflexos pupilares, o que dificulta a identificação da pupila. Já a segunda base destacada, a GI4E *database*, contém imagens apenas da região da face, adquiridas por uma webcam. A última base de imagens mencionada é elaborada pelos próprios autores, por meio de uma câmera a 10 centímetros da região dos olhos. Todas essas bases foram rotuladas manualmente e utilizadas nos experimentos de treinamento, validação e teste. Para avaliar a metodologia o autor obteve precisão média de 81% durante os seus experimentos.

Lin et al. (2019) destacam que o formato da pupila não é regular, podendo ser circular ou elíptica, dessa forma demonstram um modelo de detecção de pupilas dilatadas baseado no formato elíptico. Sendo assim, investigam a viabilidade de estimar parâmetros elípticos para a localização e detecção de pupilas dilatadas baseadas em aprendizado profundo. A *Anchor Ellipse Regression* assim denominado no estudo de Lin et al. (2019), faz uso da *Region Proposal Network*(RPN), que prevê os eixos elípticos maiores, menores e ângulos de rotação, para indicar a região da pupila. Portanto, são geradas elipses âncoras, por meio dos parâmetros de localização estimados na rede neural, e em seguida são segmentadas a região da pupila detectada. Essa abordagem realiza testes para avaliar o desempenho do método desenvolvido, usando imagens de infravermelho adquiridas através de uma Filmadora HDR-PJ790 da Sony, com a visão noturna integrada. A métrica de avaliação dos resultados, da segmentação, é baseada no coeficiente de Dice médio. Os resultados são promissores e apresentam um Dice médio de 95%.

O trabalho de Yiu et al. (2019) apresentam a DeepVog, uma metodologia computacional para a detecção de pupilas dilatadas ou contraídas em imagens adquiridas pelo *German Center for Vertigo and Balance Disorders*. Essa proposta tem por objetivo segmentar as pupilas de maneira precisa usando aprendizado profundo. Sendo assim, o autor apresenta

uma rede neural totalmente convolucional (FCNN) para a segmentação das pupilas a fim de estimar o contorno pupilar. Essa proposta é inspirada na arquitetura de rede neural U-Net e V-Net (MILLETARI; NAVAB; AHMADI, 2016). Além disso a proposta permite realizar a localização do centro da pupila e a detecção, estimativa do olhar. O trabalho foca na segmentação precisa da pupila para obter os parâmetros necessários com o intuito de finalizar a tarefa. São selecionadas três bases de dados privadas, disponibilizadas pelo *German Center*, no intuito de validar o desempenho do método. As bases contêm imagens de ressonância magnética (MRI scanner) capturadas durante um experimento neurocientífico. Duas dessas bases são usadas para treinamento e validação da rede neural, a outra é usada apenas para teste. Os resultados foram avaliados através do coeficiente de Dice médio, a fim verificar o desempenho da tarefa. Dessa forma obtiveram uma média de 96% de coeficiente de Dice na segmentação da pupila.

Petronela et al. (2019) propuseram um algoritmo de detecção de pupilas em imagens oculares ruidosas e prejudicadas pelos reflexos pupilares. Os autores não se limitam apenas à dilatação ou contração pupilar, mas apresentam a possibilidade de detectar as pupilas diante de condições adversas, independente da reação pupilar. O algoritmo desenvolvido converte a imagem de entrada em apenas um canal de cor, usa operadores morfológicos, aplica alguns filtros de suavização da imagem, obtém o histograma da imagem filtrada, e calcula o limiar para binarizar a imagem de entrada. A principal característica deste algoritmo é o uso do operador *Laplaciano* para encontrar as arestas conectadas na imagem que indicam a região da pupila, baseada na circularidade na forma binarizada. A pesquisa aponta o uso de duas bases de dados de imagens infravermelhas para validar e testar o algoritmo. A primeira base foi adquirida pelos autores do trabalho usando um dispositivo *head-mounted* desenvolvido pela equipe de pesquisa. Já a segunda base consiste em dados públicos provenientes do *dataset Casia-Iris-Lamp* (LI; SUN; TAN, 2013). Leva-se em consideração o centro da pupila detectada para estimar o desempenho do algoritmo através do cálculo da distância *Euclidiana* entre o centro da pupila rotulada manualmente e a detectada pelo algoritmo. Dos seus resultados, ele apresenta uma média do desvio padrão de 1,87 nas duas bases usadas durante a pesquisa. Esse resultado é comparado com outros trabalhos, destacando seu bom desempenho ao longo dos seus experimentos.

Kumar e Teotia (2019) desenvolveram uma técnica para rastreamento ocular, reconhecimento da íris e pupilometria, a fim de detectar e estimar o tamanho de uma pupila. A técnica é baseada na segmentação dessa região por meio de *circular active disc*, ou seja, são dois círculos concêntricos envolvendo a região da pupila e da íris. O método proposto consiste em dois passos, o primeiro é estimar o *active disc* de maneira automática, usando *normalized template matching*(NTM), e o segundo passo consiste em segmentar a região pupilar de acordo com o *circular active disc* estimado. Esse modelo se adapta conforme os limites pupilares encontrados, sendo assim, o raio deve ser sempre ajustado de acordo com as bordas pupilares. As bases de dados públicas *IIT Delhi iris* e *CasiaV3*,

contém imagens apenas das regiões dos olhos que, são rotuladas manualmente, e usadas para avaliar o desempenho do método proposto. A principal métrica de avaliação é o índice de coeficiente de Dice, onde apresenta uma média de 94,63% ao testar a base de dados *Casia-V3* e 93,08% na *IIT Delhi*.

Somuncuoglu et al. (2020) expõem a detecção das reações pupilares, apresentando uma investigação ao detectar as pupilas dilatadas em resposta ao estresse. Dessa maneira, apresenta um método baseado em aprendizado profundo para estimar os níveis de estresse por meio da detecção de pupilas. O autor consegue estimar os níveis de estresse obtendo os diâmetros pupilares, capturadas por meio de imagens de vídeos com iluminação infravermelho. Desta forma para estimar esses níveis a rede neural usada passa por algumas alterações nos parâmetros e hiperparâmetros a fim de realizar a tarefa com precisão. O método atingiu 90% de sucesso durante os experimentos realizados.

A proposta de Han et al. (2020) busca auxiliar o rastreamento ocular e a pupilometria, apresentando um processo de detecção de pupilas, fundamentada em aprendizado profundo para estimar o centro dessa região. É discutido o processo de segmentação para localizar a região de interesse, usando a arquitetura de rede neural U-Net. A arquitetura foi modificada alterando alguns parâmetros dos caminhos de codificação e decodificação, ocorrendo também alterações em hiperparâmetros da arquitetura. Durante os experimentos os autores consideraram a diminuição da complexidade da rede neural U-Net, levando em conta que as alterações realizadas desempenham bons resultados durante o processo de segmentação. Durante o desenvolvimento do trabalho os autores selecionaram três bases de dados diferentes para treinar, validar e testar o processo de segmentação da metodologia, além de rotular manualmente todas as imagens para a avaliação. Os autores construíram uma base de dados e utilizaram as bases de dados públicas *ExCuSe* e *ElSe*. Para verificar o desempenho do método o autor propõe avaliar a taxa de detecção da pupila, que é calculada a partir da distância *euclidiana* entre a pupila detectada e a rótulo real desta região, sendo assim essa métrica também é comparada com outros trabalhos com a mesma tarefa definida. Esse trabalho apresenta uma taxa de detecção da pupila de 87,3%.

O artigo de Manuri, Sanna e Petrucci (2020) mostra um algoritmo para rastreamento ocular que pode ser aplicado na pupilometria. Portanto, no trabalho é discutido o uso de operadores morfológicos durante o processo de detecção. Segundo os autores, esta é uma etapa essencial para se obter bons resultados durante a execução da tarefa. O papel desses operadores é remover os possíveis reflexos identificados na pupila, além de reduzir os ruídos que podem existir na região de interesse. O algoritmo *Pupil Detection After Isolation and Fitting* (PDIF), proposto no trabalho, foi projetado após a análise dos problemas e das limitações dos algoritmos concorrentes baseados em operadores morfológicas. Fizem uso da base de dados pública CASIA para avaliar o desempenho do PDIF. A taxa de detecção da pupila, o desvio padrão do centro da pupila rotulada, a porcentagem de

erro na estimativa da pupila e o tempo de execução do algoritmo, em milissegundos, são considerados os principais meios de avaliar o desempenho do PDIF. A técnica apresenta o tempo médio de execução de 7,10 ms menor que outros métodos comparados, e apresenta uma média de porcentagem de erro de 6,37%.

O trabalho de Temel et al. (2020) propõe RAPDNET, um método computacional para detectar diferentes reações pupilares entre os olhos quando expostos por estímulos luminosos, indicando algum distúrbio no nervo óptico que pode levar a perda visual. O trabalho propõe detectar as pupilas usando aprendizado profundo baseado em *transfer learning*, incluindo a arquitetura de rede pré-treinada Alex Net (XIAO; YAN; DENG, 2017). Deste modo, o autor localiza e detecta a região da pupila, selecionando 3.620 imagens do conjunto de dados ELSE para realizar os experimentos. O método apresentou uma média de 95,57% de precisão na tarefa proposta.

Dessa forma Bai, Wang e Wang (2021) propõem um método de segmentação baseado em *fuzzy clustering*, desenvolvido para estimar o tamanho da pupila. Para tanto, processa a imagem original do olho para remover algumas características e destacar apenas a área da pupila. Além disso, são destacadas as principais etapas de processamento na imagem do olho para obter apenas a região de interesse final, e assim estimar o tamanho da pupila dilatada ou contraída.

Kothari et al. (2021) detectam a pupila utilizando uma abordagem baseada em aprendizado profundo. Para tanto, propuseram o framework EllSeg para a segmentação pupilar, objetivando segmentar estruturas elípticas inteiras a partir de imagens da região dos olhos. Segundo os autores, a arquitetura desenvolvida consegue lidar com as oclusões causadas pelas pálpebras. Portanto a proposta é avaliar o framework junto a outras arquiteturas de rede neurais para segmentação. O DenseElNet (JÉGOU et al., 2017), RITnet (CHAUDHARY et al., 2019) e DeepVOG (YIU et al., 2019) são as arquiteturas de redes neurais que foram selecionadas para testar o framework de segmentação, a partir dos seguintes conjuntos de dados: OpenEDS, NVGaze, ExCuSe e ElSe. Os dados dos datasets foram divididos em treinamento, teste e validação. A métrica para avaliar o desempenho da segmentação usando o framework é o cálculo do IoU (Intersection Over Union) e a taxa de detecção pupilar. A DenseElNet apresenta a maior média de resultados comparado a outras arquiteturas, apresentando a pontuação de 95,4%.

Diante do exposto, é possível observar que há diversas propostas de trabalhos relacionados a pupilometria e ao rastreamento ocular, a maioria destaca que a partir da detecção da pupila é possível compreender as reações pupilares de dilatação e constrição. É importante salientar que a maior parte dos métodos apresentados, realizam a segmentação da pupila a fim de calcular alguns dados para estimar o estado da reação pupilar.

Esses trabalhos relacionados indicam que uma boa alternativa para realizar a detecção da pupila de maneira precisa é através do uso de aprendizado profundo, que

tende a trazer grandes benefícios ao realizar o processo de detecção. Na Tabela 1 estão relacionados os trabalhos apresentados.

O presente trabalho visa apresentar um método computacional capaz de detectar e estimar o tipo de reação pupilar por meio da segmentação da íris e pupila, de maneira que venha auxiliar nesse processo de análise das reações pupilares. Dessa forma o método propõe segmentar essas regiões de interesse de maneira simultânea por meio de uma arquitetura de rede neural convolucional modificada possibilitando estimar o estado pupilar em contraída ou dilatada.

Tabela 1 – Principais Trabalhos Relacionados

Autor	Dataset	Abordagem
Bastos et al. (2012)	<i>Multi Media University Iris</i> Casia-IrisV3	Algoritmo de <i>Pulling, Pushing</i> e a rede neural <i>Batch-SOM</i>
Bozomitu et al. (2016)	Privada	<i>Squares Fitting of Ellipse</i>
Wangwiwattana, Ding e Larson (2017)	Casia-Iris Privada	<i>Fuzzy Clustering</i>
Nugroho, Nasrun e Setianingsih (2017)	Privada	Transformada de Hough e Logica Fuzzy
Kiruthiga e Arumuganathan (2017)	Casia-Iris <i>Multi Media University Iris</i>	PLT e LT
Ramlee, Ramli e Noh (2017)	OpenEDS NVGaze ExCuSe ElSe	Transformada <i>wavelet</i> e Derivada Fracionária
Kitazumi e Nakazawa (2018)	UBIRIS.v2	U-Net
Lin et al. (2019)	Privada	<i>Anchor Ellipse Regression</i>
Yiu et al. (2019)	Privada	<i>DeepVog</i>
Petronela et al. (2019)	Casia-Iris-Lamp Privada	Operador Laplaciano
Kumar e Teotia (2019)	IIT Delhi iris Casia-Iris	<i>Circular Active disc</i>
Manuri, Sanna e Petrucci (2020)	ElSe/ExCuSe c	<i>Pupil Detection After Isolation and Fitting</i>
Han et al. (2020)	ExCuSe ElSe	U-Net
Temel et al. (2020)	ElSe	RAPDnet
Mazziotti et al. (2021)	Privada	MEYE
Bai, Wang e Wang (2021)	Casia-Iris-Lamp	<i>Fuzzy Clustering</i>
(GOWROJU; AARTI; KUMAR, 2021)	<i>Multi Media University Iris</i>	U-Net
Kothari et al. (2021)	OpenEDS NVGaze ExCuSe ElSe	EllSeg

3 Fundamentação Teórica

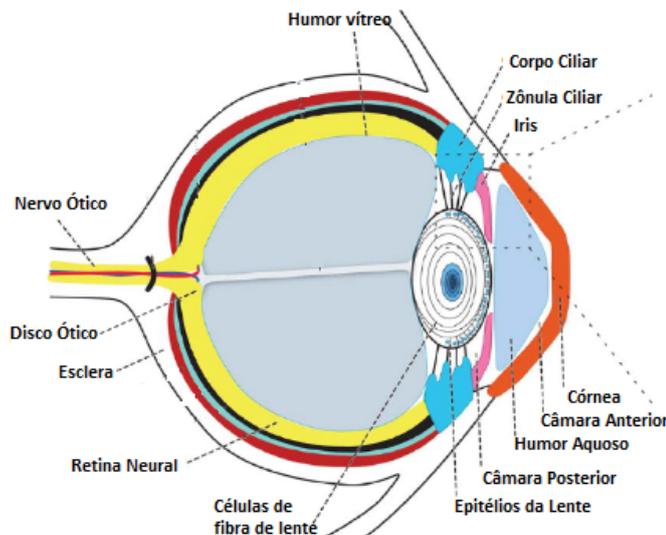
Este capítulo descreve os fundamentos teóricos necessários para o entendimento do problema abordado neste trabalho. Ademais, as técnicas que embasam a metodologia desenvolvida também estão aqui apresentadas. Inicialmente, explica-se o funcionamento do olho humano e a anatomia da íris e pupila. Com essas informações é possível entender como acontece a dilatação e a contração pupilar, aspecto fundamental para a compreensão do domínio do problema. Em seguida, explica-se os conceitos e técnicas de Processamento Digital de Imagens, Redes Neurais e principais métricas de avaliação.

3.1 Olho Humano

Conhecer a anatomia do olho humano é uma etapa importante, tendo em vista que a necessidade de entendimento de suas partes internas facilita o desenvolvimento de técnicas computacionais que visem a detecção de pupilas.

O olho humano é um sistema sensorial complexo (Figura 2) que desempenha funções de formação e processamento de imagens em conexão com o cérebro. Ele capta a energia luminosa refletida pelos objetos que penetram no sistema óptico do globo ocular, sofrem um desvio e produzem uma imagem invertida do objeto na retina. A energia luminosa é transformada em impulsos nervosos transmitidos até o cérebro, que são convertidos na imagem do objeto observado (PASSOS; ANDRADE-NETO; LEMAIRE, 2008).

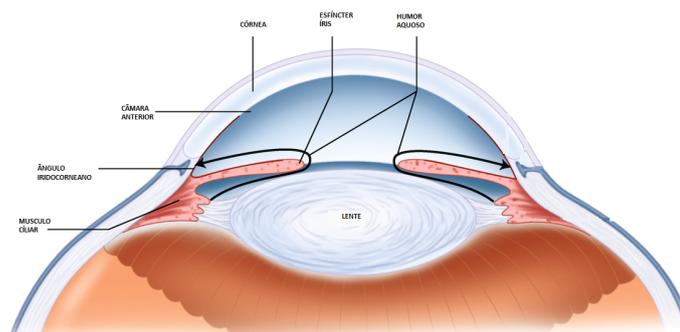
Figura 2 – Anatomia do Olho



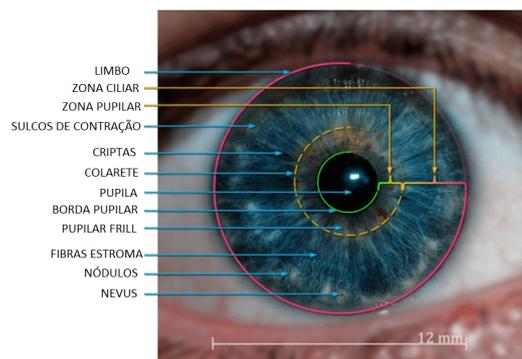
Fonte: Adaptado de (ZHU; ZHANG; RIO-TSONIS, 2012).

A região externa consiste na córnea e na esclera. A córnea reflete e transmite a luz ao cristalino e a retina, além de proteger o olho contra infecções e danos estruturais nas partes mais profundas. A esclera forma uma camada de tecido lubrificante que protege todo o globo ocular e mantém a sua forma. A córnea e a esclera estão conectadas no limbo. A parte visível da esclera é coberta por uma membrana mucosa transparente. A camada média do olho é composta pela íris, corpo ciliar e coróide. A íris controla o tamanho da pupila, portanto, a quantidade de luz que chega à retina; o corpo ciliar é um anel de tecido que envolve a íris e contém o músculo ciliar, que muda a forma do cristalino quando o olho foca em um objeto; e a coróide é uma camada vascular que fornece oxigênio e nutrientes para as camadas externas da retina. A camada interna do olho é a retina, uma estrutura complexa com camadas de neurônios que capturam e processam a luz. As três estruturas transparentes cercadas pelas camadas oculares são chamadas de humor aquoso, humor vítreo e cristalino (WILLOUGHBY et al., 2010). A Figura 3 ilustra a anatomia da íris, identificando as principais regiões internas e externas da íris.

Figura 3 – Anatomia da íris



(a)



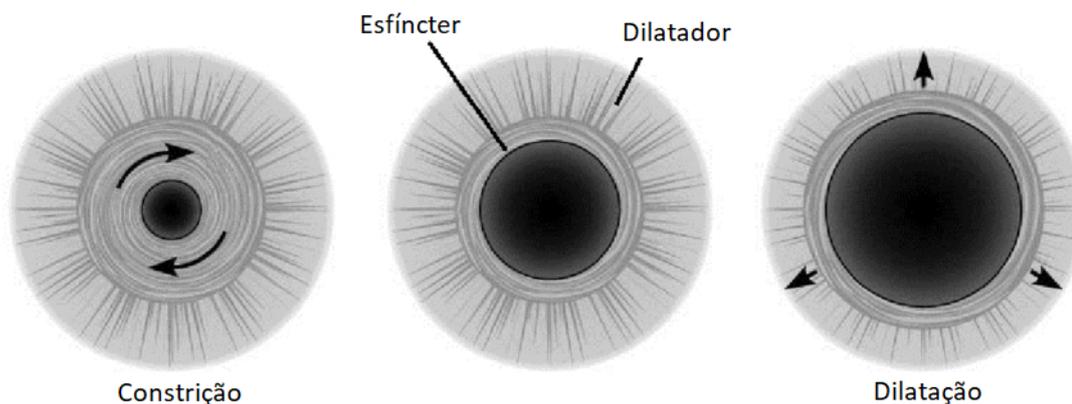
(b)

Fonte: Adaptado de (GUZMAN et al., 2013) e (SINHAROY, 2016)

A principal função da íris é ajustar o tamanho da pupila, indicando a quantidade de luz que entra na retina. A sua estrutura é complexa e consiste em muitas regiões

diferentes. A íris contém várias camadas, cada uma responsável por uma função na região. O epitélio pigmentar da íris (IPE) faz parte das camadas iniciais desta região e não contribui significativamente para a variação de cor da estrutura da íris (PELES et al., 2009). Nessa estrutura existem duas camadas conhecidas como músculos esfíncter e o músculo dilatador, ficam ao redor da íris e são responsáveis por controlar a dilatação e constrição pupilar (Figura 4).

Figura 4 – Músculos da Íris responsáveis pela Dilatação e Constrição



Fonte: Adaptado de (TOMEIO-REYES, 2015)

Além disso, a íris contém camadas que fazem parte da córnea, responsável pela proteção dessa região. A camada anterior e a camada do estroma também conhecido como tecido conectivo são as principais camadas desta região (Figura 4). A superfície do olho também pode ser dividida em zona ciliar e zona pupilar. Existem diversas espessuras para essas zonas, o que possibilita a variação da cor da estrutura. Vale mencionar que essas regiões são delimitadas por um anel de tecido chamado de colarate, que é justamente o que divide essa duas zonas na íris (PARRA et al., 2016). Existem muitos elementos texturais da íris que podem ser encontrados no olho humano saudável, são eles: criptas de *Fuch*, nódulos de *Wolfflin*, manchas pigmentares, sulcos de contração e melnose conjuntival (PARRA et al., 2016).

A pupila está localizada no centro da íris. Ela permite a entrada dos raios de luz para formação da imagem que chega ao cristalino e é focada na retina. A pupila possui um terço do tamanho da íris, podendo variar de 2 a 4 mm. O seu tamanho pode variar dependendo do esfíncter e dos músculos dilatadores que controlam a dilatação e a contração pupilar. Essa região não apresenta uma estrutura semelhante à íris, formada por camadas ou músculos, é uma região de cor escura que indica o fundo do olhos, podendo ser considerada uma abertura que promove a visão (SIROIS; BRISSON, 2014). A Figura 3 apresenta a anatomia da íris, identificando as principais regiões internas e externas da íris.

3.2 Pupilometria

A pupilometria é o estudo das mudanças do diâmetro pupilar em função de uma atividade cognitiva. Pesquisas iniciais relacionadas a dilatação pupilar realizavam estudos de casos em pacientes clínicos por meio de algumas ferramentas para entender as reações pupilares. O Berrien e Huntington (1943) usaram um telescópio de foco modificado, integrando uma luz infravermelha para avaliar o efeito emocional da mentira (SIROIS; BRISSON, 2014), mostrando um padrão de lentidão durante a dilatação e a constrição rápida mediante a decepção, desse modo o tamanho da pupila é usada para a compreensão do controle cognitivo, memória e principalmente das emoções.

