



UNIVERSIDADE FEDERAL DO MARANHÃO

Curso de Ciência da Computação

Matheus Henrique Aguiar Nunes

**Detecção automática de doenças da visão em  
imagens de reflexo vermelho utilizando Deep  
Features**

São Luís - MA

2023

Matheus Henrique Aguiar Nunes

# **Detecção automática de doenças da visão em imagens de reflexo vermelho utilizando Deep Features**

Monografia apresentada ao curso de Ciência da Computação da Universidade Federal do Maranhão, como parte dos requisitos necessários para obtenção do grau de Bacharel em Ciência da Computação.

Orientador: Prof. Dr. João Dallyson Sousa de Almeida

São Luís - MA

2023

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

Nunes, Matheus Henrique Aguiar.

Detecção automática de doenças da visão em imagens de reflexo vermelho utilizando Deep Features / Matheus Henrique Aguiar Nunes. - 2023.

47 f.

Orientador(a): João Dallyson Sousa de Almeida.

Curso de Ciência da Computação, Universidade Federal do Maranhão, São Luís - MA, 2023.

1. Deep Features. 2. Doenças da Visão. 3. Processamento de Imagens. 4. Redes Neurais. 5. Teste de Bruckner. I. Almeida, João Dallyson Sousa de. II. Título.

Matheus Henrique Aguiar Nunes

# **Detecção automática de doenças da visão em imagens de reflexo vermelho utilizando Deep Features**

Monografia apresentada ao curso de Ciência da Computação da Universidade Federal do Maranhão, como parte dos requisitos necessários para obtenção do grau de Bacharel em Ciência da Computação.

Trabalho aprovado em São Luís - MA, 06 de Janeiro de 2023:

---

**Prof. Dr. João Dallyson Sousa de Almeida**  
Orientador  
Universidade Federal do Maranhão

---

**Prof. Dr. Geraldo Braz Junior**  
Universidade Federal do Maranhão

---

**Prof. Ms. Italo Francyles Santos da Silva**  
Universidade Federal do Maranhão

São Luís - MA

2023

# Agradecimentos

Em primeiro lugar a Deus, e aos meus avós, as pessoas mais fortes que conheço. Em especial a dona Maria Antônia, que não teve a oportunidade de ver em vida, mas agora terá um neto formado. E a meu avô Manoel Aguiar, por quem tenho o maior carinho de todos.

Aos meus pais, que fizeram o possível e impossível pra me apoiar nesta jornada, meus parentes de Altamira, além dos familiares de São Luís que me acolheram por todo o tempo que precisei, em especial a minha tia Sandra Regina.

A minha noiva e futura esposa Amanda Soares, que conheci na Universidade, esteve comigo e me apoiou incondicionalmente nos momentos que mais precisei, e por isso serei eternamente grato.

Ao professor João Dallyson, por ter confiado em mim, pela paciência de me orientar (que não foi pouca) e por todo o apoio que me deu.

A professora Rose Ferreira, que me deu a oportunidade de iniciar na pesquisa e ajudou a despertar em mim o interesse pela área que trabalho hoje.

Aos amigos de UFMA, que desde o primeiro dia me acompanharam nessa trajetória com muito companheirismo. Em especial, ao pessoal do grupo CCovardes.

Aos demais professores pela paciência e colegas de São Luís, lugar que me recebeu de braços abertos.

*“Tudo o que fazemos acaba nos rendendo aprendizado.  
As passagens boas, as passagens ruins, tudo faz parte da vida.”*