Durante o processo de entendimento das reações pupilares, a pupilometria desempenha o papel de avaliar o esforço cognitivo mediante a realização de tarefas. O Winn et al. (2018b) afirma que a pupilometria contém algumas características importantes, a primeira é que é uma medida de série temporal, ou seja, tem a capacidade de mostrar todos os estágios de dilatação em pontos de tempos diferentes, no trabalho de McCloy et al. (2017) demonstra como a pupilometria mostra mudanças na dilatação em tarefas difíceis em vários momentos diferentes, um outro trabalho similar é o de Vogelzang, Hendriks e Rijn (2016) que apresenta mudanças no tempo de dilatação pupilar durante o processo de pronuncia de diferentes sentenças de uma língua. Dessa forma esses são alguns exemplos onde a pupilometria revela as mudanças nas cargas cognitivas durante a execução de uma tarefa que venha exigir um esforço cognitivo maior.

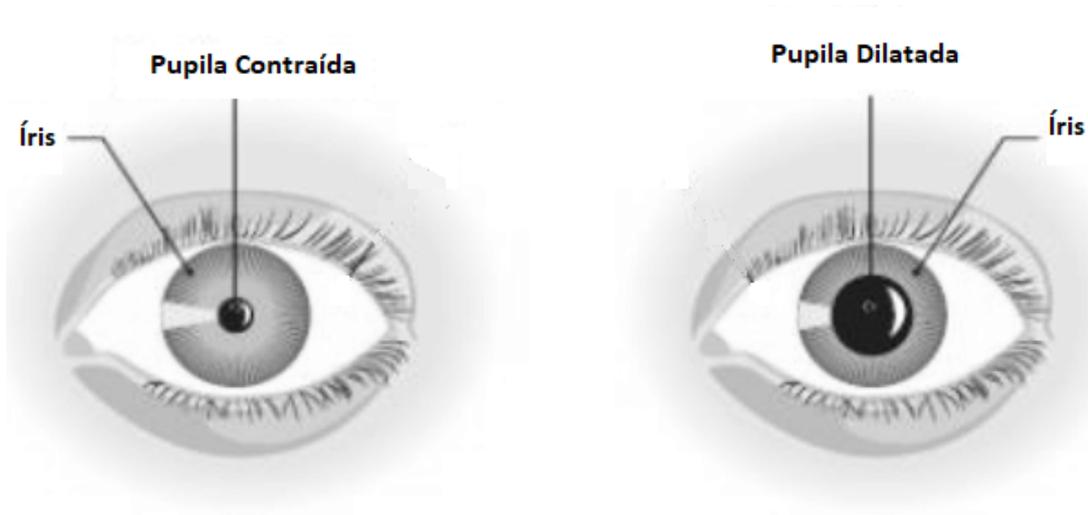
Segundo o Bremner (2004) a pupilometria não esta livre de limitações, portanto uma outra característica da pupilometria é usar as medidas encontradas para distinguir a anisocoria fisiológica da síndrome de *Horner* e estimar o campo visual. Desse modo a pupilometria é um efetivo indicador de várias percepções cognitivas e processos afetivos.

3.3 A dilatação e constrição pupilar

A dilatação e a constrição pupilar (Figura 5) podem ocorrer mediante à incidência de raios luminosos em frente a essa região. As pupilas podem ficar contraídas em condições de maior iluminação, e dilatar em condições de menor iluminação, isso é conhecido como reflexo pupilar à luz. Isso não significa que as reações pupilares (dilatação ou constrição pupilar), ocorram apenas com a influência de reflexos luminosos (SIROIS; BRISSON, 2014). Essas reações estão inteiramente ligadas aos aspectos cognitivos relacionadas ao esforço, excitação, interesse, exaustão e etc. Também estão relacionadas a diferenças individuais na capacidade de memória de trabalho e inteligência (TSUKAHARA; HARRISON; ENGLE, 2016). As reações pupilares estão ligadas ao sistema nervoso autônomo (SNA), que é responsável por controlar ações do corpo humano. Estudos revelam que a dilatação da pupila aumenta com o aumento da demanda de tarefas, refletindo as demandas ou carga pela execução

da tarefa e revelando o esforço exercido em resposta a essas demanda (GOLDINGER; PAPESH, 2012).

Figura 5 – Constrição e Dilatação Pupilar



Fonte: Adaptado de (GOLDINGER; PAPESH, 2012)

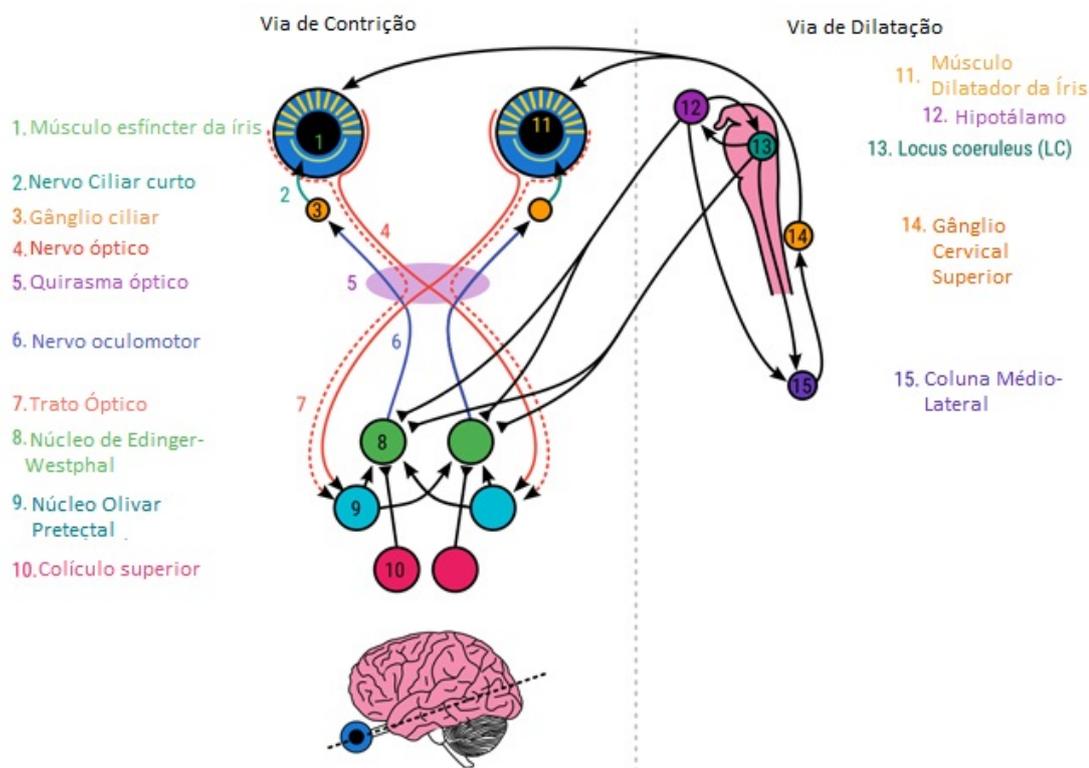
O diâmetro da pupila é um importante indicador da função cerebral, portanto neurocientistas tem relacionado o tamanho da pupila à cognição e à atividade cerebral mediante o esforço mental. Enquanto uma tarefa cognitiva é realizada, o esforço mental desperta o sistema nervoso simpático e, conseqüentemente, o diâmetro da pupila aumenta (JOSHI et al., 2016). A anatomia e a fisiologia das vias cerebrais que controlam o tamanho da pupila foram bem descritas e são conhecidas por envolver tanto o SNA quanto o somático, ou seja, a dilatação e constrição pupilar estão extremamente relacionadas ao sistema nervoso central (SNC) (EINHAUSER; KOCH; CARTER, 2010).

O sistema nervoso autônomo é responsável por regular os processos fisiológicos involuntários como a frequência cardíaca, pressão arterial, respiração, digestão e excitação sexual, e está dividido em simpático, parassimpático e entérico. O sistema nervoso simpático (SNS) e o sistema nervoso parassimpático (SNP) contêm fibras que fornece entrada sensorial e saída motora para SNC. O sistema nervoso entérico(ENS) é uma estrutura extensa, semelhante a uma teia, capaz de funcionar independentemente do restante do sistema nervoso (KAREMAKER, 2017; WAXENBAUM; VARACALLO, 2019). Então, o tamanho da pupila é controlado por duas vias neurais interconectadas, a via de constrição parassimpática e a via de dilatação simpática.

A via de constrição é controlada pelo músculo esfíncter da íris (Figura 4). Esse músculo fica ao redor da pupila, como um cordão que diminui o seu tamanho quando ela se contrai. Esse músculo é innervado (nervos) pelo sistema nervoso parassimpático, isto é, a

parte do sistema nervoso que mantém o corpo em condições estáveis. A via de dilatação é controlada pelo músculo dilatador, na qual consiste em fibras, ao redor da zona pupilar e se conectam com a região exterior e interior da íris. Quando o músculo dilatador se contrai o tamanho da pupila aumenta (Figura 4), sendo assim o músculo dilatador é controlado pelo sistema nervoso simpático, ou seja, é a parte do sistema nervoso que está envolvida com as respostas involuntárias fisiológicas do corpo. A Figura 6 mostra o funcionamento dessas vias ligadas ao SNA (MATHÔT, 2018).

Figura 6 – Vias de Contrição e Dilatação Pupilar



Fonte: Adaptado de (MATHÔT, 2018)

Dessa forma, a dilatação e a contração pupilar são objetos de estudos para a compreensão do sistema nervoso, porém a via neural relacionada à dilatação é difícil de compreensão, devido ao envolvimento de áreas do cérebro com muitos aspectos cognitivos. As vias de dilatação são expressas como um meio de identificar doenças neurológicas, pois podem apresentar algumas propriedades irregulares (tempo de dilatação e o diâmetro total da pupila), que caracterizam um mal funcionamento da cognição.

Através destas vias de dilatação pupilar, é possível diagnosticar possíveis doenças como o Alzheimer, devido o grande esforço cognitivo que pode ocorrer durante a execução de alguma tarefa intelectual (KREMEN et al., 2019). O Alzheimer pode ser detectado durante a dilatação pupilar, através de um alto esforço realizado por um indivíduo ao longo da realização de uma tarefa (GRANHOLM et al., 2017). Nota-se que através destas

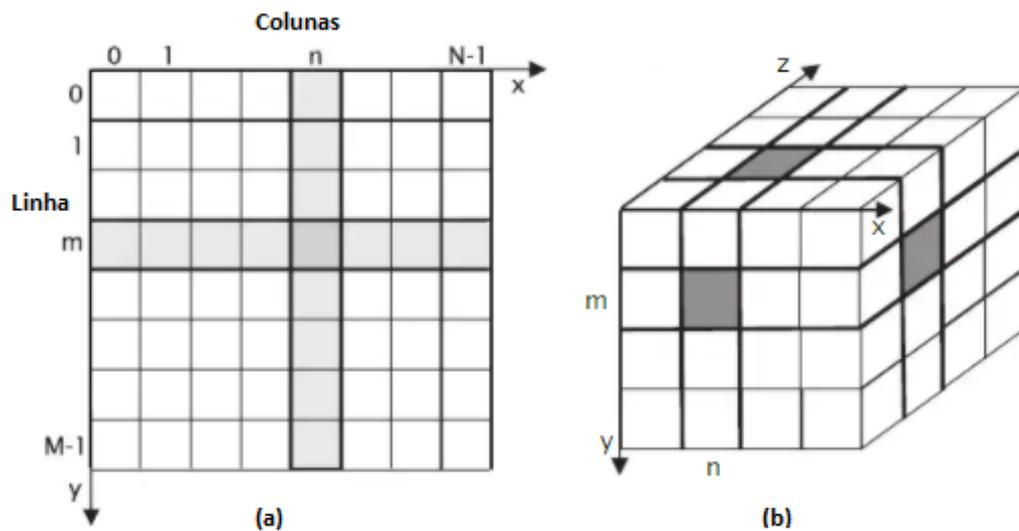
vias neurais de dilatação e constrição da pupila é possível avaliar e diagnosticar doenças neurológicas.

3.4 Processamento de Imagens

Processamento de imagem é qualquer forma de processamento de sinal para o qual a entrada é uma imagem, ou seja, é o estudo da representação e manipulação de informações pictóricas por um computador. O campo do processamento digital de imagens se refere ao processamento de imagens digitais por um computador digital visto que uma imagem digital é composta de um número finito de elementos, cada um com localização e valor específicos. Esses elementos são chamados de elementos pictóricos ou pixels. Pixel é o termo mais utilizado para representar os elementos de uma imagem digital (KOUR et al., 2012).

Uma imagem pode ser definida como uma função bidimensional, $f(x, y)$, x e y são coordenadas espaciais (plano), e a amplitude de f em qualquer par de coordenadas (x, y) é chamada de intensidade ou nível de cinza da imagem nesse ponto. Quando x, y e os valores de intensidade de f são quantidades finitas e discretas, chamamos de imagem digital (GONZALEZ; WORDS, 2010). A Figura 7 apresenta uma representação de uma imagem digital.

Figura 7 – Representação de uma Imagem Digital: (a) Imagem 2D e (b) Imagem 3D



Fonte: Adaptado de (JÄHNE, 2005)

A Figura 7 é uma representação de uma imagem digital em 2D e 3D. As imagens são constituídas por pixels, que em um computador digital são representados por meio de uma matriz de números. Sendo assim uma imagem, é constituída por linhas e colunas

($m \times n$). Cada posição $m \times n$ contém um valor numérico caracterizando uma cor que compõe a imagem (JÄHNE, 2005).

Existem etapas fundamentais para resolução de problemas de processamento de imagem, são elas: aquisição das imagens digitais, pré-processamento, segmentação, representação, descrição, reconhecimento e interpretação (GONZALEZ; WORDS, 2010). No entanto, uma metodologia pode seguir somente um subconjunto destas etapas. Esta seção apresenta as principais técnicas de processamento de imagens utilizadas no desenvolvimento deste trabalho, aplicadas em tarefas de melhoramento, representação e análise.

3.4.1 Modelos de Representação da Cores

As cores em processamento de imagens descrevem as propriedades de um objeto, essas propriedades são inerentes e distinguíveis pelo olho humano e são classificados como cromaticidade e brilho. A cromaticidade consiste em dois parâmetros, geralmente especificados como matiz e saturação. O sistema visual humano tem vários recursos que a caracterizam, tais como, a constância de cor e constância de brilho (WONG et al., 2015).

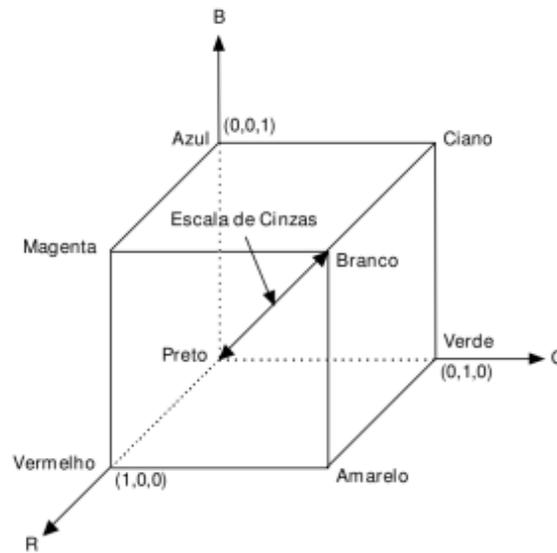
A constância de cor é aplicada à cromaticidade de imagens coloridas como o *automatic white balance* (AWB). Por outro lado, a constância de brilho é aplicada ao brilho das imagens coloridas como o aumento do contraste (WENG; CHEN; FUH, 2005).

Uma imagem colorida representa o mais alto nível de informação comparada a outras formas. Um modelo de cor é necessário para analisar dados presentes em uma imagem colorida (BHURCHANDI; NAWGHARE; RAY, 2009). Existem vários modelos amplamente utilizados em Computação Gráfica como modelo RGB, modelo HSI, CMY modelo etc. Cada modelo de cor tem uma natureza diferente para definir um pixel de uma imagem. O modelo de cor mais utilizado no processamento de imagens é o modelo de cor RGB (ROY; MITRA; SETUA, 2014).

No modelo RGB, cada pixel é representado por três valores distintos, caracterizando as cores: vermelha, verde e azul. A combinação dessa três cores forma um novo valor representando a cor, onde essa cor é definida em um espaço Euclidiano dimensional e contém coordenadas para um valor específico. A Figura 8 apresenta um exemplo do modelo de cor RGB.

O espaço de cores RGB pode ser representado como um cubo normalizado, contendo valores que representam as cores no intervalo $[0,1]$. Os valores correspondentes a cinza estão na diagonal principal dos valores proporcional a cor preta $(0,0,0)$, e no canto oposto estão os valores que equivalem à cor branca $(1,1,1)$. O RGB é considerado como o modelo de cor base para a maioria das imagens (IBRAHEEM et al., 2012).

Figura 8 – Modelo de cor RGB



Fonte: Adaptado (FILHO; NETO, 1999)

3.4.2 Operações Morfológicas

A morfologia matemática é uma ferramenta utilizada para extrair componentes de imagens que são úteis na representação e descrição da forma de uma região, além disso, as técnicas morfológicas são usadas para pré ou pós-processamento. As operações morfológicas são transformações lógicas com base na comparação de pixel próximos com um padrão predeterminado, elas se concentram em imagens binárias. As operações morfológicas oferecem uma visão unificada e é uma poderosa abordagem para inúmeras necessidades de processamento de imagem (ALAZAWEE; ABDEL-QADER; ABDEL-QADER, 2015).

Na imagem binária o conjunto de todos os pixels brancos é uma descrição morfológica completa da imagem. Então, os conjuntos em qualquer equação são membros do espaço inteiro 2D E^2 (GONZALEZ; WORDS, 2010). Como resultado, cada elemento do conjunto é um vetor 2D cuja coordenadas são as coordenadas (x, y) de um pixel branco ou preto na imagem. Os conceitos de reflexão e translação de conjuntos são considerados os conceitos mais importantes da morfologia da matemática por serem os principais e mais utilizados nas operações morfológicas (ALAZAWEE; ABDEL-QADER; ABDEL-QADER, 2015). A reflexão é definida na Equação 3.1.

$$\hat{B} = \{w/w = -b \text{ para } b \in B\} \quad (3.1)$$

O \hat{B} denota a reflexão de um conjunto B . B , é o conjunto de pixel (x, y) representando um objeto na imagem, w é formado multiplicando cada um dos elementos do conjunto B

por -1 , e b é um elemento de B . Então, \widehat{B} será o conjunto de pontos em B com novas coordenadas de $(-x, -y)$. A translação é expressa na Equação 2:

$$(B)_z = \{C/C = b + z \quad b \in B\} \quad (3.2)$$

O $(B)_z$ denota a translação de um conjunto B pelo ponto $Z = (z_1, z_2)$, e C é formado pela translação de cada um dos elementos de B por Z . Assim $(B)_z$ é o conjunto de ponto B cujo (x, z) as coordenadas foram substituídas por $(x + z_1, y + z_2)$ (ALAZAWEE; ABDEL-QADER; ABDEL-QADER, 2015).

3.4.2.1 Operação Morfológica de Dilatação

A operação de dilatação em processamento de imagens refere-se ao alongamento da imagem binária e é controlada por um elemento estruturante, isto é, definida matematicamente como uma operação de conjunto. Quando X está dilatado por Y , é denotado como $X \oplus Y$ e é definido como:

$$X \oplus Y = \{z | (\widehat{Y})_z \cap X \neq \emptyset\} \quad (3.3)$$

O Y denota o elemento estruturante, \widehat{Y} representa reflexão da coleção Y e \emptyset é o conjunto nulo. X dilatado por Y resulta em um conjunto que compreende a posição de origem do elemento estruturante (JOSHI; JAIN, 2020).

A dilatação foi proposta por Minkowski (1903), como uma operação teórica para caracterizar medidas integrais de certos conjuntos abertos (esparcos). A dilatação como uma operação de processamento de imagem foi empregada por vários pesquisadores iniciais no processamento de imagem como operações de suavização. A Dilatação é um operador de imagem para extração de forma e estimativa de parâmetros de imagem dessa forma todas as primeiras aplicações lidavam apenas com imagens binárias (HARALICK; STERNBERG; ZHUANG, 1987). A Figura 9 ilustra uma instância da operação de dilatação.

Na Figura 9 o primeiro operando A é considerado a imagem em análise, enquanto o segundo operando B é referido como o elemento estruturante, a ser pensado como constituindo um único parâmetro de forma da transformação de dilatação.

3.4.2.2 Operação Morfológica de Erosão

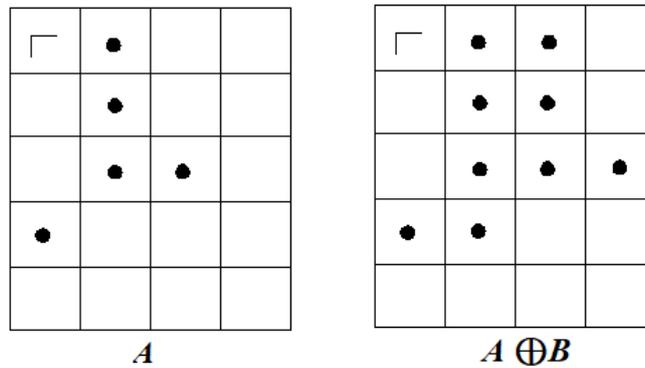
É o oposto do processo de dilatação. Isto refere-se ao encolhimento ou afinamento da imagem binária e é controlado por elemento estruturante. Quando X é erodido por Y , é denotado como $X \ominus Y$, e é definido conforme apresentado na Equação 3.4.

$$X \ominus Y = \{z | (Y)_z \cap X^c \neq \emptyset\} \quad (3.4)$$

Figura 9 – Ilustração da operação de Dilatação

$$A = \{(0,1), (1,1), (2,1), (2,2), (3,0)\}$$

$$B = \{(0,0), (0,1)\}$$



$$A \oplus B = \{(0,1), (1,1), (2,1), (2,2), (3,0), (0,2), (1,2), (2,2), (2,3), (3,1)\}$$

Fonte: Adaptado de (HARALICK; STERNBERG; ZHUANG, 1987)

O Y denota elemento de estrutura, X^c denota o suplemento da coleção X e \emptyset é o conjunto nulo. Dessa forma, a erosão e a transformação morfológica combinam dois conjuntos usando a subtração vetorial de elementos do conjunto. Se A e B são conjuntos no espaço N euclidiano, então a erosão de A por B é o conjunto de todos os elementos x para os quais $x + b \in A$ para cada $b \in B$ (HARALICK; STERNBERG; ZHUANG, 1987). A Figura 10 apresenta uma ilustração da operação de erosão.

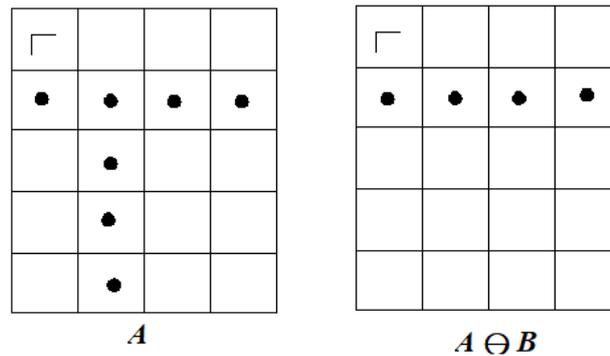
A transformação de erosão em uma imagem binária é uma generalização da transformada da distância proposta por Rosenfeld e Pfaltz (1968), na literatura. Sendo assim Chen e Haralick (1995) expõe que transformação de erosão de A em relação a um elemento estruturante K produz uma imagem em escala de cinza onde o nível de cinza de cada pixel $x \in A$ é a distância generalizada de x para o plano de fundo da imagem, ou seja, n é o maior inteiro positivo.

Neste trabalho as operações morfológicas são usadas para remoção de falsos positivos do resultado da segmentação da íris e pupila, durante a etapa de pós-processamento das imagens segmentadas, conforme será apresentada na subseção 4.3. Desse modo antes de realizar a aplicação dos operadores, é necessário encontrar os falsos positivos de acordo com a área dos pixel conectados da imagem segmentada. Apenas os operadores de erosão

Figura 10 – Ilustração da Operação de Erosão

$$A = \{(1, 0), (1, 1), (1, 2), (1, 3), (1, 4), (1, 5), (2, 1), (3, 1), (4, 1), (5, 1)\}$$

$$B = \{(0, 0), (0, 1)\}$$



$$A \ominus B = \{(1, 0), (1, 1), (1, 2), (1, 3), (1, 4)\}$$

Fonte: Adaptado de (HARALICK; STERNBERG; ZHUANG, 1987)

e dilatação são usadas durante o desenvolvimento do trabalho.

3.4.3 Segmentação

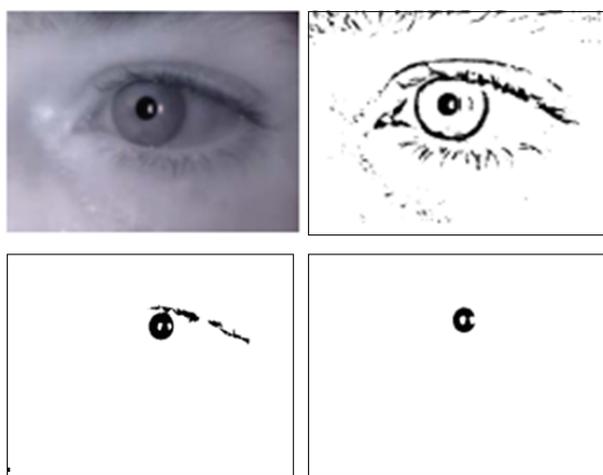
A segmentação de imagens é um passo fundamental no processamento digital de imagens e desempenha um papel importante na tecnologia de processamento digitais em imagem. Essa técnica extrai uma região de interesse que tem um impacto importante na tarefa de reconhecimento de imagens (MARUSINA; VOLGAREVA; SIZIKOV, 2015). Existem vários meios de realizar o processo de segmentação de imagens. Cheng e Li (2021) discutem o método de segmentação por limiar, onde o princípio deste método é processar uma imagem a ser segmentada determinando o valor limiar. Quando objetos com o mesmo pixel são encontrados, eles são segmentados uniformemente. O método de segmentação por limiar é na verdade a seguinte transformação da imagem de entrada f para uma imagem de saída g :

$$g(i, j) = \begin{cases} 1, & f(i, j) \geq T \\ 0, & f(i, j) < T \end{cases} \quad (3.5)$$

no qual T é o valor limite, para os elementos da imagem do objeto, $g(i, j) = 1$, e para os elementos de imagem do fundo $g(i, j) = 0$.

Um outro meio de realizar a segmentação segundo Cheng e Li (2021) é por meio do método de segmentação por *cluster* que exhibe diferentes regiões de uma imagem de acordo com um conjunto de imagens, além disso, segmenta imagens em escala de cinza. Existem duas abordagens diferentes para realizar a segmentação baseada em *cluster*, o *K-means* e *Fuzzy K-means* (KAUR, 2014). A Figura 11 mostra um exemplo do processo de segmentação usando um limiar.

Figura 11 – Exemplo do processo de segmentação da pupila usando limiar em imagens infravermelhas dos olhos.



Fonte: Adaptado de (NAMBIAR, 2020)

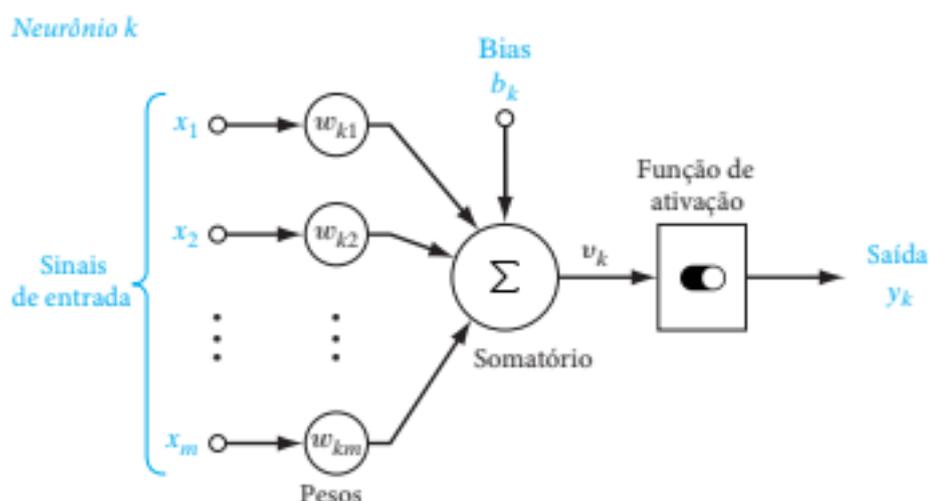
Além de técnicas de processamento de imagens para realizar a tarefa de segmentação de imagens, o uso do aprendizado profundo é o mais moderno método aplicado para este tipo de tarefa, dessa forma arquiteturas de redes neurais surgem com a finalidade de realizar esse processo de maneira eficaz. A segmentação de imagens usando rede neurais consiste em inserir uma imagem e seus segmentos em uma estrutura complexa contendo neurônios e camadas, cujo os rótulos de entrada na rede neural indicam os pixels da região de interesse a ser segmentada (MOHAMED et al., 2020). Uma rede neural contém uma estrutura única é usada em várias tarefas da visão computacional.

Neste trabalho é adotado o processo de segmentação usando uma arquitetura de rede neural convolucional conforme será apresentado no próximo capítulo. Dessa forma o objetivo é obter as regiões segmentadas da íris e pupila a fim de aplicar outras etapas da metodologia durante o desenvolvimento da pesquisa, portanto, é a partir dessas regiões obtidas no processo de segmentação que será possível definir o estado da reação pupilar.

3.5 Redes Neurais

Uma rede neural é uma máquina projetada para representar a forma com que o cérebro executa uma tarefa específica e é composta por uma série de unidades simples de processamento que coletam e armazenam o conhecimento obtido a partir de um conjunto de dados de entrada, a organização de uma rede neural pode simular o sistema neural biológico para fazer respostas interativas com objetos do mundo real (NINGNING, 2010). A essas unidades de processamento dá-se o nome de neurônios artificiais. Na rede neural, os neurônios são interligados por meio de conexões chamadas de sinapses. As sinapses entre os neurônios são fundamentais para o processo de aprendizado no qual a rede adquire conhecimento (HAYKIN, 2010). A Figura 12 apresenta estrutura de um neurônio artificial.

Figura 12 – Estrutura de um Neurônio Artificial



Fonte: Adaptado de (HAYKIN, 2010)

A Figura 12 apresenta um exemplo de um neurônio que é uma unidade de processamento de informações fundamental para a operação de uma rede neural. Nesta Figura são apresentadas três elementos básicos de um modelo de rede neural, o primeiro constitui um conjunto de sinapses ou links de conexão, cada um caracterizado por um peso ou força própria. A segunda etapa soma os sinais de entradas ponderadas pelas respectivas forças sinápticas do neurônio. As operações aqui descritas constituem um combinador linear e na última etapa é usada uma função de ativação para limitar a amplitude da saída de um neurônio.

Existem as funções de ativação lineares e não-lineares. Em relação a esta última categoria, citam-se como exemplos a função sigmoide, tangente hiperbólica e as recentemente propostas *Rectified Linear Units* (ReLU) e a *Leaky ReLU*, cuja velocidade de convergência

em relação às supracitadas é superior (MAAS; HANNUN; NG, 2013). Essas funções são expressas respectivamente pelas Equações 3.6 e 3.7.

$$f(n) = \max(0, x) \quad (3.6)$$

$$f(n) = \max(0.1, x) \quad (3.7)$$

As redes neurais artificiais (RNA) fazem parte de um dos métodos de *machine learning*, e visam simular o funcionamento do cérebro humano (RAWAT; WANG, 2017) baseado em várias funções matemáticas, contendo um conjunto de neurônios e suas ligações. O principal objetivo de uma RNA é executar tarefas computacionais fundamentadas em ações cotidianas do ser humano, a exemplo da classificação de animais, pessoas e objetos. Esses tipos de atividades são complexas para uma máquina, e por esse motivo, técnicas baseadas em aprendizado profundo são desenvolvidas para obter o melhor desempenho em tarefas de visão computacional.

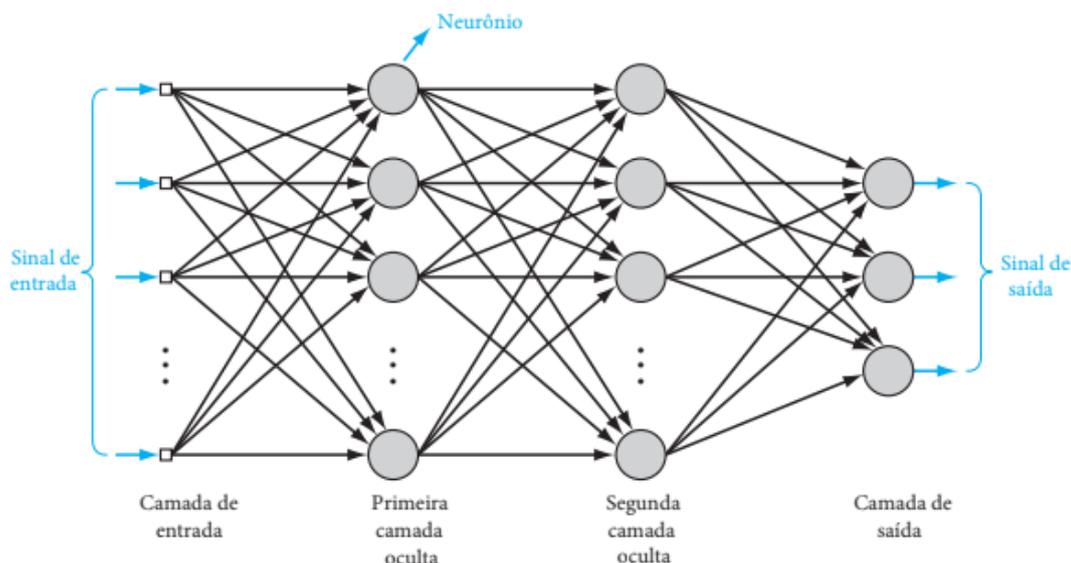
3.6 Redes *Multilayer Perceptron*

A Perceptron é a forma mais básica da arquitetura de rede neural com zero camadas ocultas, usada com mais destaque para classificação de padrões linearmente separáveis. O modelo de rede neural Perceptron se enquadra na categoria de aprendizagem supervisionada e pode ser referida como um classificador binário. O algoritmo é vantajoso no cenário em que queremos priorizar o impacto das características selecionadas sobre os outros (usando pesos). Esse algoritmo consiste em um único Neurônio *McCulloch-Pitts* contendo sináptica ajustável pesos e *bias* (SINGH; BANERJEE, 2019).

Redes de *Multilayer Perceptron*(MLP), são redes neurais que apresentam uma ou mais camadas de neurônios ocultos, que não fazem parte da entrada ou saída da rede (HAYKIN, 2004).

Segundo Haykin (2009), este algoritmo também é baseado em um único neurônio linear com pesos ajustáveis, o que limita o poder de computação do algoritmo. Para superar as limitações práticas do *perceptron* e do algoritmo *Least Mean Square* (LMS), surgiu uma estrutura de rede conhecida como perceptron multicamadas. Os três pontos a seguir destacam as características básicas dos perceptrons multicamadas:

- O modelo de cada neurônio da rede inclui uma função de ativação não linear que é diferenciável.
- A rede contém uma ou mais camadas ocultas, uma de entrada e outra de saída.

Figura 13 – Rede Neural *Multilayer Perceptron*

Fonte: Adaptado de (HAYKIN, 2009)

- A rede exibe um alto grau de conectividade, cuja extensão é determinada pelos pesos sinápticos da rede

A Figura 13 apresenta a arquitetura de rede neural *Multilayer Perceptron*, dois tipos de sinais são identificados segundo Haykin (2009).

- **Função Sinal:** Uma função sinal é um sinal de entrada (estímulo) na extremidade da rede, propaga-se para frente (neurônio por neurônio) através do rede, e emerge na extremidade de saída da rede como um sinal de saída. O sinal é calculado como funções das entradas e pesos associados aplicada a esse neurônio.
- **Sinais de Erro:** Um sinal de erro se origina em um neurônio de saída da rede e propaga para trás (camada por camada) através da rede. Esse termo surgiu porque sua computação por cada neurônio da rede envolve uma função dependente de erro.

As redes MPL's usam a regra de aprendizado por correção de erro, através do algoritmo de retropropagação de erro, ou seja, o *backpropagation*. Esse tipo de rede consiste em dois passos considerados como o processamento direto e processamento reverso. O processamento direto é o processo de propagação camada por camada, do efeito causado pela entrada aplicada a rede neural, possibilitando a não alteração dos pesos durante esse tipo de processamento. Já no processamento reverso é aplicado o *backpropagation*, dessa forma possibilita o ajuste dos peso de acordo com uma regra de correção baseada na saída correta esperada. Esse sinal de erro é calculada na saída da rede neural e é propagada de maneira reversa camada por camada (HAYKIN, 2004).

Segundo Haykin (2009) o sinal de erro é representado por:

$$e_j(n) = s_j - y_j(n) \quad (3.8)$$

Onde j representa um neurônio e $s_j(n)$ corresponde à resposta desejada para um determinado neurônio da camada de saída, o n corresponde à interação (apresentação do n -ésimo padrão de treinamento). Sendo assim, esse sinal de erro é propagado de maneira reversa durante o treinamento da rede MPL.

Durante o processo são computados o erro quadrático médio 3.10 e o erro quadrático instantâneo 3.9, onde a soma dos dois é uma função de todos os pesos sinápticos da rede que definem a função de transferência ou mapeamento de entrada e saída da rede. Dessa forma, os ajustes dos pesos acontecem mediante o erro encontrado para cada padrão apresentado a rede.

$$\xi(n) = \frac{1}{2} \sum_{j \in C} e_j^2(n), \quad (3.9)$$

$$\xi_{a,v}(n) = \frac{1}{N} \sum_{n=1}^N \xi(n), \quad (3.10)$$

Para minimizar o erro, é necessário aplicar aos pesos uma correção proporcional ao oposto do gradiente de erro, sendo representado essa regra da seguinte forma:

$$w_{ij}(n+1) = w_{ji}(n) - \eta(n) \frac{\partial \xi(n)}{\partial w_{ij}(n)}, \quad (3.11)$$

De maneira geral a MLP durante o *backpropagation* são calculado os gradientes locais para cada neurônio na rede a fim de ajustar os pesos durante as interações no treinamento.

3.7 Aprendizado Profundo

O aprendizado profundo é um subcampo do aprendizado de máquina, que tenta aprender abstrações de alto nível em dados, utilizando arquiteturas hierárquicas. É uma abordagem emergente e tem sido amplamente aplicada em domínios de inteligência artificial, como análise semântica, aprendizado por transferência, processamento de linguagem natural, visão computacional entre outras (REN; XU, 2015). O aprendizado profundo está inserido em quase todos os campos de pesquisas como, a biologia, psicologia, medicina e segurança da informação devido apresentar, executar e realizar a análise de dados de maneira precisa em uma determinada tarefa existente (WU et al., 2017).

Várias abordagens de aprendizado profundo têm sido amplamente revisadas e discutidas. Alguns autores como Ciresan, Meier e Schmidhuber (2012) enfatizam as importantes inspirações e contribuições técnicas em um formato de linha do tempo histórico, já Bengio (2013) examinou os desafios do aprendizado profundo e propôs algumas pesquisas relacionadas. As redes profundas têm se mostrado bem sucedidas para tarefas de visão computacional porque elas podem extrair recursos apropriados enquanto executam em conjunto de tarefas. A ideia por trás da aprendizagem profunda vem do potencial das camadas ocultas. Por usar muitas camadas de processamento, a aprendizagem profunda tem mais chances de aprender a representação correta para um conjunto de dados. Por meio do algoritmo de retropropagação de erro, a aprendizagem profunda pode descobrir estruturas complexas em grandes conjuntos de dados. (LECUN; BENGIO; HINTON, 2015).

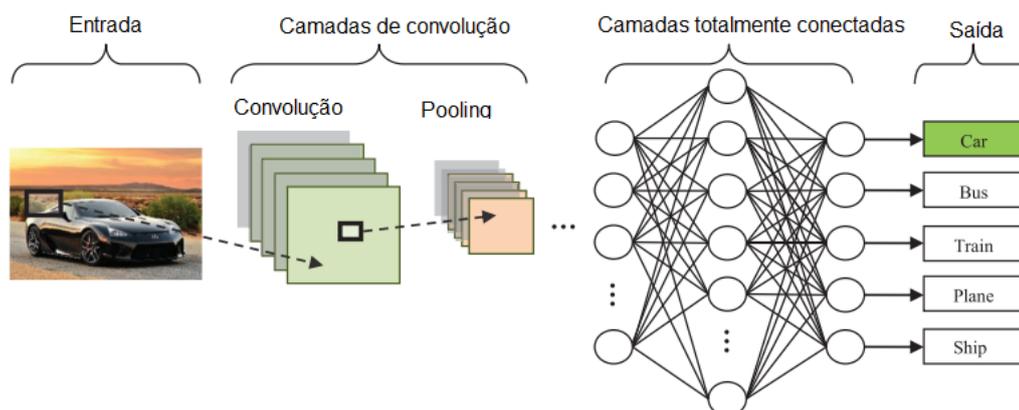
O uso do aprendizado profundo, possibilita a inserção de mais camadas ocultas em redes neurais artificiais, a fim de gerar modelos específicos para uma determinada tarefa, com a intenção de alcançar um desempenho de ponta, dessa forma são propostas arquiteturas de rede neurais profundas. Devido a esse alto poder e desempenho do aprendizado profundo, os avanços na tarefa de segmentação e detecção são precisas ao usar arquiteturas de Redes Neurais Convolucionais específicas. Na seção a seguir serão apresentadas as Redes neurais Convolucionais.

3.8 Redes Neurais Convolucionais

Com o avanço nos campos de redes neurais profundas várias arquiteturas de RNA são propostas para pesquisas e produção de tecnologias com inteligência artificial. As redes neurais convolucionais (CNN) se constituem em uma dessas arquiteturas de redes neurais profundas, caracterizadas por uma organização de várias camadas convolucionais, camadas de ativação, de agrupamento e entre outras camadas em um único esqueleto (ELHASSOUNY; SMARANDACHE, 2019). As CNNs são arquiteturas do tipo *feedforward* onde o fluxo de informações ocorre somente em uma direção, de suas entradas para suas saídas.

As CNNs, de acordo com LeCun, Kavukcuoglu e Farabet (2010), são modelos biologicamente inspirados que podem aprender características em multiestágios. A entrada e a saída de cada estágio são compostas de conjuntos de matrizes denominados mapas de características. Por exemplo, se a entrada for uma imagem colorida, cada mapa de características seria um vetor 2D contendo um canal de cor da imagem de entrada (para uma entrada de áudio, cada mapa de características seria um vetor 1D e para um vídeo ou imagem volumétrica, seria um vetor 3D). Na saída, cada mapa de características representa um recurso específico extraído em todos os locais do objeto de entrada. Ao final, a arquitetura possui uma camada responsável pela discriminação/rotulação de cada

Figura 14 – Exemplo típico de uma CNN



Fonte: Adaptado de (RAWAT; WANG, 2017)

imagem.

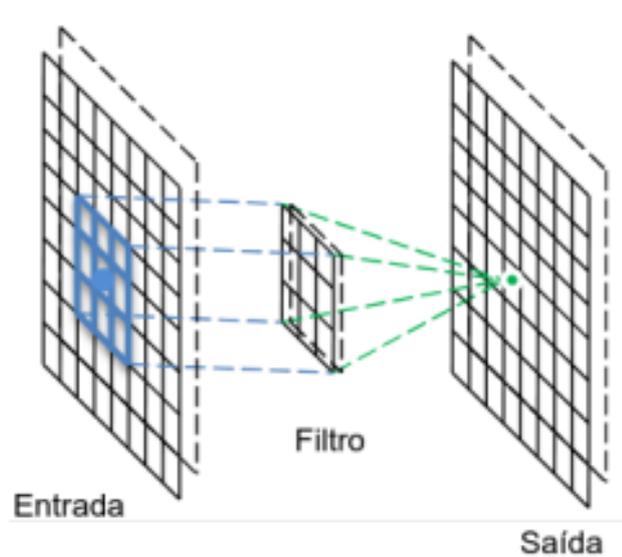
As CNN's são poderosas e profundos algoritmos de aprendizagem, capazes de lidar com milhões de parâmetros, convolvendo as imagens 2D com filtros de vários tamanhos, extraíndo e aprendendo máximo de características de uma imagem em uma determinada tarefa. (CHAUHAN; GHANSHALA; JOSHI, 2018).

Em geral, a arquitetura de uma CNN é formada por camadas convolucionais (*convolutional*) e de *pooling* (ou *subsampling*), que são agrupadas em módulos, ou blocos. E uma ou mais camadas totalmente conectadas (*fully connected layer*), sempre ao final da rede. Os módulos são frequentemente empilhados uns sobre os outros para formar um modelo profundo. A Figura 14 ilustra a arquitetura típica de uma CNN para uma tarefa de classificação de uma imagem. Uma imagem é inserida diretamente na rede e isso é seguido por vários estágios de convolução e *pooling*. Posteriormente, as representações dessas operações alimentam uma ou mais camadas totalmente conectadas. Finalmente, a última camada totalmente conectada produz o rótulo indicativo da classe (RAWAT; WANG, 2017).

A CNN geralmente utiliza em sua arquitetura três tipos de camadas: convolução, subamostragem e completamente conectada. As camadas de convolução e subamostragem compõem as camadas iniciais da rede, e são implicitamente responsáveis pela extração de características. A camada completamente conectada, por fim, localiza-se no final da rede conforme a Figura 14 (GUO et al., 2016).

As CNN's podem ser aplicadas em várias tarefas específicas no aprendizado profundo, elas incluem classificação de Imagens e Segmentação, Detecção de Objetos, Processamento de Vídeo, Processamento de Linguagem Natural e Fala Reconhecimento. Dessa forma, são propostas arquiteturas de CNN's diferentes a fim de cumprir determinadas tarefas.

Figura 15 – Ilustração da Camada de Convolução



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

Neste trabalho, são utilizadas duas arquiteturas de redes neurais convolucionais, são elas: VGG-16 e U-Net. As duas arquiteturas serão apresentadas nas subseções 3.8.3 e 3.8.4.

3.8.1 Camada de Convolução

A camada de convolução é responsável pela maioria dos cálculos envolvidos. Os principais parâmetros de uma camada de convolução são um conjunto de filtros ou kernel convolvidos com os mapas de características com a finalidade de produzir um mapa bidimensional separado. Os neurônios das camadas convolucionais são organizados em mapas de características. Cada neurônio em um mapa de características tem um campo receptivo, que é conectado a uma vizinhança de neurônios na camada anterior através de um conjunto de pesos treináveis.

Os mapas de características produzidos compartilham pesos gerados por toda a rede neural para reduzir a sua complexidade mantendo o número baixo de parâmetros (ZEILER; FERGUS, 2014). A Figura 15 mostra uma ilustração da camada de convolução.

Na camada de convolução são definidas alguns hiperparâmetros para a extração de característica, como o número de filtros e o *stride* que indica o movimento do filtro de acordo com o campo receptivo usado na camada conforme mostra a Figura 15 (ALOYSIUS; GEETHA, 2017). o k -ésimo mapa de características de saída Y_k pode ser calculado como:

$$Y_k = f(W_k * x), \quad (3.12)$$

onde a imagem de entrada é denotada por x ; o filtro convolucional relacionado ao k -ésimo mapa de características é denotado por W_k ; o sinal de multiplicação neste contexto refere-se ao operador convolucional 2D, que é usado para calcular o produto interno do modelo de filtro em cada local da imagem de entrada; E $f(x)$ representa a função de ativação não linear (YU et al., 2014b).

As funções de ativação não-lineares permitem a extração de características não-lineares. Tradicionalmente, eram utilizadas as funções sigmóide e tangente hiperbólica. Recentemente, a função Relu tornou-se popular conforme a Equação 3.6

3.8.2 Camada de *Pooling*

A camada *Pooling* tem a função de reduzir a resolução espacial dos mapas de características extraídos da camada de convolução (Figura 16). O *pooling* seleciona o máximo de recurso representativo de uma pequena região. O emprego da camada de *pooling* em CNN's visa obter invariância para mudanças na posição ou nas condições de iluminação, robustez à desordem e compactidade de representação. Em geral, a camada de *pooling* resume as saídas de grupos vizinhos de neurônios no mesmo mapa de kernel. (YU et al., 2014a)

Nas CNNs, existem dois métodos convencionais de *pooling*, incluindo *max pooling* e *average pooling*. O *max pooling* seleciona o maior elemento em cada região de agrupamento como:

$$Y_{kij} = \max_{(p,q) \in \mathfrak{R}_{ij}} x_{kpq}, \quad (3.13)$$

Onde a saída da operação de agrupamento, associada ao k -ésimo mapa de características, é denotada por Y_{kij} , x_{kpq} denota o elemento na localização (p, q) contido pela região de agrupamento \mathfrak{R}_{ij} , que incorpora um campo receptivo ao redor da posição (i, j) (YU et al., 2014a). Já em relação ao método de *pooling* médio, ele toma a média aritmética dos elementos em cada região de *pooling* como:

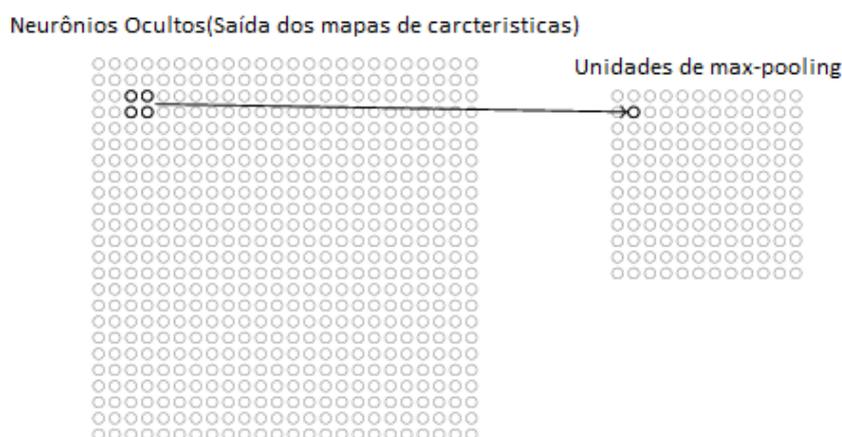
$$Y_{kij} = \frac{1}{|R_{ij}|} \sum_{(p,q) \in R_{ij}} x_{kpq}, \quad (3.14)$$

onde $|R_{kij}|$ min representa o tamanho da região de *pooling* R_{ij} . (YU et al., 2014a)

3.8.3 VGG16

A VGG-16 é uma arquitetura de rede neural proposta por Simonyan e Zisserman (2015), amplamente usada no conjunto de dados ImageNet (DENG et al., 2009). Esse

Figura 16 – Ilustração do Pooling após a camada de convolução



Fonte: Adaptado de (ACADEMY, 2022)

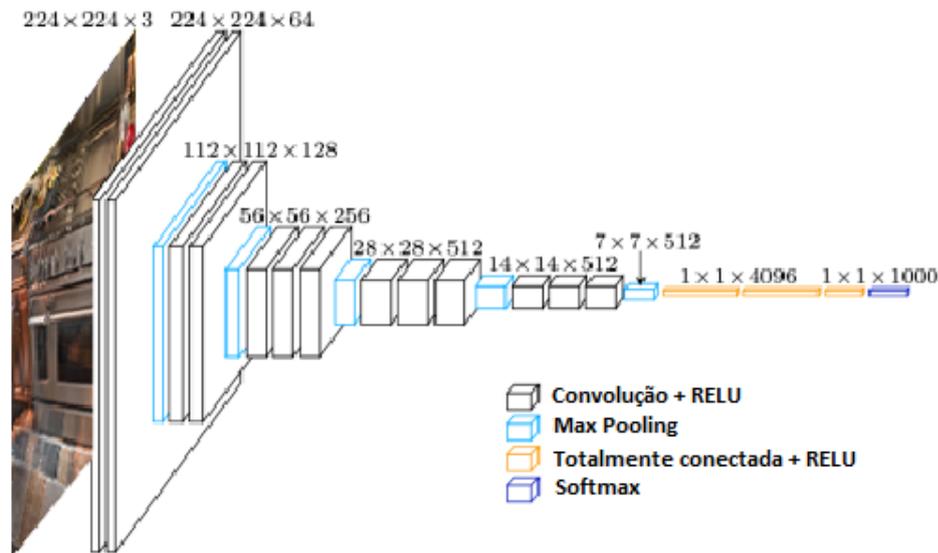
modelo apresentou um bom desempenho neste conjunto de dados demonstrando um excelente desempenho durante a tarefa de classificação.

Essa arquitetura tem 16 camadas variando entre camadas de convolução e camadas *pooling*. A arquitetura recebe uma imagem fixa 224x224, que passa por uma pilha de camadas convolucionais (Conv), onde é aplicada um filtro com um campo receptivo pequeno de 3×3 . Em uma das configurações também é usada filtros de convolução 1×1 , que podem ser vistos como uma transformação linear dos canais de entrada (seguida de não linearidade). O hiperparâmetro *stride* de convolução é fixo em 1 pixel. O agrupamento espacial é feito por cinco camadas max-pooling antecedidas por camadas de convolução (nem todas as camadas de convolução são seguidas de camadas *pooling*). O pool máximo é realizado em uma janela de 2×2 pixels, com o *stride* de tamanho 2. Na Figura 17 é apresentada a estrutura da VGG16. (SIMONYAN; ZISSERMAN, 2015).

Uma pilha de camadas convolucionais é seguida por três camadas totalmente conectadas (FC): as duas primeiras têm 4096 neurônios cada, a terceira contém 1000 neurônios (um para cada classe de classificação (SIMONYAN; ZISSERMAN, 2015)). Ao final de cada uma dessas camadas existe uma camada oculta de ativação usando *RELU*. E Por fim, na estrutura clássica a última camada é uma softmax tendo em vista que o problema enfrentado é de classificação.

A configuração das camadas totalmente conectadas é a mesma usada em todas as redes. Todas as camadas ocultas são equipadas com *RELU* (KRIZHEVSKY; SUTSKEVER; HINTON, 2012). Na Tabela 2 é apresentada as configurações da arquitetura VGG16.

Figura 17 – Arquitetura VGG16



Fonte: Adaptado de (LI; GUO, 2018)

A arquitetura VGG16 apresenta bons resultados em relação a outras arquiteturas de rede neural convolucionais, apresentando um desempenho superior em relação a AlexNet (KRIZHEVSKY; SUTSKEVER; HINTON, 2012).

Essa arquitetura é integrada no bloco de codificação da rede neural U-Net a fim de compor e estruturar a arquitetura de rede neural usada neste trabalho, conforme será ilustrada na Figura 23 da subseção 4.2.

3.8.4 U-Net

A U-Net (RONNEBERGER; FISCHER; BROX, 2015) é uma arquitetura de rede neural profunda para tarefas de segmentação de imagens baseada na classificação dos pixels. Essa arquitetura tem uma estrutura simétrica de codificador e decodificador com conexões de salto. O codificador contém uma sequência de camadas de convolução e camadas *pooling*, enquanto o decodificador realiza a expansão e usa convoluções transpostas ou de convoluções (JURDI et al., 2020).

A Figura 18 apresenta a arquitetura da rede neural constituída pelo caminho de contração (lado esquerdo) e pelo caminho expansivo (lado direito). A via de contração segue a arquitetura típica de uma rede convolucional. Consiste na aplicação quatro blocos compostas por duas convoluções 3×3 , seguidas por uma camada oculta usando a função de ativação ReLU e uma camada de max pooling 2×2 com o *stride* de tamanho 2 para *downsampling*. A cada etapa do *downsampling* o número de característica extraída é dobrado. O caminho expansivo está dividido em quatro blocos, cada bloco é composto por uma convolução transposta 2×2 , conexões de salto que concatenam a saída das camadas de

Tabela 2 – Arquitetura VGG16

Configuração da VGG_16				
A	A-LRN	B	C	D
11 weight layers	11 weight layers	13 weight layers	16 weight layers	16 weight layers
input (224 × 224 RGB image)				
conv3-64	conv3-64 LRN	conv3-64 conv3-64	conv3-64 conv3-64	conv3-64 conv3-64
maxpool				
conv3-128	conv3-128	conv3-128 conv3-128	conv3-128 conv3-128	conv3-128 conv3-128
maxpool				
conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256 conv1-256	conv3-256 conv3-256 conv3-256
maxpool				
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 conv1-512	conv3-512 conv3-512 conv3-512
maxpool				
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 conv1-512	conv3-512 conv3-512 conv3-512
maxpool				
FC-4096				
FC-4096				
FC-1000				
soft-max				

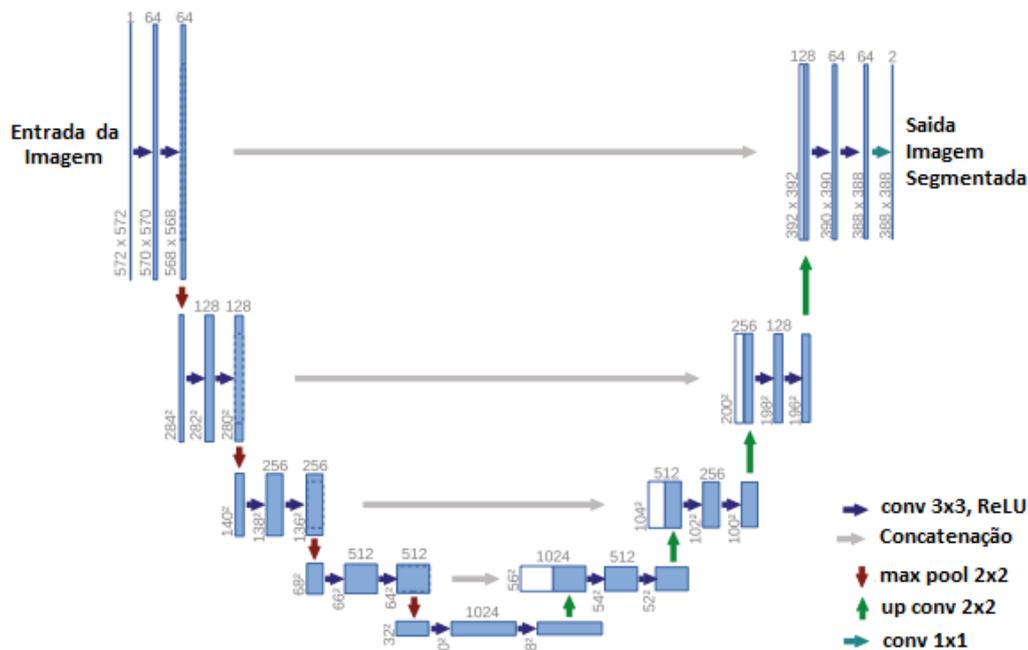
Fonte: Adaptado de (SIMONYAN; ZISSERMAN, 2015)

convolução transpostas juntamente com os mapas de características do codificador, duas camadas de convolução 3x3 seguidas de uma função de ativação Relu (RONNEBERGER; FISCHER; BROX, 2015).

Na camada final há uma convolução 1x1 usada para mapear cada vetor de característica dos 64 componentes para o número desejado de classes. No total, a rede possui 23 camadas convolucionais. Para permitir um mosaico perfeito do mapa de segmentação de saída, é importante selecionar o tamanho do bloco de entrada de modo que todas as operações de *max pooling* 2x2 são aplicados a uma camada com um mesmo tamanho x e y (RONNEBERGER; FISCHER; BROX, 2015).

A rede U-Net foi desenvolvida com base nos trabalhos de Long, Shelhamer e Darrell

Figura 18 – Arquitetura da Rede Neural U-Net



Fonte: Adaptada de (RONNEBERGER; FISCHER; BROX, 2015)

(2015) usando redes totalmente convolucionais. Sua implementação alcançou melhor desempenho do que o melhor anterior no desafio *International Symposium on Biomedical Imaging (ISBI) 2012* (JURDI et al., 2020), superando o estado da arte na época por um considerável margem. A estrutura da U-Net resulta em uma rede semelhante a uma forma de *U*, propagando as informações contextuais ao longo da rede, que permite segmentar objetos em uma área usando o contexto de uma área de sobreposição maior (JURDI et al., 2020).

Existem diversas adaptações e variações da arquitetura de rede U-Net apresentando um excelente desempenho em imagens biomédicas. A FD-Unet é uma readaptação da arquitetura, visando obter bons resultados de segmentação em imagens da região dos olhos. Essa alteração da arquitetura usa convolução dilatada em vez da convolução original, para extrair mais características globais das imagens processadas (ZHANG et al., 2019).

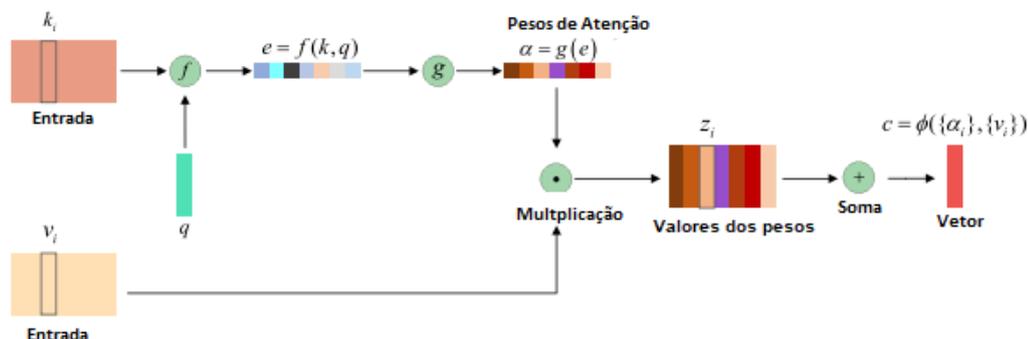
A arquitetura de rede neural convolucional U-Net é usada durante o processo de segmentação da íris e pupila deste trabalho. Durante o desenvolvimento desta metodologia são apresentadas e implementadas algumas modificações nesta arquitetura de rede neural a fim melhorar o desempenho da etapa de segmentação conforme apresentado no Capítulo 4, portanto, a U-Net padrão é alterada com a intenção de desempenhar bons resultados ao segmentar as regiões de interesse.

3.8.5 Mecanismo de Atenção Redes Neurais

A atenção é uma função cognitiva complexa, que é indispensável para o ser humano, e consiste no processo de concentração em alguma informação quando necessário. Este mesmo princípio de atenção pode ser empregado em redes neurais através do mecanismo de atenção. O mecanismo de atenção pode processar informações mais importantes mesmo com recurso computacionais limitados (NIU; ZHONG; YU, 2021).

A aplicação do mecanismo de atenção na visão computacional, é comumente usada por pesquisadores da visão computacional. Mnih et al. (2014) propõe o uso do mecanismo de atenção em redes neurais recorrentes(RNN) para classificar imagens, já Bahdanau, Cho e Bengio (2014) usou a o mecanismo de atenção em tarefas de tradução automática de texto, além disso o mecanismo é inserido cada vez mais em arquitetura de redes neurais para a realização de tarefas, como geração de títulos ou legendas automáticas em imagens, classificação de texto, reconhecimento da ação humana e reconhecimento de um discurso (XU et al., 2015), (LIU; GUO, 2019),(CHOROWSKI et al., 2014), (SONG et al., 2017). A Figura 19 mostra o mecanismo de atenção.

Figura 19 – Mecanismo de atenção



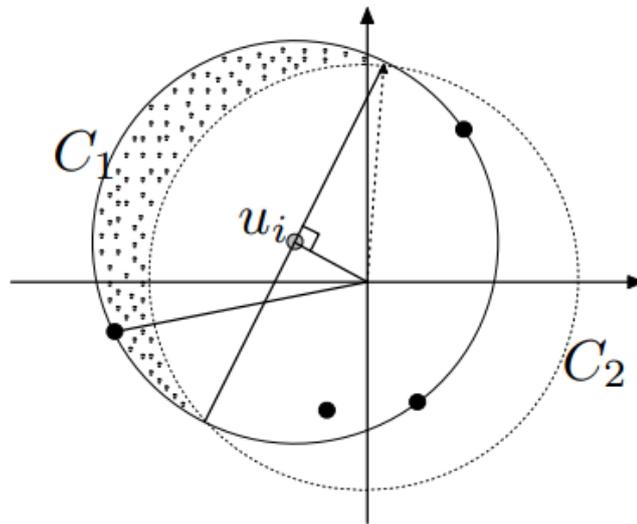
Fonte: Adaptado de (NIU; ZHONG; YU, 2021)

O mecanismo de atenção conforme da Figura 19, consiste em um modelo de atenção unificado definido por Niu, Zhong e Yu (2021), que compreende na parte central compartilhada pela maioria dos modelos de atenção encontrados na literatura pesquisada. Ao calcular a distribuição de atenção, a rede neural primeiro codifica o recurso de dados de origem definida como K onde pode ser expressa em várias representações de acordo com tarefas específicas e arquiteturas neurais. Quando a rede neural calcula vetores de contexto (geração de um vetor de comprimento fixo), geralmente é necessário introduzir uma nova representação para as características dos dados, chamado de V . (NIU; ZHONG; YU, 2021). O mecanismo de atenção é inserida junto a arquitetura de rede neural convolucional para a segmentação da íris e pupilas conforme será apresentado na subseção 4.2.

3.9 Smallest Circle

O *Smallest Circle* é um algoritmo, que busca encontrar o menor círculo baseado em um conjunto de pontos específico. O algoritmo foi proposto em 1987 por Sylvester (1857) consiste em um conjunto P de n pontos é distribuído em um plano 2D, o objetivo é identificar um ponto c (não necessariamente um membro de P) tal que a distância euclidiana dos membros de P de c é mínima entre todas as outras pontos no plano ou seja, o objetivo é relatar o centro do círculo de raio mínimo que pode incluir todos os pontos em P (ROY et al., 2006). A Figura 20 ilustra o *smallest circle*.

Figura 20 – Ilustração do *Smallest Circle*



Fonte: Adaptado de (LIN; FRANCIS; MAGGIORE, 2022)

Na Figura 20, o círculo sólido C_1 é o menor círculo envolvendo os pontos 0 e y_{ij} , $j \in N_i(x)$. O círculo pontilhado C_2 está centrado na origem e passa pelos pontos de interseção entre C_1 e seu diâmetro, que é perpendicular a u_i . Se houver algum y_{ij} no área sombreada fechada, então um deles atinge o distância máxima da origem. Por outro lado, existe pelo menos um $j \in N_i(x)$ tal que y_{ij} está no semicírculo fechado de C_1 , pois caso contrário não é o menor círculo. Portanto, y_{ij} está na área sombreada fechada se $j \in I_i(x)$. Além disso, o ângulo entre u_i e y_{ij} (y_{ij} está no área sombreada) é menor que $\pi/2$. Isso implica que máximo $\max(x_i - y_j)^T u_i \leq 0$ (LIN; FRANCIS; MAGGIORE, 2022).

Segundo Xu, Freund e Sun (2003) o *Smallest Circle* surge para ser aplicado em várias áreas tais como: na construção de ambientes ou locais compartilhados, ou seja, ao construir um hospital é necessário atender um número de comunidades, o hospital deve ficar no centro dessas comunidades de forma que essa região forme um círculo minimizando a distância do hospital ao público; o *Smallest Circle* pode ser usado para resolver problemas

na ciência, no reconhecimento de padrões através da identificação de alguns pontos de referência de objetos ou imagens, na biologia ao analisar proteínas, na computação gráfica e entre outras áreas.

Esse algoritmo foi generalizado ao ser introduzida a noção de problemas do tipo LP (*linear-programming-like*), logo em seguida mostrou propriedades aplicáveis em teoremas de *Helly-Type*. Além disso o algoritmo é objeto de estudo em pesquisas relacionadas a geometria e aplicações computacionais. (FISCHER; GARTNER, 2003)(FISCHER; GÄRTNER; KUTZ, 2003)(GILL; GERSHON; KIM, 2005). O smallest circle é inserido na etapa de pós-processamento do resultado da segmentação da íris e pupila com a finalidade de ajustar essas regiões, tornado-os mais próximos das marcações *Ground Truth*.

3.10 Métricas de Avaliação

Esta seção apresenta as principais métricas usadas para quantificar o desempenho do método proposto com o intuito de avaliar e validar os resultados obtidos. A fim de analisar as regiões segmentadas da íris e pupila geradas a partir do método discutido neste trabalho, as principais métricas selecionadas são: Coeficiente de similaridade de Dice e o desvio padrão.

O coeficiente de similaridade de Dice (DICE), é constantemente discutida para avaliar o processo de segmentação (TAHA; HANBURY, 2015), dessa forma este coeficiente, também pode ser chamada de índice de sobreposição. Em várias tarefas o coeficiente de Dice é selecionado para a avaliação conforme apresentado por Zou et al. (2004), onde, a métrica valida a segmentação de lesões de substâncias brancas em ressonâncias magnéticas. A Equação 3.15 apresenta o coeficiente de similaridade de Dice.

$$DICE = \frac{2 * |X \cap Y|}{|X| + |Y|} \quad (3.15)$$

O índice X e Y representam as máscaras das imagens segmentadas rotuladas manualmente e as regiões segmentadas previstas pelo método proposto, dessa forma o DICE mede a sobreposição espacial entre duas imagens binárias e seus valores que variam entre 0 (sem sobreposição) a 1 (sobreposição perfeita) (TRIYANI et al., 2016). O Rohlfing (2011) afirma que o DICE obtido representa o tamanho da estrutura da imagem, ou seja, quanto menor a região do resultado da segmentação, menor é o DICE de acordo com o tamanho do rótulo original da imagem.

O desvio padrão é um índice de quão próximo os pontos de dados individuais se agrupam em torno da média (HASSANI; GHODSI; HOWELL, 2010). A Equação 3.16 representa o cálculo do desvio padrão, em que n é a quantidade de dados, e M_A é a média aritmética dos dados calculados e x_i é o valor de uma dado na posição i de um conjunto.

$$Dv = \sqrt{\frac{\sum_{i=1}^n (x_i - M_A)^2}{n}} \quad (3.16)$$

No calculo do desvio padrão, a diferença dos valores da média representa a variação dos dados de distribuição do conjunto. Dessa forma, a coleção de dados é espalhada ao redor do valor médio(LEE; IN; LEE, 2015).

3.11 Considerações Finais

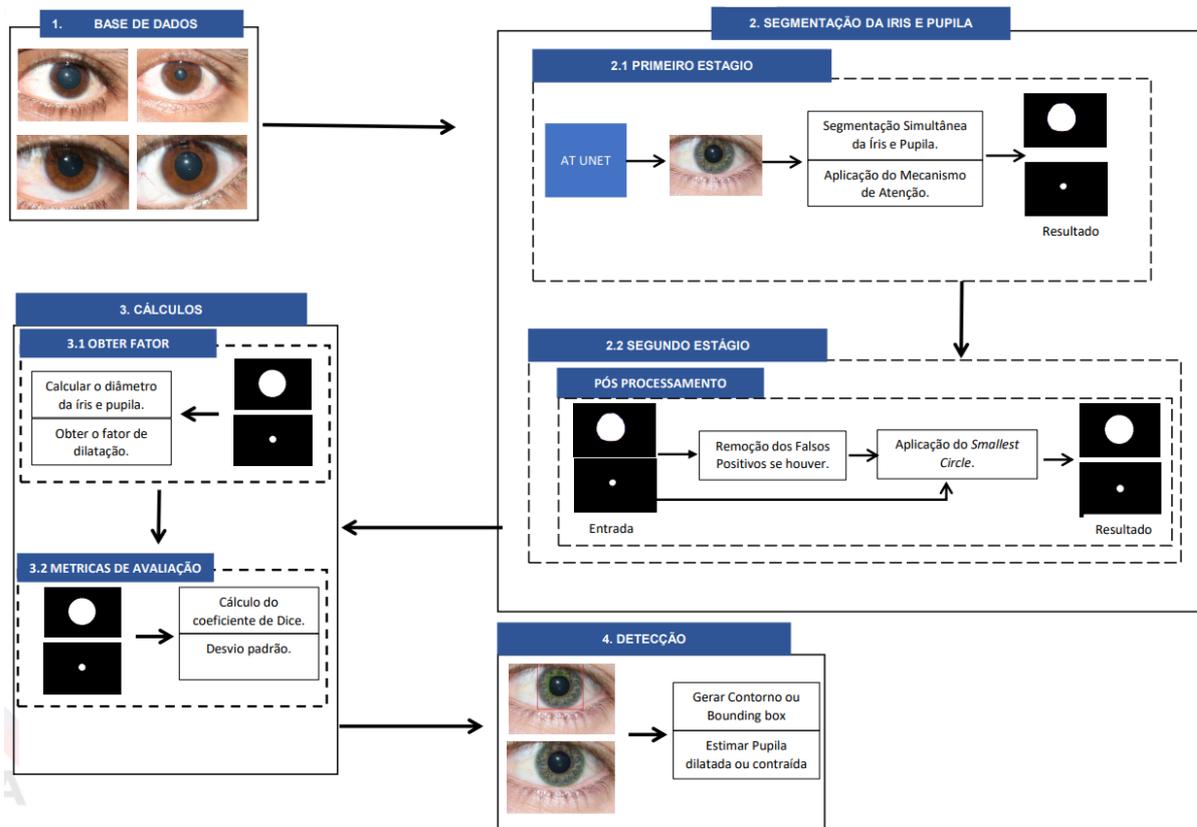
Neste capítulo, foram apresentados os conceitos e fundamentos teóricos imprescindíveis para a compreensão e aplicação do método proposto. Abordaram-se tópicos relacionadas a pupilometria, e as reações pupilares, resultando em potenciais temas para propor métodos computacionais. Além disso são destacados os conceitos relacionados ao aprendizado profundo, técnicas de processamento de imagens, redes neurais e métricas de avaliação.

No próximo capítulo, serão explanados os materiais e métodos utilizados para propor este trabalho, cujo objetivo é detectar e indicar pupilas dilatadas ou contraídas por meio da segmentação de imagens.

4 Metodologia

A metodologia proposta neste trabalho visa a utilização da arquitetura de rede neural U-Net para segmentar e detectar a íris e pupila dilatada em imagens da região dos olhos. Para tanto a proposta original da arquitetura U-Net é modificada, através da substituição do codificador por uma arquitetura de rede neural VGG-16 e o ajuste do decodificador original da rede U-Net, implementando o mecanismo de atenção entre as camadas de convolução transposta e concatenação, dessa forma, a arquitetura modificada denominada de At-UNet segmenta simultaneamente a região da íris e pupila de uma imagem de entrada. As imagens segmentadas são pós-processadas com o algoritmo *Smallest Circle* e ajustadas para calcular os seus diâmetros, a fim de distinguir a pupila dilatada e contraída por meio do cálculo do fator de dilatação e constrição. A Figura 21 apresenta as etapas da metodologia que serão detalhadas nas próximas seções.

Figura 21 – Metodologia Proposta

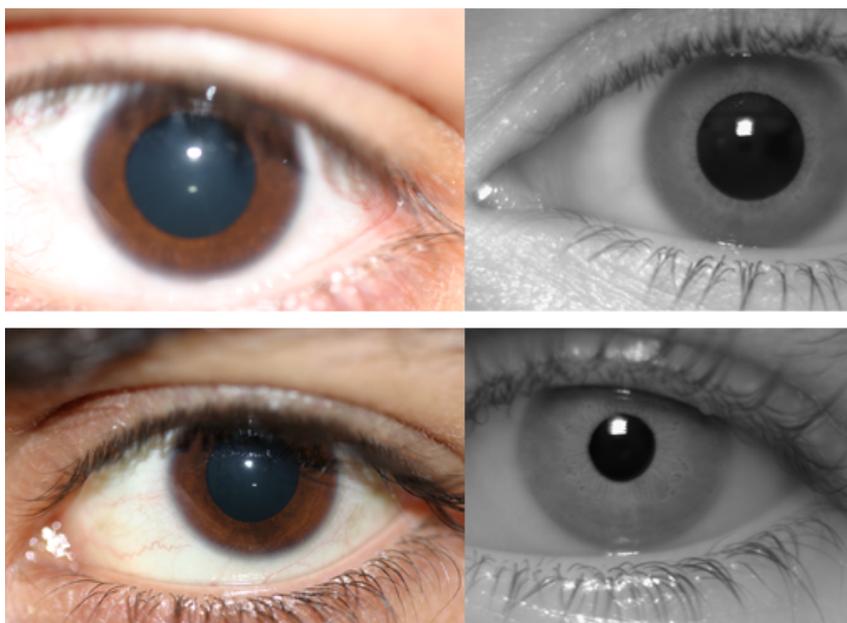


Fonte: Elaborada pelo Autor

4.1 Aquisição da Base de Imagens

O conjunto de dados *UTIRIS database* (HOSSEINI; ARAABI; SOLTANIAN-ZADEH, 2010) da Universidade do Teerã é usada neste trabalho a fim de obter resultados relacionados ao objetivo estabelecido deste trabalho. Este conjunto está disponível em um dos seus repositórios públicos, com a finalidade de promover desenvolvimento científico em processamento de imagens digitais. A Figura 22 apresenta a base de dados UTIRIS.

Figura 22 – UTIRIS Dataset



A base contém o total de 1.540 imagens, correspondendo em 770 em formato RGB e 770 em formato infravermelho da região periocular, composta por cílios, pálpebras, sobrancelhas e a região dos olhos. A universidade realizou duas sessões de aquisição de 79 indivíduos, usando dois dispositivos diferentes. Na primeira sessão o dispositivo consistia em uma câmera *Canon EOS 10D*, com a finalidade de adquirir imagens no formato RGB. Já na segunda sessão a câmera *ISG Lightwise LW* é selecionada para adquirir imagens infravermelhas (HOSSEINI; ARAABI; SOLTANIAN-ZADEH, 2010). São obtidas 5 ou 6 imagens do lado direito e esquerdo de cada região periocular de um indivíduo, totalizado 770 imagens por cada sessão de aquisição.

O conjunto de dados contém imagens de pupilas dilatadas e contraídas. Os dois tipos de imagens tem características diferentes conforme apresentada na Tabela 3. Neste trabalho são consideradas apenas as imagens no formato RGB.

O principal motivo de utilizar essa base de dados, deve-se ao fato de que contém imagens da região dos olhos em RGB e apresentar a pupila em seus estados de dilatação, constrição ou sem algum tipo de reação (ou seja, pupila com diâmetro normal). Uma outra motivação para usar as imagens em RGB, é que ao analisar qualitativamente essas imagens, nota-se que em vários momentos a pupila está oclusa por reflexos espaciares

Tabela 3 – Características da base UTIRIS

Imagem	Dimensão(px)	Formato	Espaço de Cor
Coloridas	2048 x 1360	JPEG	RGB
Infravermelhos	1000 x 776	BMP	GRAY

causadas durante a sessão de aquisição, dessa forma o processo de detecção dessa região torna-se uma tarefa desafiadora.

O conjunto de dados está organizado em diretórios, contendo as imagens coloridas e infravermelho da pupila dilatada ou contraída de um indivíduo. A base não fornece o diâmetro da íris ou da pupila, e nem disponibiliza as máscaras binárias dessas regiões, portanto os dados são rotulados manualmente, com a finalidade de obter o diâmetro da íris e pupila, baseados na marcação das regiões anotadas.

Durante o processo de anotação, em todas as imagens do conjunto de dados, o diâmetro é calculado baseado na Equação 4.1. A relação entre o diâmetro da pupila e o diâmetro da íris também é calculada para identificar o tipo de reação de pupilar. Essa relação é definida neste trabalho como o fator de dilatação, conforme apresentado na subseção 4.4. O fator de dilatação é baseada na equação apresentada por Hasegawa et al. (2012), que define uma medida relativa do quanto a pupila está dilatada. Essa medida relativa estabelecida pelo autor considera os raios da íris e pupila a fim de definir o estado pupilar. Dessa forma é necessário calcular o diâmetro da íris e pupila de maneira que seja possível definir o fator de dilatação.

$$Dm = 2 * \sqrt{A}/\pi \quad (4.1)$$

O Dm é o diâmetro real da marcação (*Ground Truth*) gerada, e A é área do rótulo da imagem, que consiste no total de pixels conectados daquela região com o valor 1 considerando que a imagem é binária. Sendo assim é possível definir o diâmetro da das regiões de interesse.

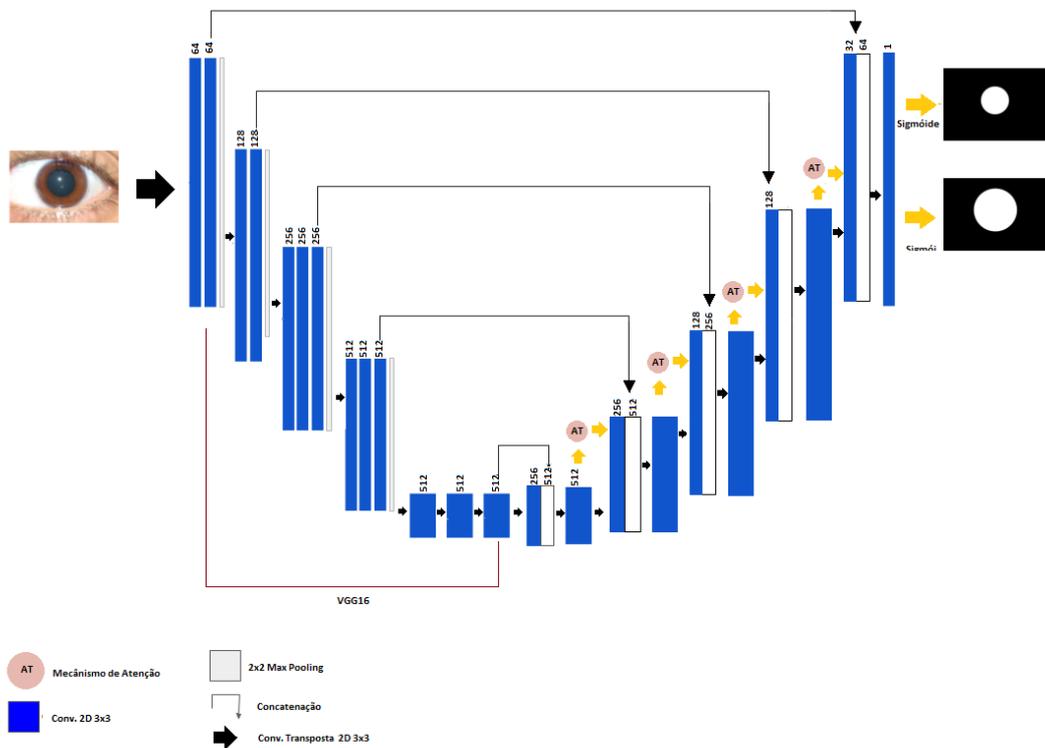
4.2 Segmentação da Íris e Pupila

A segmentação neste trabalho é usada para destacar a região da íris e pupila a fim de obter algumas informações para definição da reação pupilar. Logo esta proposta apresenta a arquitetura de rede neural At-Unet multitarefas baseada na arquitetura de rede neural U-Net (RONNEBERGER; FISCHER; BROX, 2015). Desse modo propõe-se segmentar as regiões de interesse a partir deste método.

A rede neural At-Unet de multitarefa, realiza a segmentação da íris e da pupila da região dos olhos de maneira simultânea, gerando duas imagens segmentadas da região

periocular. Durante todo o processo de segmentação, a arquitetura dispõe de uma operação de decodificação das imagens usando o mecanismo de atenção (OKTAY et al., 2018). A Figura 23 apresenta a arquitetura de rede neural At-Unet multitarefa.

Figura 23 – Arquitetura de Rede Neural Proposta(At-Unet)



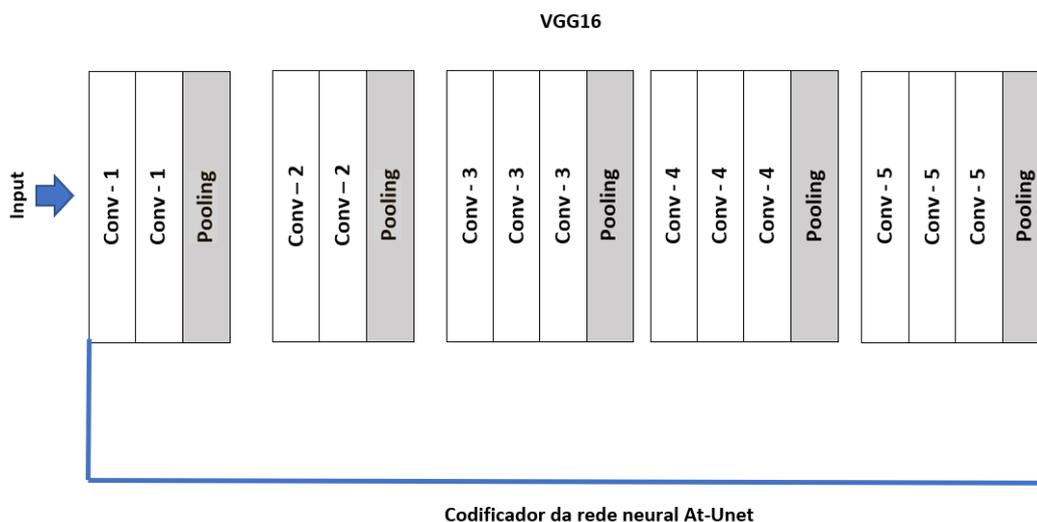
Fonte: Elaborada pelo Autor

A arquitetura At-Unet multitarefa, conforme a Figura 23, também é baseada na rede neural U-Net, para a segmentação de imagens médicas. Esta arquitetura tem dois caminhos para realizar a segmentação de imagens, porém o caminho de codificação, ou codificador, é constituída pela arquitetura de rede neural VGG16 (SIMONYAN; ZISSERMAN, 2015), por apresentar bom desempenho no conjunto de dados ImageNet (DENG et al., 2009). A alteração neste caminho, visa melhorar o desempenho desta rede neural durante a extração de características da imagem. Os pesos pré-treinados do conjunto de dados ImageNet, são carregados no codificador da rede neural proposta.

O codificador da rede neural At-Unet é baseado na VGG16 que consiste em 16 camadas divididas em cinco estágios que contém duas ou três camadas de convolução conforme a Figura 24, com filtros de tamanho 3x3, seguidas de normalização em lote e camadas de ativação Relu. A cada estágio de duas ou três camadas de convoluções, são adicionadas camadas de *pooling* com filtros de tamanho 2x2.

O caminho de decodificação amplia as dimensões da imagem para o seu tamanho original, depois de classificar os pixels, correspondentes à região de interesse. A decodificação

Figura 24 – Codificador da rede At-Unet



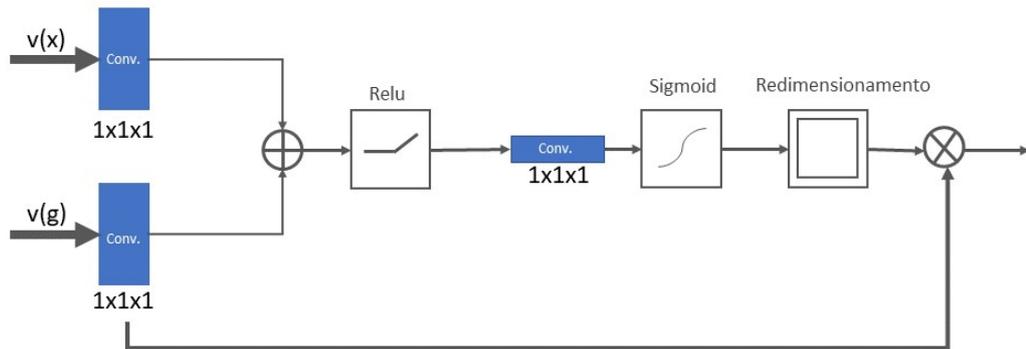
Fonte: Elaborada pelo Autor.

da At-Unet é composta por etapas de *upsampling* para a expansão das dimensões dos mapas de características a fim da rede neural melhorar o aprendizado durante a expansão dos mapas de características concatenados. Isto é feito por meio das conexões de saltos que concatenam as características extraídas do codificador com o decodificador.

Os mecanismos de atenção são aplicados durante a expansão dos mapas de característica. Este mecanismo está posicionado antes da concatenação dos mapas de características do decodificar e do codificador, entre as camadas de convolução transpostas conforme apresentado na Figura 23. O objetivo do mecanismo de atenção é melhorar o poder de generalização da rede neural, dando uma ênfase maior a região de interesse durante o *upsampling*.

Antes de seguir o fluxo da rede neural, o mecanismo de atenção recebe dois vetores de características que irão passar por camadas de convolução 1x1 individuais, de modo que os dois vetores tenham as mesmas dimensões. Após essa operação os dois vetores são somados, elemento por elemento, e encaminhados a uma camada de ativação Relu por mais uma camada de convolução 1x1. O vetor resultante passa pela função de ativação sigmoide, produzindo coeficientes de atenção (pesos) e dando importância aos valores mais próximos. Por fim, os coeficientes de atenção são ampliados para a dimensão do vetor original, usando uma interpolação trilinear (OKTAY et al., 2018). Esses coeficientes são multiplicados com o vetor original, elemento por elemento, escalando o vetor de acordo com os valores de importância definidas. Sendo assim, o vetor resultante é transmitido normalmente na conexão de salto da rede neural At-Unet de multitarefas. A Figura 25 apresenta o mecanismo de atenção usado na rede neural.

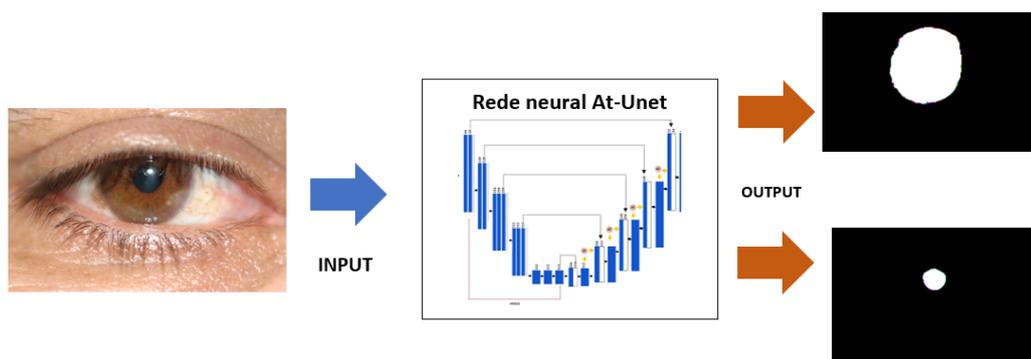
Figura 25 – Mecanismo de Atenção da At-Unet



Fonte: Elaborada pelo Autor.

Em seguida, são aplicadas operações de convolução regulares, camadas de ativação Relu e normalização em lote. No estágio final da decodificação são aplicados 32 filtros em uma sequência de camadas de convolução 3×3 , seguidas de camadas Relu e normalização em lote, aplicando-se ao final uma camada de convolução 1×1 com 3 filtros, seguido da função de ativação sigmoide. O resultado final desta rede neural multitarefas é a previsão da região da íris e da pupila simultaneamente conforme o exemplo da Figura 26. Após a essa etapa de segmentação, aplica-se o pós-processamento nas imagens de saída da rede At-Unet a fim de melhorar os resultados de segmentação.

Figura 26 – Exemplo do processo de segmentação com a rede neural At-Unet



Fonte: Elaborada pelo Autor

4.3 Pós-Processamento da Segmentação

Foi verificada a necessidade de incluir uma etapa de pós-processamento após a segmentação da íris e pupila. Esta etapa visa tratar a imagem segmentada a fim de remover alguns artefatos que se encontram nas regiões segmentadas e realizar os devidos ajustes para aproximar o resultado final da segmentação ao rótulo original. O pós-processamento na tarefa de segmentação é fundamental para que esta metodologia atinja bons resultados, portanto, esta etapa é considerada crucial em métodos de segmentação de imagens (DHAWAN; BODANI; GARG, 2019; ROSS-HOWE; TIZHOOSH, 2018; PALUS, 2007). A Figura 27 exemplifica alguns desses artefatos.

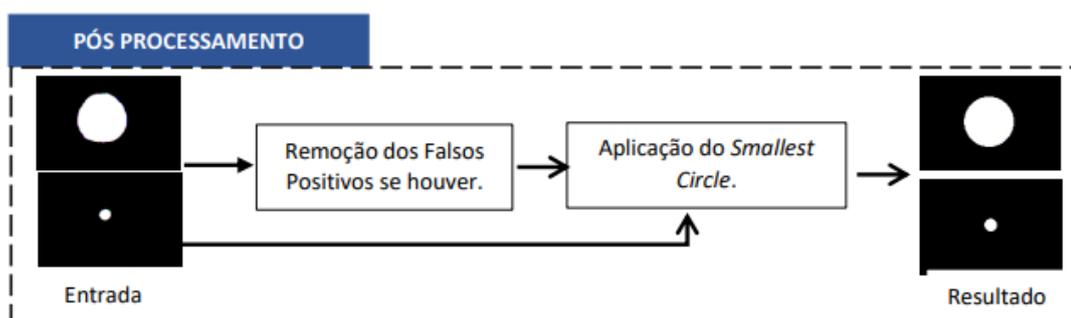
Figura 27 – Imagens segmentadas com áreas que não fazem parte da região da íris.



Fonte: Elaborada pelo Autor

A etapa de pós-processamento consiste em duas subetapas a fim de realizar os ajustes necessários no resultado de segmentação da íris e pupila. Na Figura 28 são destacadas as duas etapas aplicadas no resultado da segmentação. A primeira subetapa consiste na remoção das áreas indesejadas (falsos positivos) que podem surgir quando a região dos olhos é segmentada conforme a Figura 27. Durante os experimentos notou-se que apenas as imagens segmentadas da região da íris apresentaram esses falsos positivos, desse modo são removidas as regiões segmentadas com menor área dentre todas as regiões em uma mesma imagem. Este ação parte do princípio de que a íris é sempre a maior região segmentada na imagem.

Figura 28 – Ilustração das subetapas do pós-processamento.



Fonte: Elaborada pelo autor.

Calcula-se a área total de cada região segmentada, distinguindo quais das regiões são menores em relação a área maior (a íris). Logo, nas menores áreas são aplicadas algumas operações morfológicas, a primeira consiste na erosão dos pequenos falsos positivos com o elemento estruturante de tamanho 5x5 no formato elíptico/circular realizando apenas uma interação sob toda a imagem. Em seguida é aplicada a operação de dilatação com o elemento estruturante de tamanho 5x5 com o formato elíptico/circular com duas interação sob a imagem após remover as regiões de interesse. Após todas essas operações verifica-se novamente a existência dos falsos positivos, para que aconteça todo o processo novamente de acordo com as configurações já citadas.

Após esta etapa de remoção, o algoritmo *Smallest Circle* é aplicado com o propósito de melhorar o delineamento da circunferência que representa a íris e a pupila. O algoritmo busca resolver um problema conhecido como problema dos círculos mínimos. Ele busca encontrar o menor círculo possível $D = [(x, y), r]$ em um conjunto de pontos de um plano euclidiano, onde x e y são as coordenadas do ponto e r é o raio do círculo a ser criado. A proposta é gerar um círculo perfeito baseado nos pontos de pixels classificados corretamente como pertencendo à íris e à pupila na saída da segmentação, corrigindo o formato dessa regiões segmentadas que podem estar deformadas, devido às oclusões causadas pelos cílios e pálpebras da região dos olhos. A Figura 29 apresenta exemplos da aplicação do algoritmo nas imagens segmentadas.

4.4 Fator de dilatação Pupilar

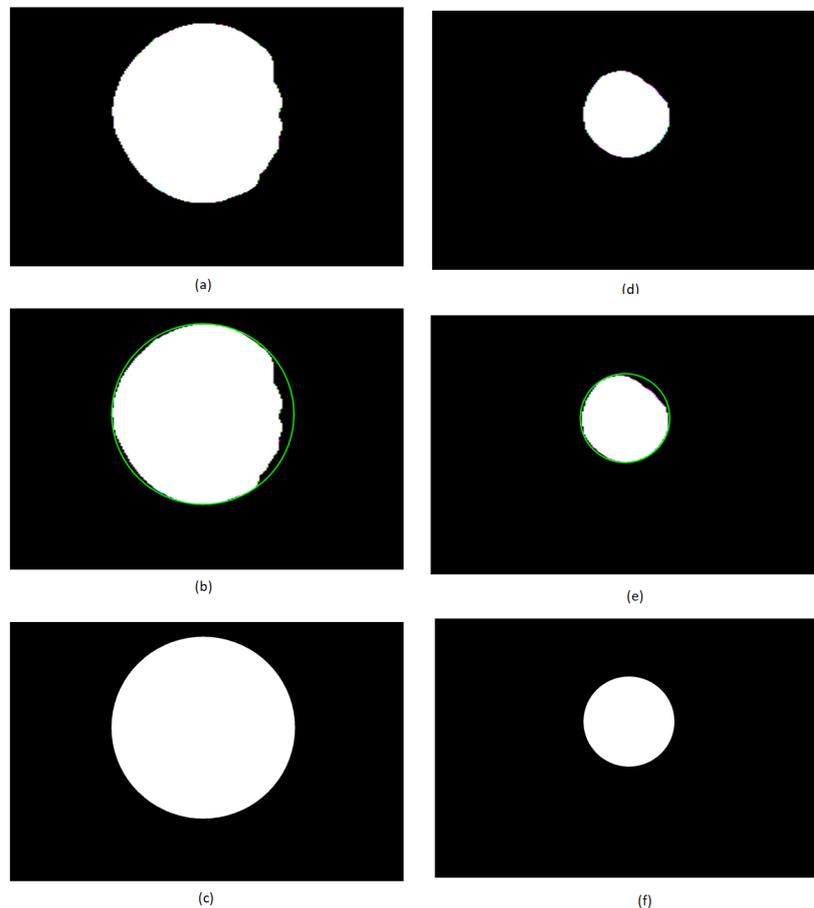
A etapa de segmentação é essencial para a identificação de pupilas dilatadas ou contraídas em imagens da região dos olhos. Conforme apresentado na subseção 4.1, o conjunto de dados foi rotulado manualmente gerando as marcações da íris e da pupila de cada imagem. Para cada imagem rotulada foi calculado o diâmetro da íris e da pupila a fim de distinguir o tamanho de cada região, e assim definir o fator de dilatação.

O fator de dilatação proposto nesta metodologia é a relação entre o diâmetro da pupila e o diâmetro da íris conforme a Equação 4.2. Dessa forma ela é baseada na equação apresentada por Hasegawa et al. (2012). Diferentemente da proposta de Hasegawa et al. (2012), são consideradas neste trabalho, apenas os diâmetros calculados a fim de definir o intervalo entre 0 e 1, e indicar o grau de dilatação da pupila. Considera-se que quanto mais próximo de 1, maior a dilatação pupilar. Esse fato foi observado após analisar cada imagem do conjunto de dados, considerando o fator de dilatação estimado.

$$Ft = \frac{D_p}{D_i} \quad (4.2)$$

Ao concluir a análise propõe-se a definição dos intervalos do grau de dilatação, conforme o conjunto de dados apresentado, dessa forma o grau de dilatação é dividido

Figura 29 – Regiões Segmentadas da íris e pupila aplicando o algoritmo *Smallest Circle*.(a) íris segmentada. (b) aplicação do algoritmo na imagem segmentada. (c) Imagem resultante da íris após a aplicação do algoritmo. (d) Pupila segmentada. (e) Aplicação do algoritmo na imagem segmentada. (f) Resultado Final da pupila segmentada após a aplicação do algoritmo



Fonte: Elaborada pelo Autor.

em três: Pupila contraída; Pupila de tamanho normal; Pupila dilatada. Cada grau de dilatação identificado é baseado no fator de dilatação calculado. A Tabela 4 apresenta os intervalos do fator de dilatação para cada grau identificado.

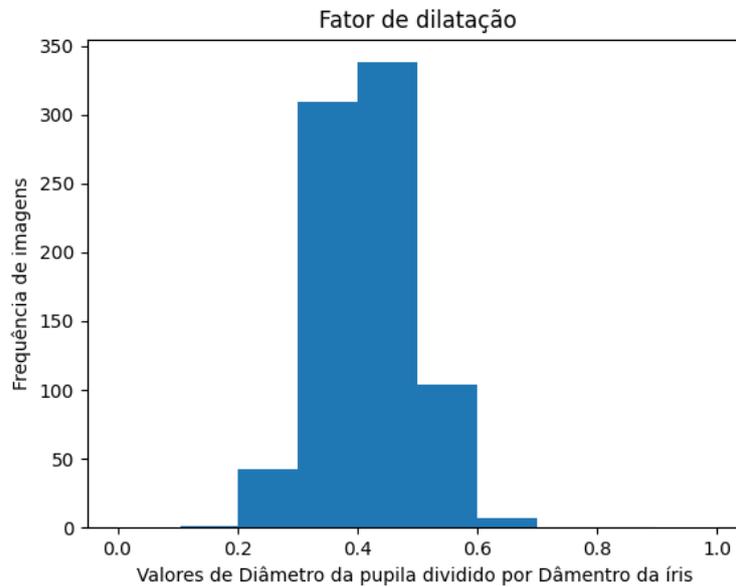
Tabela 4 – Dilatação Pupilar

Intervalo do Fator de dilatação	Grau de dilatação
0,0 - 0,29	Pupila contraída
0,30 - 0,39	Pupila de tamanho normal
0,4 - 1,0	Pupila dilatada

Os resultados desta pesquisa são baseados no fator de dilatação, a fim de identificar as pupilas dilatadas ou contraídas no conjunto de dados. Sendo assim, o foco é detectar as pupilas com um intervalo entre 0.0 e 0.29 e 0.4 e 1.0. O gráfico da Figura 30 apresenta a

distribuição do conjunto de dados entre os intervalos do fator de dilatação.

Figura 30 – Intervalos do fator de dilatação no conjunto de dados



Fonte: Elaborada pelo Autor

O gráfico aponta que na maioria das imagens do conjunto de dados, o fator de dilatação é acima de 0.4, indicando que a maior parte das imagens da base tem a pupila dilatada, conforme a divisão proposta neste trabalho. A Figura 31 apresenta o fator de dilatação de algumas imagens usadas durante os experimentos.

Figura 31 – Imagens do conjunto de dados com o fator de dilatação



Fonte: Elaborada pelo Autor

O fator de dilatação, conforme apresentada na Figura 31, distingue as imagens com pupilas dilatadas, contraídas e de tamanho normal, sendo assim esse fator é considerado essencial para detectar pupilas dilatadas. Na próxima seção serão demonstrados os experimentos e resultados alcançados a partir da metodologia proposta.

5 Resultados

O propósito deste capítulo é descrever os experimentos realizados no processo de segmentação da íris e pupila para estimar o tipo de reação pupilar. Para validar o desempenho desta metodologia todos os experimentos foram realizados usando a base de dados apresentada na seção 4.1. Dessa forma, o conjunto de dados UTIRIS é ajustado a fim de dividir os dados aleatoriamente em treinamento, validação e teste. Desse modo são selecionadas as 770 imagens em formato RGB conforme apresentado na Subseção 4.1, além disso neste conjunto são adicionadas os rótulos gerados manualmente da íris e pupila, totalizado 2.310 imagens, dessa maneira são, 637 (com 1274 rótulos) para treinamento, 70 (com 140 rótulos) para validação e 63 (com 126 rótulos) para teste do modelo gerado.

5.1 Segmentação da Iris e Pupila

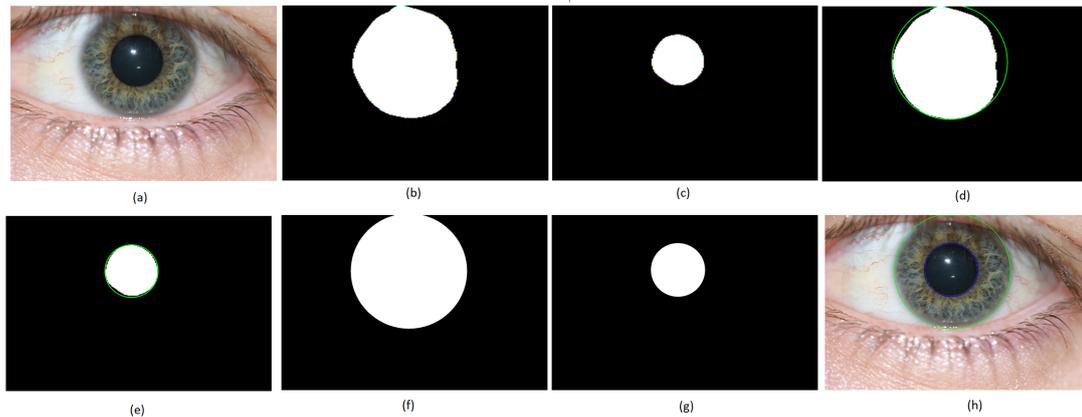
Esta seção descreve os resultados alcançados na etapa de segmentação da íris e pupila, a partir da arquitetura de rede neural proposta, desse modo, os diâmetros dessas regiões de interesse são calculados com o propósito de obter o fator de dilatação e assim verificar o quão próximo os resultados adquiridos estão dos dados rotulados.

Dessa forma, as configurações estabelecidas para o treinamento da rede At-Unet de multitarefa consistem no uso do método de otimização Adam, com a taxa de aprendizagem de 0.00001, o *decay* de 0.000001 e o tamanho do *batch* igual a três. Além disso a função de perda de entropia cruzada é selecionada para verificar o desempenho da rede neural durante o treinamento em 100 épocas. Durante o processo de treinamento o *Data Augmentation* é configurado na arquitetura de rede neural, com a finalidade de fazer o aumento de dados e assim evitar o *overfitting*, sendo assim os dados são rotacionados em 90° , deslocados verticalmente e horizontalmente em até 5%, são aplicados *flips* horizontais e verticais e um zoom de 30% em cada imagem. No decorrer do treinamento nota-se que a função de *loss* convergiu para o mínimo até a época 25, a partir de então a perda de entropia cruzada deixa de minimizar constantemente durante 10 épocas, encerrando automaticamente o treinamento da rede neural antes de alcançar a época 100.

A Figura 32 apresenta todo o processo de segmentação da íris e pupila de uma imagem de entrada na rede neural, logo é possível visualizar todas as etapas principais até alcançar a detecção das regiões.

A imagem (a) da Figura 32 passa pelo processo de segmentação gerando duas saídas simultaneamente conforme a imagem (b) e (c). Dessa forma a rede neural proposta gera dois resultados de segmentação, indicando que uma imagem representa a região

Figura 32 – Processo de Segmentação da Íris e Pupila. (a) Imagem de entrada para segmentação. (b) Resultado da segmentação da íris. (c) Resultado da segmentação da pupila. (d) Pós-processamento do resultado de segmentação da íris. (e) Pós-processamento do resultado da segmentação da pupila. (f) Imagem segmentada da íris pós-processada. (g) Imagem segmentada da pupila pós-processada. (h) Íris e pupila detectada



da íris segmentada e outra aponta a região segmentada da pupila. Essas imagens são pós-processadas conforme apresentado na Seção 4.3, sendo assim as imagens (d) e (e) representam os ajustes realizados durante essa etapa incluindo a aplicação do algoritmo *smallest circle*. Logo em seguida o resultado final após o processo de pós-processamento dessas regiões são representadas nas imagens (f) e (g), possibilitando a obtenção dos diâmetros necessários para definir o tipo de reação pupilar. A imagem (h) configura a região da íris e pupila detectada após todo o processo de segmentação, gerando uma borda em volta das regiões de interesse.

O coeficiente de Dice foi selecionado como métrica de validação de desempenho da segmentação. Essa métrica mede a sobreposição espacial entre duas imagens binárias, e dessa forma avalia o desempenho da segmentação e seus valores que variam de 0 (sem sobreposição) a 1 (sobreposição perfeita) (TRIYANI et al., 2016). Todas as imagens de testes são diferentes das imagens de treinamento da rede neural, mas fazem parte da mesma base de dados.

A Tabela 5 apresenta os resultados da segmentação da íris e pupila, indicando a média de coeficiente de Dice por imagem de teste durante o experimento realizado, além de demonstrar o desvio padrão desta medida de modo a compreender a contribuição do pós-processamento com *Smallest Circle*. Um experimento de ablação é realizado, removendo a etapa de ajuste da íris e pupila segmentada com o algoritmo *Smallest Circle*, a fim de verificar o desempenho da segmentação com e sem a aplicação dessa etapa.

O coeficiente de Dice é calculado por imagem para a avaliação qualitativa e observação do comportamento na segmentação da íris e pupila. Em seguida, calcula-se a média de coeficiente de Dice na base de teste, para obter um resultado global do método

Tabela 5 – Resultados da segmentação usando a Rede Neural de Multitarefa At-Unet

Metodologia	Íris		Pupila	
	Dice(%)	Desvio Padrão(%)	Dice (%)	Desvio Padrão(%)
Multitarefa At-Unet + Módulo Atenção <i>sem Smallest Circle</i>	97.00	0.0162	95.06	0.0232
Multitarefa At-Unet + Módulo de Atenção <i>com Smallest Circle</i>	97.35	0.0139	96.16	0.0198

proposto. Dessa maneira os resultados iniciais são apresentados conforme a Tabela 5, notando que o desempenho da segmentação da íris e pupila apresenta uma diferença ao aplicar o algoritmo *Smallest Circle* durante a etapa de pós-processamento.

A média de coeficiente de Dice mínimo em imagens de testes ao aplicar o algoritmo *Smallest Circle* é de 97.35% ao segmentar íris, e de 96.16% ao segmentar a pupila. Sem a aplicação do algoritmo o Dice médio em imagens de teste por base diminui em pelo menos 0.35% para íris e de 1.1% para a pupila, apontando que a etapa de pós-processamento tem um impacto maior no resultado da segmentação da região da pupila. Desse modo esta etapa é importante para o bom desempenho do resultado da segmentação das regiões de interesse.

Um fato a destacar é que o desvio padrão diminui a cada experimento, conforme a Tabela 5, indicando também que quanto menor o desvio padrão mais uniforme está o coeficiente de Dice por imagem de acordo com a média obtida, desse modo é caracterizado uma boa performance no processo de segmentação. Durante os experimentos observou-se que algumas imagens segmentadas da região da íris apresentavam alguns falsos positivos, conforme citado na Seção 4.3, dessa forma em análise qualitativa dessas imagens, verificou-se que essas áreas representavam alguma parte da região dos olhos, como sobrancelhas ou cílios.

5.2 Cálculo do Fator de Dilatação

Durante todo o processo de segmentação são avaliados aspectos da dilatação pupilar conforme a apresentado na Seção 4.4. O diâmetro da íris e pupila, são calculados para encontrar o fator de dilatação utilizado a fim de distinguir a pupila dilatada e não dilatada.

Neste experimento, 10% da base UTIRIS foi utilizada para teste conforme apresentado nesta seção. A Tabela 6 apresenta os resultados de algumas imagens de teste para avaliar o fator de dilatação calculado durante a execução deste experimento.

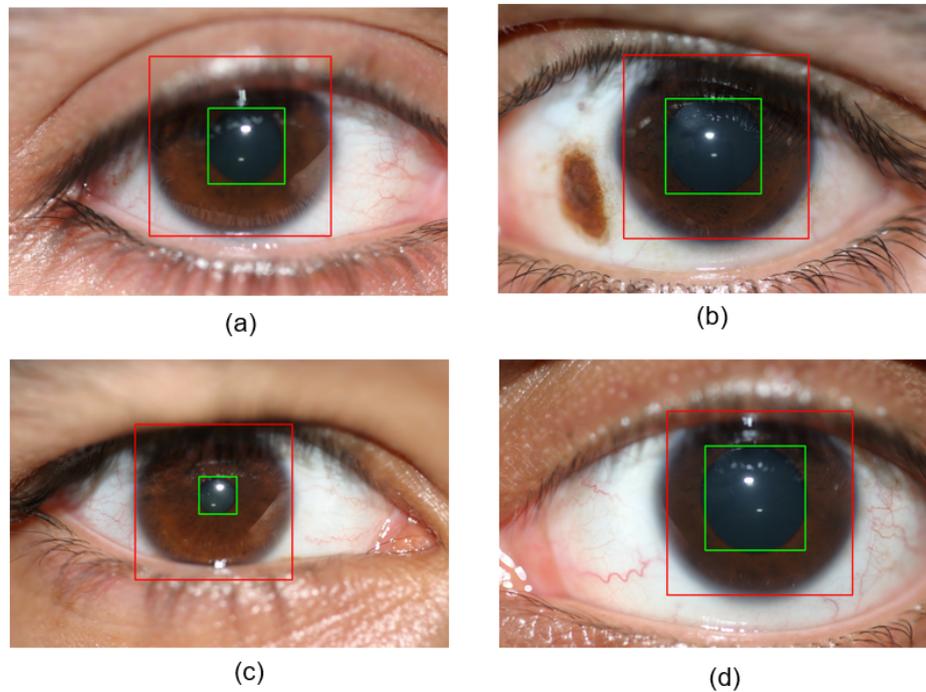
Tabela 6 – Resultados do cálculo do fator de dilatação

Imagem	Fator de dilatação (0-1)	Erro do Fator	Tipo de reação estimada	Fator Real
1	0.4332	0.0046	Dilatada	0.4378
2	0.3758	0.0111	Normal	0.3819
3	0.2298	0.0086	Contraída	0.2212
4	0.3347	0.0011	Normal	0.3358
5	0.229	0.0086	Contraída	0.2206
6	0.3812	0.0321	Normal	0.4191
7	0.3997	0.0383	Normal	0.4311
8	0.4041	0.0012	Dilatada	0.4053
9	0.4121	0.0102	Dilatada	0.4223
10	0.2982	0.0564	Dilatada	0.2821
11	0.4562	0.0068	Dilatada	0.4494
12	0.3938	0.0220	Normal	0.4158
13	0.4927	0.0162	Dilatada	0.5089
14	0.4959	0.0568	Dilatada	0.5527
15	0.2554	0.0008	Contraída	0.2546

O erro é calculado com a finalidade de indicar o quão próximo o fator de dilatação calculado (Subseção 4.4) está do fator real das imagens de teste, logo, este erro é adquirido estimando a diferença entre esses fatores. Destaca-se que este valor está entre 0.0011 e 0.0564, indicando que, a diferença obtida entre o fator real de dilatação e o fator de dilatação calculada baseado no resultado da segmentação predito, são mínimos. As imagens **2, 6, 7, 10 12 e 14**, apresentam uma margem de erro maior em relação as outras imagens de teste, já as imagens **1, 3, 8 e 15** apontam que margem de erro é mínima indicando um excelente desempenho durante todos os processos discutidos para a segmentação e identificação da reação pupilar.

As imagens **1, 7, 8, 9, 11 e 12** da tabela, são de pupilas dilatadas, logo, apenas as imagens **6, 7 e 12** indicam incorretamente o tipo de reação pupilar baseada no fator calculado. Já as imagens **3, 5, 10 e 15** são de pupilas contraídas, mas apenas na imagem **10** o fator de dilatação definiu incorretamente o tipo de reação pupilar. As imagens **2 e 4** são de pupilas classificadas sem reação pupilar, ou seja, em estado normal. Verifica-se então que, em apenas quatro imagens o tipo de estado pupilar foi definido incorretamente. O fator de dilatação é calculado para todo o conjunto de teste da base de dados. A média de erro do calculo é de 0.0167, destacando assim a eficiência do resultado da segmentação para indicar o tipo de reação pupilar. Na Figura 33 apresenta as imagens classificadas incorretamente de acordo com o fator de dilatação calculado conforme citado anteriormente.

Figura 33 – Imagens de teste classificadas incorretamente o tipo de reação pupilar de acordo com o fator de dilatação calculado. (a) Imagem n^o 6 de teste com pupila dilatada. (b) Imagem n^o 7 de teste com pupila dilatada. (c) Imagem n^o 10 com pupila contraída. (d) Imagem n^o 12 de teste com pupila dilatada.



5.3 Validação Cruzada

A validação cruzada é usada para avaliar os resultados da segmentação da íris e pupila por todo o conjunto de dados. Sendo assim, a base UTIRIS é subdividida em 5 grupos ou seja $k = 5$. A diferença deste experimento está fundamentada no princípio de que serão realizadas cinco execuções de treinamento e teste, considerando cada grupo do conjunto de dados formado. Os resultados da validação cruzada, a fim de avaliar a rede Multitarefa At-Unet, são apresentados na Tabela 7.

Tabela 7 – Resultados da Validação Cruzada da rede At-Unet de multitarefas

Grupos de treinamento (k=5)	Conjunto de Dados	Íris(%)		Pupila (%)	
		Dice	Desvio Padrão	Dice	Desvio Padrão
k1	UTIRIS	96.93 %	0.0207	95.79 %	0.0376
k2	UTIRIS	96.98 %	0.0189	96,25%	0.0227
k3	UTIRIS	97.69 %	0.0107	96.62%	0.0214
k4	UTIRIS	97.11 %	0.0169	95.18%	0.0510
k5	UTIRIS	97.41 %	0.0165	97.17%	0.0197

Para realizar a validação cruzada de um conjunto de dados k é necessário dividi-lo em grupos, portanto neste trabalho a base UTIRIS é dividida em cinco partes, definida por $k = 5$. Cada grupo está indicado pela nomenclatura k1, k2, k3, k4 e k5, logo cada subconjunto é dividido em treinamento, validação e teste, sendo assim os dados são distribuídos aleatoriamente para cada k .

Em cada k um treinamento e teste é realizado, dessa forma são cinco execuções durante a validação cruzada. Portanto, a média de Dice da validação cruzada em cada k é de 97,22% para segmentação da íris e 96.20% na segmentação da pupila.

O pior desempenho durante este experimento na tarefa de segmentação segundo a Tabela 7, é identificado no grupo $k1$ e $k2$, onde a média de coeficiente de Dice da íris e da pupila do grupo $k1$ é menor em relação aos outros grupos do conjunto de dados, e a média de coeficiente de Dice do grupo $k4$ em relação a segmentação da pupila é menor em relação aos outros resultados adquiridos em cada grupo. No entanto, o desempenho do At-Unet é satisfatória durante a execução da tarefa.

5.4 Comparação com outros Trabalhos

De modo a situar a proposta no estado da arte atual, os resultados são comparados com trabalhos que realizam a segmentação da região dos olhos, em específico a íris nas bases de dados UBIRIS.v2 (PROENCA et al., 2010) e MICHE-I (MARSICO et al., 2015). A U-Net, FCDNN (HU et al., 2020), Refinanet (HOFBAUER; JALILIAN; UHL, 2019) e FCEDNS (JALILIAN; UHL, 2017) apresentam métodos de segmentação distintos, selecionados para avaliar o desempenho desta metodologia levando em consideração os dados de teste dessas bases.

Esses conjuntos de dados públicos contêm imagens da região dos olhos em quantidades diferentes e disponibilizam as marcações *Ground Truth* da íris e da pupila. As pupilas dessas imagens não estão dilatadas, sendo assim não será necessário calcular o fator de dilatação dessas imagens, o foco deste experimento é avaliar o desempenho da rede de multitarefas At-Unet no processo de segmentação para validar os resultados e comparar com outros trabalhos da literatura. A base MICHE-I contém 680 imagens de tamanhos distintos, que variam entre 400x297 pixels a 400x400 pixels. Já a base UBIRIS.v2 contém 483 imagens com o tamanho fixo de 400x300 pixels. Portanto estes conjuntos são divididos conforme o primeiro experimento realizado na base de dados UTIRIS.

Em comparação com resultados de outros trabalhos apresentados na literatura, no qual realiza a segmentação da região periocular conforme apresentado na Tabela 8, a metodologia apresenta um bom desempenho. Assim como os resultados iniciais da Tabela 5 na base UTIRIS, calcula-se o coeficiente de Dice médio por base de teste disponibilizada pelos autores a fim de validar a metodologia. O coeficiente de Dice médio deste trabalho testado na base MICHE-I é 95.32% e na base UBIRIS.v2 é de 94.14%, aproximadamente 3.91% maior que a proposta do CNNHT. Ao comparar com o FCDNN identifica-se que os resultados estão bem próximos 0.24% maior na média de coeficiente de Dice.

A segmentação baseada apenas na rede U-Net conforme apresentada na Tabela 8, expõe resultados abaixo do esperado comparado a metodologia deste trabalho. Dessa

Tabela 8 – Comparação de diferentes abordagens para a segmentação da região periocular.

Método	Database	Coefficiente de Dice (%)
(RONNEBERGER; FISCHER; BROX, 2015)	MICHE-I	90.25
(U-Net)	UBIRIS.v2	91.77
(HOFBAUER; JALILIAN; UHL, 2019)	MICHE-I	91.41
(Refinanet)	UBIRIS.v2	90.34
(JALILIAN; UHL, 2017)	MICHE-I	-
(FCEDNS)	UBIRIS.v2	91.82
(HU et al., 2020)	MICHE-I	-
(FCDNN)	UBIRIS.v2	93.90
At-Unet	MICHI-I	95.32
(Método proposto)	UBIRIS.v2	94.14

maneira a inclusão de novas parâmetros e técnicas dentro da arquitetura da rede neural U-Net promove um desempenho satisfatório na tarefa de segmentação.

5.5 Hardware e Softwares Utilizados

A metodologia apresentada neste capítulo foi desenvolvida utilizando a linguagem de programação Python, na versão 3.8 (ROSSUM; JR, 1995). Esta linguagem foi escolhida por oferecer grande suporte ao desenvolvimento de sistemas baseados em Aprendizado de Máquina, e ao processamento e análise de dados científicos. Dentre as principais bibliotecas em Python usadas no desenvolvimento desta pesquisa estão: Keras (CHOLLET et al., 2015), na versão 2.4.3 e Tensorflow na versão 2.3.1 (ABADI et al., 2016), ambas usadas para o treinamento de modelos de aprendizagem profunda. O hardware utilizado para o desenvolvimento de boa parte da metodologia conta com sistema operacional windows 10, com processador Intel Core i7-7700, 16GB de memória RAM e 1TB de HD e 500 GB de SSD, assim como para o treinamento da CNN, utilizou-se a GPU NVIDIA GeForce GTX 1660 Ti.

6 Conclusão

A pupilometria é uma técnica comumente usada por médicos e neurocientistas para compreender as reações pupilares de dilatação e contração. Esta técnica é usada a fim de avaliar alguns aspectos cognitivos que podem ser identificados mediante a essas reações que podem acontecer devido aos reflexos luminosos ou fatores internos relacionados ao sistema nervoso central. Além disso, a pupilometria é usada em alguns exames clínicos pré-operatórios além de auxiliar no diagnóstico de algum tipo de doença neurológica ou ocular.

Dessa forma, é necessário introduzir equipamentos capazes de realizar esse processo de detecção e a medição de pupilas dilatadas ou contraídas mediante ao alto custo para obter um pupilômetro automatizado. Além disso, ao realizar a pupilometria de maneira manual apenas com um objeto que reflete a luz aos olhos não é possível obter informações precisas, como o diâmetro da pupila, a latência da dilatação ou contração e a definição do estado final da reação pupilar.

Diante disso este trabalho propôs um método computacional para avaliar as repostas pupilares dos indivíduos, a fim de auxiliar os médicos e psicólogos cognitivos a medir a dilatação ou contração final das pupilas dos olhos de um determinado paciente, possibilitando a análise desses resultados e, dessa forma, tornar possível o entendimento das implicações dessas reações pupilares.

O método apresentado é baseado em aprendizado profundo devido à possibilidade de todo processo ser realizado de maneira automatizada, tornando uma ferramenta precisa e de baixo custo. Portanto, ao longo do trabalho são apresentadas algumas etapas para que seja possível obter informações da pupila. A proposta apresenta e discute todo o processo de segmentação para que aconteça a detecção da pupila e obtenha dados relacionados a dilatação e contração pupilar.

As principais etapas da proposta são: segmentar a pupila e a íris por meio da arquitetura de rede neural denominada de At-Unet de Multitarefa com mecanismo de atenção, possibilitando a segmentação da íris e pupila de maneira simultânea; além disso, após o resultado da segmentação foi necessário aplicar a etapa de pós-processamento para remoção de algumas regiões que não fazem parte da íris ou pupila dos olhos, e dessa forma ajustar as regiões segmentadas a partir da aplicação do algoritmo *Smallest Circle*, para que essas regiões fiquem mais próximas das marcações geradas; e por fim a última etapa consiste em obter algumas informações dessas regiões como o diâmetro da íris, pupila e o fator de dilatação que determina qual o tipo de reação pupilar é detectada.

Sendo estabelecidas estas etapas, vários experimentos são realizados a fim de

verificar o desempenho desse processo. O resultado da segmentação da íris e pupila considerando todas as etapas apresentadas, é de 97,35% e 96,16%, respectivamente, esses números são obtidos a partir das 90 imagens da base de teste do conjunto de dados UTIRIS. Os resultados são avaliados a partir do coeficiente de Dice e o desvio padrão do resultado da segmentação. Destaca-se que a etapa de pós-processamento realizada com o algoritmo *Smallest Circle* contribui para um incremento dos resultados do método. Além disso, a rede At-Unet colabora para o bom desempenho na tarefa de segmentação.

Destaca-se que o erro calculado do fator de dilatação predito durante todos os experimentos, apresenta uma média 0.0167 de diferença em relação ao fator de dilatação real calculado de acordo as marcações geradas no conjunto de dados. Este fator é o que determina o estado final da pupila, indicando se está dilatada ou contraída. Além disso calcula-se os diâmetros de cada estado pupilar identificado. Portanto, os resultados para essa tarefa são promissores visto que a implementação da rede neural contribui para futuras pesquisas relacionadas a avaliação das reações pupilares.

6.1 Trabalhos Futuros

Como sugestão de trabalhos futuros indicamos a utilização dos resultados fornecidos por este método para segmentação da pupila possibilitando identificar reações pupilares para avaliar os estados cognitivos de uma pessoa a ponto de diagnosticar algum distúrbio neurológico. Desse modo sugere-se também:

- Adquirir um conjunto de imagens de vídeos para realizar a avaliação pupilar.
- Realizar o diagnóstico de doenças neurológicas a partir das reações pupilares estimadas.
- Aumentar o conjunto de dados de dilatação pupilar para construir um modelo de aprendizado mais robusto, incluindo imagens dos olhos em iluminações variadas.
- Incluir a etapa de localização dos olhos para detectar a região dos olhos em imagens de vídeos e estimar o estado pupilar.

6.2 Produções Científicas

Como resultado desta pesquisa, um artigo científico com resultados preliminares foi publicado, conforme mostra a Tabela 9.

Tabela 9 – Artigos publicados que possuem relação com o método proposto.

Título	Congresso	Qualis
Computational Methodology for Iris Segmentation and Detection in Images from the Eyes Region Using Convolutional Neural Networks	WVC 2021	B4

Referências

- ABADI, M.; BARHAM, P.; CHEN, J.; CHEN, Z.; DAVIS, A.; DEAN, J.; DEVIN, M.; GHEMAWAT, S.; IRVING, G.; ISARD, M.; KUDLUR, M.; LEVENBERG, J.; MONGA, R.; MOORE, S.; MURRAY, D. G.; STEINER, B.; TUCKER, P.; VASUDEVAN, V.; WARDEN, P.; WICKE, M.; YU, Y.; ZHENG, X. **TensorFlow: A System for Large-Scale Machine Learning**. In: Proceedings of the 12th USENIX Conference on Operating Systems Design and Implementation. USA: USENIX Association, 2016. (OSDI'16), p. 265–283. ISBN 9781931971331.
- ACADEMY, D. S. **Deep Learning Book**. [S.l.: s.n.], 2022. Disponível em: <<https://www.deeplearningbook.com.br>>. Acesso em 31/03/2022.
- ALAZAWEE, W. S.; ABDEL-QADER, I.; ABDEL-QADER, J. **Using morphological operations — Erosion based algorithm for edge detection**. In: 2015 IEEE International Conference on Electro/Information Technology (EIT). [S.l.: s.n.], 2015. p. 521–525.
- ALOYSIUS, N.; GEETHA, M. **A review on deep convolutional neural networks**. In: 2017 International Conference on Communication and Signal Processing (ICCCSP). [S.l.: s.n.], 2017. p. 0588–0592.
- BAHDANAU, D.; CHO, K.; BENGIO, Y. **Neural Machine Translation by Jointly Learning to Align and Translate**. ArXiv, v. 1409, 09 2014.
- BAI, K.; WANG, J.; WANG, H. **A Pupil Segmentation Algorithm Based on Fuzzy Clustering of Distributed Information**. Sensors (Basel, Switzerland), v. 21, 2021.
- BASTOS, C. A. C. M.; TSANG, I. R.; VASCONCELOS, G. S.; CAVALCANTI, G. D. C. **Pupil segmentation using Pulling ; Pushing and BSOM neural network**. In: 2012 IEEE International Conference on Systems, Man, and Cybernetics (SMC). [S.l.: s.n.], 2012. p. 2359–2364.
- BENGIO, Y. **Deep Learning of Representations: Looking Forward**. 05 2013.
- BERRIEN, F. K.; HUNTINGTON, G. H. **An exploratory study of pupillary responses during deception**. Journal of Experimental Psychology, p. 443–449, 1943.
- BHURCHANDI, K.; NAWGHARE, P.; RAY, A. **An Analytical Approach for Sampling the RGB Color Space Considering Physiological Limitations of Human Vision and its Application for Color Image Analysis**. 04 2009.
- BOZOMITU, R. G.; PĂȘĂRICĂ, A.; CEHAN, V.; ROTARIU, C.; COCA, E. **Eye pupil detection using the least squares technique**. In: 2016 39th International Spring Seminar on Electronics Technology (ISSE). [S.l.: s.n.], 2016. p. 439–442.
- BREMNER, F. **Pupil assessment in optic nerve disorders**. Eye (London, England), v. 18, p. 1175–81, 12 2004.

- CHAGLASIAN, E.; AKBAR, S.; PROBST, L. **Pupil measurement using the Colvard pupillometer and a standard pupil card with a cobalt blue filter penlight**. *Journal of cataract and refractive surgery*, v. 32, p. 255–60, 02 2006.
- CHAUDHARY, A. K.; KOTHARI, R.; ACHARYA, M.; DANGI, S.; NAIR, N.; BAILEY, R.; KANAN, C.; DIAZ, G.; PELZ, J. B. **RITnet: Real-time Semantic Segmentation of the Eye for Gaze Tracking**. In: 2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW). [S.l.: s.n.], 2019. p. 3698–3702.
- CHAUHAN, R.; GHANSHALA, K.; JOSHI, R. **Convolutional Neural Network (CNN) for Image Detection and Recognition**. In: . [S.l.: s.n.], 2018. p. 278–282.
- CHEN, S.; HARALICK, R. **Recursive erosion, dilation, opening, and closing transforms**. *IEEE Transactions on Image Processing*, v. 4, n. 3, p. 335–345, 1995.
- CHENG, Y.; LI, B. **Image Segmentation Technology and Its Application in Digital Image Processing**. In: 2021 IEEE Asia-Pacific Conference on Image Processing, Electronics and Computers (IPEC). [S.l.: s.n.], 2021. p. 1174–1177.
- CHOLLET, F. et al. **Keras**. GitHub, 2015. Disponível em: <<https://github.com/fchollet/keras>>. Acessado em 18/03/2022.
- CHOROWSKI, J.; BAHDANAU, D.; CHO, K.; BENGIO, Y. **End-to-end continuous speech recognition using attention-based recurrent nn: First results**. In: NIPS 2014 Workshop on Deep Learning, December 2014. [S.l.: s.n.], 2014.
- CIRESAN, D. C.; MEIER, U.; SCHMIDHUBER, J. **Transfer learning for Latin and Chinese characters with Deep Neural Networks**. The 2012 International Joint Conference on Neural Networks (IJCNN), p. 1–6, 2012.
- de Vries, L.; FOUQUAET, I.; BOETS, B.; NAULAERS, G.; STEYAERT, J. **Autism spectrum disorder and pupillometry: A systematic review and meta-analysis**. *Neuroscience e Biobehavioral Reviews*, v. 120, p. 479–508, 2021. ISSN 0149-7634. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S014976342030590X>>. Acessado em 22/03/2022.
- DENG, J.; DONG, W.; SOCHER, R.; LI, L.-J.; LI, K.; FEI-FEI, L. **ImageNet: A large-scale hierarchical image database**. In: 2009 IEEE Conference on Computer Vision and Pattern Recognition. [S.l.: s.n.], 2009. p. 248–255.
- DHAWAN, A.; BODANI, P.; GARG, V. **Post Processing of Image Segmentation using Conditional Random Fields**. p. 729–734, 2019.
- DU, R.; MEEKER, M.; BACCHETTI, P.; LARSON, M.; HOLLAND, M.; MANLEY, G. **Evaluation of the Portable Infrared Pupillometer**. *Neurosurgery*, v. 57, p. 198–203; discussion 198, 08 2005.
- DUDA, R. O.; HART, P. E. Use of the hough transformation to detect lines and curves in pictures. *Commun. ACM*, Association for Computing Machinery, New York, NY, USA, v. 15, n. 1, p. 11–15, jan 1972. ISSN 0001-0782. Disponível em: <<https://doi.org/10.1145/361237.361242>>.

EINHAUSER, W.; KOCH, C.; CARTER, O. **Pupil dilation betrays the timing of decisions**. *Frontiers in Human Neuroscience*, v. 4, 2010. ISSN 1662-5161. Disponível em: <<https://www.frontiersin.org/article/10.3389/fnhum.2010.00018>>. Acessado em 18/03/2022.

ELHASSOUNY, A.; SMARANDACHE, F. **Trends in deep convolutional neural Networks architectures: a review**. In: 2019 International Conference of Computer Science and Renewable Energies (ICCSRE). [S.l.: s.n.], 2019. p. 1–8.

FILHO, O. M.; NETO, H. V. **Processamento digital de imagens**. [S.l.]: Brasport, 1999.

FISCHER, K.; GARTNER, B. **The Smallest Enclosing Ball of Balls: Combinatorial Structure and Algorithms**. In: Proceedings of the Nineteenth Annual Symposium on Computational Geometry. New York, NY, USA: Association for Computing Machinery, 2003. (SCG '03), p. 292–301. ISBN 1581136633.

FISCHER, K.; GÄRTNER, B.; KUTZ, M. **Fast Smallest-Enclosing-Ball Computation in High Dimensions**. In: . [S.l.: s.n.], 2003. p. 630–641. ISBN 978-3-540-20064-2.

GILL, B.; GERSHON, E.; KIM, M.-S. **Computing the Minimum Enclosing Circle of a Set of Planar Curves**. *Computer-aided Design and Applications*, v. 2, p. 301–308, 2005.

GIZA, E.; FOTIOU, D.; BOSTANTJOPOULOU, S.; KATSAROU, Z.; KARLOVASITOU, A. **Pupil Light Reflex in Parkinson's Disease: Evaluation With Pupillometry**. *The International journal of neuroscience*, v. 121, p. 37–43, 11 2010.

GOLDINGER, S. D.; PAPESH, M. H. **Pupil Dilation Reflects the Creation and Retrieval of Memories**. *Current Directions in Psychological Science*, v. 21, n. 2, p. 90–95, 2012. Disponível em: "<<https://doi.org/10.1177/0963721412436811>>". Acessado em 19/03/2022.

GONZALEZ, R. C.; WORDS, R. E. **Processamento Digital de imagens**. [S.l.: s.n.], 2010.

GOWROJU, S.; AARTI; KUMAR, S. **Robust Pupil Segmentation using UNET and Morphological Image Processing**. p. 105–109, 2021.

GRANHOLM, E.; PANIZZON, M.; ELMAN, J.; JAK, A.; HAUGER, R.; BONDI, M.; LYONS, M.; FRANZ, C.; KREMEN, W. **Pupillary Responses as a Biomarker of Early Risk for Alzheimer's Disease**. *Journal of Alzheimer's Disease*, v. 56, p. 1–10, 02 2017.

GUO, Y.; LIU, Y.; OERLEMANS, A.; LAO, S.; WU, S.; LEW, M. S. **Deep learning for visual understanding: A review**. *Neurocomputing*, v. 187, p. 27–48, 2016. ISSN 0925-2312. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0925231215017634>>. Acessado em 19/03/2022.

GUZMAN, C. P.; GONG, T.; NONGPIUR, M. E.; PERERA, S. A.; HOW, A. C.; LEE, H. K.; CHENG, L.; HE, M.; BASKARAN, M.; AUNG, T. **Anterior Segment Optical Coherence Tomography Parameters in Subtypes of Primary Angle Closure**. *Investigative Ophthalmology e Visual Science*, v. 54, n. 8, p. 5281–5286, 08 2013.

ISSN 1552-5783. Disponível em: <<https://doi.org/10.1167/iovs.13-12285>>. Acessado em 20/03/2022.

HALL, C. A.; CHILCOTT, R. P. **Eyeing up the Future of the Pupillary Light Reflex in Neurodiagnostics**. *Diagnostics*, v. 8, n. 1, 2018. ISSN 2075-4418. Disponível em: <<https://www.mdpi.com/2075-4418/8/1/19>>. Acessado em 15/03/2022.

HAN, S. Y.; KWON, H. J.; KIM, Y.; CHO, N. I. **Noise-Robust Pupil Center Detection Through CNN-Based Segmentation With Shape-Prior Loss**. *IEEE Access*, v. 8, p. 64739–64749, 2020.

HARALICK, R. M.; STERNBERG, S. R.; ZHUANG, X. **Image Analysis Using Mathematical Morphology**. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-9, n. 4, p. 532–550, 1987.

HASEGAWA, R.; BOWYER, K.; FLYNN, P. J.; ORTIZ, E.; STARK, L.; HUGHES, K. **Synthetic eye images for pupil dilation mitigation**. 2012 IEEE Fifth International Conference on Biometrics: Theory, Applications and Systems (BTAS), p. 339–345, 2012.

HASSANI, H.; GHODSI, M.; HOWELL, G. **A note on standard deviation and standard error**. *Teaching Mathematics and Its Applications: International Journal of the IMA*, v. 29, n. 2, p. 108–112, 2010.

HAYKIN, S. **Kalman Filtering and Neural Networks**. [S.l.]: Wiley, 2004. (Adaptive and Cognitive Dynamic Systems: Signal Processing, Learning, Communications and Control). Disponível em: <<https://books.google.com.br/books?id=-h9ryhE6LCIC>>. Acessado em 20/03/2022. ISBN 9780471464211.

HAYKIN, S. **Neural Networks and Learning Machines**. In: . [S.l.: s.n.], 2010.

HAYKIN, S. S. **Neural networks and learning machines**. Third. Upper Saddle River, NJ: Pearson Education, 2009.

HOFBAUER, H.; JALILIAN, E.; UHL, A. **Exploiting superior CNN-based iris segmentation for better recognition accuracy**. *Pattern Recognition Letters*, v. 120, p. 17–23, 2019. ISSN 0167-8655. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0167865518309395>>. Acessado em 20/04/2022.

HOSSEINI, M.; ARAABI, B.; SOLTANIAN-ZADEH, H. **Pigment Melanin: Pattern for Iris Recognition**. *Instrumentation and Measurement, IEEE Transactions on*, v. 59, n. 4, p. 792–804, april 2010. ISSN 0018-9456.

HU, Q.; YIN, S.; NI, H.; HUANG, Y. **An End to End Deep Neural Network for Iris Recognition**. *Procedia Computer Science*, v. 174, p. 505–517, 01 2020.

IBRAHEEM, N.; HASAN, M.; KHAN, R. Z.; MISHRA, P. **Understanding Color Models: A Review**. *ARNP Journal of Science and Technology*, v. 2, 01 2012.

JÄHNE, B. **Digital Image Processing**. [S.l.]: Springer Berlin Heidelberg, 2005. (EngineeringPro collection). Disponível em: <<https://books.google.com.br/books?id=GvWYBrDXMNkC>>. Acessado em 20/03/2022. ISBN 9783540275633.

JALILIAN, E.; UHL, A. **Iris Segmentation Using Fully Convolutional Encoder–Decoder Networks**. In: . [S.l.: s.n.], 2017.

- JOSHI, N.; JAIN, S. **A Robust Approach for Application of Morphological Operations on MRI**. In: 2020 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO). [S.l.: s.n.], 2020. p. 585–589.
- JOSHI, S.; LI, Y.; KALWANI, R.; GOLD, J. **Relationships between Pupil Diameter and Neuronal Activity in the Locus Coeruleus, Colliculi, and Cingulate Cortex**. *Neuron*, v. 89, n. 1, p. 221–234, 2016. ISSN 0896-6273. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S089662731501034X>>. Acessado em 18/03/2022.
- JU YANG HEE KYUNG, H. J.-M. Y. Y. **Efficacy of digital pupillometry for diagnosis of Horner syndrome**. *Public Library of Science*, v. 13, p. 1932–6203, 06 2017.
- JURDI, R. E.; PETITJEAN, C.; HONEINE, P.; ABDALLAH, F. **BB-UNet: U-Net With Bounding Box Prior**. *IEEE Journal of Selected Topics in Signal Processing*, v. 14, n. 6, p. 1189–1198, 2020.
- JéGOU, S.; DROZDZAL, M.; VAZQUEZ, D.; ROMERO, A.; BENGIO, Y. **The One Hundred Layers Tiramisu: Fully Convolutional DenseNets for Semantic Segmentation**. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). [S.l.: s.n.], 2017. p. 1175–1183.
- KAREMAKER, J. **An introduction into autonomic nervous function**. *Physiological Measurement*, v. 38, 03 2017.
- KAUR, A. **A Review Paper on Image Segmentation and its Various Techniques in Image Processing**. In: . [S.l.: s.n.], 2014.
- KIRUTHIGA, A. R.; ARUMUGANATHAN, R. **Smoothing of iris images and pupil segmentation using fractional derivative and wavelet transform**. In: 2017 Fourth International Conference on Signal Processing, Communication and Networking (ICSCN). [S.l.: s.n.], 2017. p. 1–6.
- KITAZUMI, K.; NAKAZAWA, A. **Robust Pupil Segmentation and Center Detection from Visible Light Images Using Convolutional Neural Network**. In: . [S.l.: s.n.], 2018. p. 862–868.
- KOTHARI, R. S.; CHAUDHARY, A. K.; BAILEY, R. J.; PELZ, J. B.; DIAZ, G. J. **EllSeg: An Ellipse Segmentation Framework for Robust Gaze Tracking**. *IEEE Transactions on Visualization e Computer Graphics*, IEEE Computer Society, Los Alamitos, CA, USA, v. 27, n. 05, p. 2757–2767, may 2021. ISSN 1941-0506.
- KOUR, A.; YV, V.; MAHESHWARI, V.; PRASHAR, D. **A Review on Image Processing**. *International Journal of Electronics Communication and Computer Engineering*, v. 4, p. 2278–4209, 01 2012.
- KREMEN, W. S.; PANIZZON, M. S.; ELMAN, J. A.; GRANHOLM, E. L.; ANDREASSEN, O. A.; DALE, A. M.; GILLESPIE, N. A.; GUSTAVSON, D. E.; LOGUE, M. W.; LYONS, M. J.; NEALE, M. C.; REYNOLDS, C. A.; WHITSEL, N.; FRANZ, C. E. **Pupillary dilation responses as a midlife indicator of risk for Alzheimer’s disease: association with Alzheimer’s disease polygenic risk**. *Neurobiology of Aging*, v. 83, p. 114–121, 2019. ISSN 0197-4580. Disponível em:

<<https://www.sciencedirect.com/science/article/pii/S0197458019303215>>. Acessado em 20/03/2022.

KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. E. **ImageNet Classification with Deep Convolutional Neural Networks**. In: PEREIRA, F.; BURGESS, C. J. C.; BOTTOU, L.; WEINBERGER, K. Q. (Ed.). *Advances in Neural Information Processing Systems*. [S.l.]: Curran Associates, Inc., 2012. v. 25. Disponível em: <<https://proceedings.neurips.cc/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf>>. Acessado em 22/03/2022.

KUMAR, J. R. H.; TEOTIA, K. **Automatic Pupil Segmentation Based On Circular Active Discs**. In: *TENCON 2019 - 2019 IEEE Region 10 Conference (TENCON)*. [S.l.: s.n.], 2019. p. 1932–1936.

LARSON, M.; BEHRENDTS, M. **Portable Infrared Pupillometry: A Review**. *Anesthesia and analgesia*, v. 120, p. 1242–53, 06 2015.

LECUN, Y.; BENGIO, Y.; HINTON, G. **Deep Learning**. *Nature*, v. 521, p. 436–44, 05 2015.

LECUN, Y.; KAVUKCUOGLU, K.; FARABET, C. **Convolutional networks and applications in vision**. *Proceedings of 2010 IEEE International Symposium on Circuits and Systems*, p. 253–256, 2010.

LEE, D. K.; IN, J.; LEE, S. Standard deviation and standard error of the mean. *Korean Journal of Anesthesiology*, v. 68, p. 220 – 223, 2015.

LI, X.; SUN, Z.; TAN, T. **Predict and improve iris recognition performance based on pairwise image quality assessment**. In: *2013 International Conference on Biometrics (ICB)*. [S.l.: s.n.], 2013. p. 1–6.

LI, Y. T.; GUO, J.-I. **A VGG-16 Based Faster RCNN Model for PCB Error Inspection in Industrial AOI Applications**. *2018 IEEE International Conference on Consumer Electronics-Taiwan (ICCE-TW)*, p. 1–2, 2018.

LIN, H.-H.; LI, Z.-Y.; SHIH, M.-H.; SUN, Y.-N.; SHEN, T.-L. **Pupil Localization for Ophthalmic Diagnosis Using Anchor Ellipse Regression**. In: *2019 16th International Conference on Machine Vision Applications (MVA)*. [S.l.: s.n.], 2019. p. 1–5.

LIN, Z.; FRANCIS, B.; MAGGIORE, M. **On the state agreement problem for multiple nonlinear dynamical systems**. 02 2022.

LIU, G.; GUO, J. **Bidirectional LSTM with attention mechanism and convolutional layer for text classification**. *Neurocomputing*, v. 337, p. 325–338, 2019. ISSN 0925-2312. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0925231219301067>>. Acessado em 26/03/2022.

LONG, J.; SHELHAMER, E.; DARRELL, T. **Fully convolutional networks for semantic segmentation**. In: *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Los Alamitos, CA, USA: IEEE Computer Society, 2015. p. 3431–3440. ISSN 1063-6919.

LYNCH, G. **Using Pupillometry to Assess the Atypical Pupillary Light Reflex and LC-NE System in ASD**. *Behavioral Sciences*, v. 8, n. 11, 2018. ISSN 2076-328X.

- MAAS, A. L.; HANNUN, A. Y.; NG, A. Y. **Rectifier nonlinearities improve neural network acoustic models**. 2013.
- MANURI, F.; SANNA, A.; PETRUCCI, C. P. **PDIF: Pupil Detection After Isolation and Fitting**. IEEE Access, v. 8, p. 30826–30837, 2020.
- MARSICO, M. D.; NAPPI, M.; RICCIO, D.; WECHSLER, H. **Mobile Iris Challenge Evaluation (MICHE) – I, Biometric iris dataset and protocols**. Pattern Recognition Letters, v. 57, 02 2015.
- MARUSINA, M.; VOLGAREVA, A.; SIZIKOV, V. **Noise suppression in the task of distinguishing the contours and segmentation of tomographic images**. Journal of Optical Technology, v. 82, p. 673, 10 2015.
- MATHÔT, S. **Pupillometry: Psychology, Physiology, and Function**. Journal of Cognition, v. 1, 02 2018.
- MAZZIOTTI, R.; CARRARA, F.; VIGLIONE, A.; LUPORI, L.; VERDE, L.; BENEDETTO, A.; RICCI, G.; SAGONA, G.; AMATO, G.; PIZZORUSSO, T. **MEYE: Web App for Translational and Real-Time Pupillometry**. eNeuro, v. 8, p. ENEURO.0122–21.2021, 09 2021.
- MCCLOY, D. R.; LAU, B. K.; LARSON, E.; PRATT, K. A. I.; LEE, A. K. C. **Pupillometry shows the effort of auditory attention switching**. The Journal of the Acoustical Society of America, v. 141 4, p. 2440, 2017.
- MILLETARI, F.; NAVAB, N.; AHMADI, S.-A. **V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation**. In: . [S.l.: s.n.], 2016. p. 565–571.
- MINKOWSKI, H. **Volumen und Oberfläche**. Mathematische Annalen, v. 57, p. 447–495, 1903.
- MNIH, V.; HEES, N.; GRAVES, A.; KAVUKCUOGLU, K. **Recurrent Models of Visual Attention**. Advances in Neural Information Processing Systems, v. 3, 06 2014.
- MOHAMED, C.; NSIRI, B.; ABDELMAJID, S.; ABDELGHANI, E. M.; BRAHIM, B. **Deep Convolutional Networks for Image Segmentation: Application to Optic Disc detection**. p. 1–3, 2020.
- NAMBIAR, A. **Brain Tumor Segmentation using Image Processing**. International Journal for Research in Applied Science and Engineering Technology, v. 8, p. 2072–2075, 06 2020.
- NINGNING, G. **Artificial neural network management system**. In: The 2nd International Conference on Information Science and Engineering. [S.l.: s.n.], 2010. p. 1891–1894.
- NIU, Z.; ZHONG, G.; YU, H. **A review on the attention mechanism of deep learning**. Neurocomputing, v. 452, p. 48–62, 2021. ISSN 0925-2312. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S092523122100477X>>. Acessado em 22/04/2022.

- NOWAK, W.; ŻAROWSKA, A.; SZUL-PIETRZAK, E.; O, M. **System and measurement method for binocular pupillometry to study pupil size variability**. Biomedical engineering online, v. 13, p. 69, 06 2014.
- NUGROHO, R. H.; NASRUN, M.; SETIANINGSIH, C. **Lie detector with pupil dilation and eye blinks using hough transform and frame difference method with fuzzy logic**. In: 2017 International Conference on Control, Electronics, Renewable Energy and Communications (ICCREC). [S.l.: s.n.], 2017. p. 40–45.
- OKTAY, O.; SCHLEMPER, J.; FOLGOC, L.; LEE, M.; HEINRICH, M.; MISAWA, K.; MORI, K.; MCDONAGH, S.; HAMMERLA, N.; KAINZ, B.; GLOCKER, B.; RUECKERT, D. **Attention U-Net: Learning Where to Look for the Pancreas**. 04 2018.
- PADMAVATHI; KUPPUSAMY, M. **An Innovative Technique to Evaluate Quantitative Pupillary Light Reflex by Dynamic Pupillometry using Infrared Videography**. Journal of Clinical and Diagnostic Research, v. 13, p. CM01–CM03, 04 2019.
- PALUS, H. **Preprocessing and postprocessing in process of colour image segmentation**. 01 2007.
- PARRA, E.; EDWARDS, M.; CHA, D.; ROY, K.; JOHNSON, M. **Analysis of iris surface features in populations of diverse ancestry**. Royal Society Open Science, v. 3, 01 2016.
- PASSOS, E. C.; ANDRADE-NETO, A. V.; LEMAIRE, T. **Comportamento Ótico do Olho Humano e suas Ametropias**. Caderno de física da uefs 06 (01 e 02), v. 2008, p. 7–18, 2008.
- PELES, D. N.; HONG, L.; HU, D.-N.; ITO, S.; NEMANICH, R. J.; SIMON, J. D. **Human Iridal Stroma Melanosomes of Varying Pheomelanin Contents Possess a Common Eumelanin Outer Surface**. The Journal of Physical Chemistry B, v. 113, n. 32, p. 11346–11351, 2009. Disponível em: <<https://doi.org/10.1021/jp904138n>>. Acessado em 23/03/2022.
- PETRONELA, B.; ARCADIE, C.; BONTEANU, B. R. G.; GABRIEL. **A New Robust Pupil Detection Algorithm for Eye Tracking Based Human-Computer Interface**. 2019.
- PIAGGIO, D.; NAMM, G.; MELILLO, P.; SIMONELLI, F.; IADANZA, E.; PECCHIA, L. **Pupillometry via smartphone for low-resource settings**. Biocybernetics and Biomedical Engineering, v. 41, n. 3, p. 891–902, 2021. ISSN 0208-5216.
- PROENCA, H.; FILIPE, S.; SANTOS, R.; OLIVEIRA, J.; ALEXANDRE, L. A. **The UBIRIS.v2: A Database of Visible Wavelength Iris Images Captured On-the-Move and At-a-Distance**. IEEE Trans. Pattern Anal. Mach. Intell., IEEE Computer Society, USA, v. 32, n. 8, p. 1529–1535, aug 2010. ISSN 0162-8828.
- RAMLEE, I. D. R.; RAMLI, A.; NOH, Z. M. **Pupil Segmentation of Abnormal Eye using Image Enhancement in Spatial Domain**. IOP Conference Series: Materials Science and Engineering, v. 210, p. 012031, 06 2017.

- RAWAT, W.; WANG, Z. **Deep Convolutional Neural Networks for Image Classification: A Comprehensive Review**. *Neural Computation*, v. 29, n. 9, p. 2352–2449, 2017. PMID: 28599112.
- REN, J. S. J.; XU, L. **On Vectorization of Deep Convolutional Neural Networks for Vision Tasks**. [S.l.: s.n.], 2015.
- ROHLFING, T. Image similarity and tissue overlaps as surrogates for image registration accuracy: Widely used but unreliable. *IEEE transactions on medical imaging*, v. 31, p. 153–63, 08 2011.
- RONNEBERGER, O.; FISCHER, P.; BROX, T. U-net: Convolutional networks for biomedical image segmentation. In: NAVAB, N.; HORNEGGER, J.; WELLS, W. M.; FRANGLI, A. F. (Ed.). *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*. Cham: Springer International Publishing, 2015. p. 234–241.
- ROSENFELD, A.; PFALTZ, J. **Distance functions on digital pictures**. *Pattern Recognition*, v. 1, n. 1, p. 33–61, 1968. ISSN 0031-3203. Disponível em: 20/03/2022 <<https://www.sciencedirect.com/science/article/pii/0031320368900137>>. Acessado em.
- ROSS-HOWE, S.; TIZHOOSH, H. R. **The Effects of Image Pre- and Post-Processing, Wavelet Decomposition, and Local Binary Patterns on U-Nets for Skin Lesion Segmentation**. p. 1–8, 2018.
- ROSSUM, G. V.; JR, F. L. D. **Python reference manual**. [S.l.]: Centrum voor Wiskunde en Informatica Amsterdam, 1995.
- ROY, S.; KARMAKAR, A.; DAS, S.; NANDY, S. **Constrained Minimum Enclosing Circle with Center on a Query Line Segment**. In: . [S.l.: s.n.], 2006. v. 42, p. 765–776. ISBN 978-3-540-37791-7.
- ROY, S.; MITRA, A.; SETUA, S. K. **Color Image Representation Using Multivector**. In: 2014 5th International Conference on Intelligent Systems, Modelling and Simulation. [S.l.: s.n.], 2014. p. 357–363.
- SIMONYAN, K.; ZISSERMAN, A. **Very Deep Convolutional Networks for Large-Scale Image Recognition**. CoRR, abs/1409.1556, 2015.
- SINGH, J.; BANERJEE, R. **A Study on Single and Multi-layer Perceptron Neural Network**. p. 35–40, 2019.
- SINHAROY, I. **Scheimpflug with computational imaging to extend the depth of field of iris recognition systems**. Tese (Doutorado), 12 2016.
- SIROIS, S.; BRISSON, J. **Pupillometry**. *Wiley Interdisciplinary Reviews: Cognitive Science*, v. 5, 11 2014.
- SOMUNCUOGLU, A.; PURUTÇUOĞLU, V.; ARI, F.; GOKCAY, D. **Investigation on the Use of Hidden Layers, Different Numbers of Neurons and Different Activation Functions to Detect Pupil Dilation Responses to Stress**. In: . [S.l.: s.n.], 2020. p. 1–4.

- SONG, S.; LAN, C.; XING, J.; ZENG, W.; LIU, J. **An End-to-End Spatio-Temporal Attention Model for Human Action Recognition from Skeleton Data**. In: AAAI. [S.l.: s.n.], 2017.
- SOOYEOUN HONG JEONG-HO, Y. J. Y. **Analysis of pupillometer results according to disease stage in patients with Parkinson's disease**. Scientific Reports, v. 11, 2021. ISSN 2045-2322. Disponível em: <<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8429555/>>.
- SYLVESTER, J. J. **A question in the geometry of situation**. Quarterly Journal of Pure and Applied Mathematics, v. 1, n. 1, p. 79–80, 1857.
- TAHA, A. A.; HANBURY, A. **Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool**. BMC Medical Imaging, v. 15, 2015.
- TEMEL, D.; MATHEW, M. J.; ALREGIB, G.; KHALIFA, Y. M. **Relative Afferent Pupillary Defect Screening Through Transfer Learning**. IEEE Journal of Biomedical and Health Informatics, v. 24, n. 3, p. 788–795, 2020.
- TOMEIO-REYES, I. **Robust Iris Recognition using Decision Fusion and Degradation Modelling**. Tese (Doutorado), 09 2015.
- TRIYANI, Y.; NUGROHO, H. A.; RAHMAWATY, M.; ARDIYANTO, I.; CHORIDAH, L. **Performance analysis of image segmentation for breast ultrasound images**. p. 1–6, 2016.
- TSUKAHARA, J.; HARRISON, T.; ENGLE, R. **The relationship between baseline pupil size and intelligence**. Cognitive Psychology, v. 91, p. 109–123, 11 2016.
- VENKATESH, Y.; RAJA, S.; RAMYA, N. **Multiple contour extraction from graylevel images using an artificial neural network**. IEEE Transactions on Image Processing, v. 15, n. 4, p. 892–899, 2006.
- VOGELZANG, M.; HENDRIKS, P.; RIJN, H. van. **Pupillary responses reflect ambiguity resolution in pronoun processing**. Language, Cognition and Neuroscience, Routledge, v. 31, n. 7, p. 876–885, 2016. Disponível em: <<https://doi.org/10.1080/23273798.2016.1155718>>. Acessado em? 18/03/2022.
- WANGWIWATTANA, C.; DING, X.; LARSON, E. C. **PupilNet, Measuring Task Evoked Pupillary Response using Commodity RGB Tablet Cameras**. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, v. 1, p. 1 – 26, 2017.
- WAXENBAUM, J.; VARACALLO, M. **Anatomy, Autonomic Nervous System**. 04 2019.
- WENG, C.-C.; CHEN, H.; FUH, C.-S. **A novel automatic white balance method for digital still cameras**. Proc IEEE Int Symp Circuits System, v. 4, p. 3801 – 3804 Vol. 4, 06 2005.
- WILLOUGHBY, C.; MD, D.; FERRARI, S.; MD, A.; LANDAU, K.; OMIDI, Y. **Anatomy and physiology of the human eye: Effects of mucopolysaccharidoses disease on structure and function - a review**. Clinical e Experimental Ophthalmology, v. 38, p. 2 – 11, 08 2010.

- WINN, M. B.; WENDT, D.; KOELEWIJN, T.; KUCHINSKY, S. E. **Best Practices and Advice for Using Pupillometry to Measure Listening Effort: An Introduction for Those Who Want to Get Started**. Trends in Hearing, v. 22, p. 2331216518800869, 2018. PMID: 30261825.
- WINN, M. B.; WENDT, D.; KOELEWIJN, T.; KUCHINSKY, S. E. **Best Practices and Advice for Using Pupillometry to Measure Listening Effort: An Introduction for Those Who Want to Get Started**. Trends in Hearing, v. 22, p. 2331216518800869, 2018. Disponível em: <<https://doi.org/10.1177/2331216518800869>>. Acessado em 22/03/2022.
- WONG, S.-L.; ARII, S.; PARAMESRAN, R.; TAGUCHI, A. **Color image processing system based on color/brightness constancy**. p. 21–24, 2015.
- WU, Q.; LIU, Y.; LI, Q.; JIN, S.; LI, F. **The application of deep learning in computer vision**. p. 6522–6527, 2017.
- XIAO, L.; YAN, Q.; DENG, S. **Scene classification with improved AlexNet model**. In: 2017 12th International Conference on Intelligent Systems and Knowledge Engineering (ISKE). [S.l.: s.n.], 2017. p. 1–6.
- XU, K.; BA, J.; KIROS, R.; CHO, K.; COURVILLE, A.; SALAKHUDINOV, R.; ZEMEL, R.; BENGIO, Y. **Show, Attend and Tell: Neural Image Caption Generation with Visual Attention**. In: BACH, F.; BLEI, D. (Ed.). *Proceedings of the 32nd International Conference on Machine Learning*. Lille, France: PMLR, 2015. (Proceedings of Machine Learning Research, v. 37), p. 2048–2057. Disponível em: <<https://proceedings.mlr.press/v37/xuc15.html>>. Acessado em 22/03/2022.
- XU, S.; FREUND, R.; SUN, J. **Solution Methodologies for the Smallest Enclosing Circle Problem**. Computational Optimization and Applications, v. 25, 12 2003.
- YIU, Y.-H.; ABOULATTA, M.; RAISER, T.; OPHEY, L.; FLANAGIN, V. L.; zu Eulenburg, P.; AHMADI, S.-A. **DeepVOG: Open-source pupil segmentation and gaze estimation in neuroscience using deep learning**. Journal of Neuroscience Methods, v. 324, p. 108307, 2019. ISSN 0165-0270. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0165027019301578>>. Acessado em 18/03/2022.
- YU, D.; WANG, H.; CHEN, P.; WEI, Z. **Mixed Pooling for Convolutional Neural Networks**. In: . [S.l.: s.n.], 2014. p. 364–375. ISBN 978-3-319-11739-3.
- YU, W.; YANG, K.; BAI, Y.; YAO, H.; RUI, Y. **DNN Flow: DNN feature pyramid based image matching**. p. 109.1–109.10, 01 2014.
- ZEILER, M. D.; FERGUS, R. *Visualizing and understanding convolutional networks*. In: In Computer Vision–ECCV 2014. [S.l.]: Springer, 2014. p. 818–833.
- ZHANG, W.; LU, X.; GU, Y.; LIU, Y.; MENG, X.; LI, J. **A Robust Iris Segmentation Scheme Based on Improved U-Net**. IEEE Access, v. 7, p. 85082–85089, 2019.
- ZHU, J.; ZHANG, E.; RIO-TSONIS, K. **Eye Anatomy**. In: _____. [S.l.: s.n.], 2012. ISBN 0470016175.

ZOU, K.; WARFIELD, S.; BHARATHA, A.; TEMPANY, C.; KAUS, M.; HAKER, S.; WELLS, W.; JOLESZ, F.; KIKINIS, R. Statistical validation of image segmentation quality based on a spatial overlap index. *Academic radiology*, v. 11, p. 178–89, 02 2004.