**Gugu Liberato**

# Resumo

O Teste de Bruckner, popularmente conhecido por exame do reflexo vermelho, é um método simples e indolor de diagnóstico que possui o objetivo de detectar patologias da visão. Por meio da observação do reflexo retiniano vermelho, realizada com um equipamento chamado oftalmoscópio direto, características das estruturas internas são identificáveis e podem indicar uma possível saúde ocular comprometida. Informações do Censo Oftalmológico 2021 apontam que mesmo com a saturação de pessoal especializado nas capitais e um constante movimento de êxodo urbano, ainda há deficiência de profissionais da área em várias regiões. No Estado do Maranhão, por exemplo, a proporção de habitantes para um oftalmologista é três vezes maior que a média nacional. Além disso, 15 municípios acima de 40 mil habitantes no estado não possuem sequer a presença de um profissional, o que dificulta o acesso rápido aos devidos cuidados. Nestes casos, o suporte do Teste de Bruckner é de grande valia aos profissionais, melhorando a qualidade da triagem em regiões mais distantes dos grandes centros. A metodologia combina o uso de descritores de características baseados em aprendizado profundo e classificadores para identificar a presença de patologias em imagens de reflexo vermelho, e os estudos apontaram um destaque no uso da rede neural convolucional *DeepLoc* e um *ensemble* entre os classificadores Regressão Logística, Random Forest e SVM, alcançando 93,20% de precisão, sensibilidade de 93,20% e especificidade de 84,50%.

**Palavras-chaves:** Teste de Bruckner, Doenças da Visão, Processamento de Imagens, Deep Features, Redes Neurais.

# Abstract

Bruckner's Test, popularly known as red reflex exam, is a simple and painless diagnosis method aimed at early detection of vision related diseases. Observing retinal red reflex with an equipment called direct ophthalmoscope, internal structural features are identified and can alert to a compromised eye health. Information of 2021 Brazilian Ophthalmological Census point that even with the professional saturation on capitals and a constant urban exodus, there is still missing ophthalmologists in some regions. In State of Maranhão, for example, the proportion between inhabitants and ophthalmologists is three times bigger than the national average. Furthermore, 15 municipalities with population higher than 40.000 doesn't even have the presence of one professional, what makes fast access to needed care even more difficult. The methodology combines characteristic descriptors based on deep learning and classifiers to identify the presence of pathologies in red reflex images, and studies highlighted the use of the convolucional neural network DeepLoc and an ensemble combining Linear Regression, Random Forest and SVM classifiers, achieving 93,20% of precision, sensibility of 93,20% and specificity of 84,50%.

**Keywords::** Bruckner's Test, Vision Diseases, Image processing, Deep Features, Neural Networks.

# Lista de ilustrações

Figura 1 – Aplicação do Teste do Olhinho realizada em um ambulatório. . . . .	17
Figura 2 – Ilustração da arquitetura da rede VGG-16. . . . .	22
Figura 3 – Ilustração da arquitetura da rede Inception v3. . . . .	23
Figura 4 – Ilustração da arquitetura da rede <i>DeepLoc</i> . . . . .	23
Figura 5 – Ilustração da aplicação da técnica de Análise de Componentes Principais. . . . .	25
Figura 6 – Pseudo-códigos para treinamento e teste do algoritmo kNN. . . . .	26
Figura 7 – Separação de classes por meio de hiperplanos, método usado pelo algoritmo SVM. . . . .	27
Figura 8 – (a) Imagem classificada com patologias e (b) imagem classificada sem patologias. . . . .	29
Figura 9 – Apresentação do fluxograma criado na plataforma Orange. . . . .	30
Figura 10 – Diagrama ilustrando os passos desta pesquisa. . . . .	31
Figura 11 – (a) Imagem original com reflexo vermelho, (b) imagem de máscara contendo a região de interesse extraída a partir de (a) e (c) imagem pós-aplicação da máscara. . . . .	31
Figura 12 – Imagens (a) original e (b) com a aplicação da máscara com a região de interesse, da classe sem patologias. . . . .	39
Figura 13 – Imagens (a) original e (b) com a aplicação da máscara com a região de interesse, da classe com patologias. . . . .	39
Figura 14 – Imagens (a) original e (b) com a aplicação da máscara com a região de interesse, da classe com patologia e que gerou classificações falhas. . . . .	40

# Lista de tabelas

Tabela 1 – Resultados da pesquisa de Pinto (2020). . . . .	15
Tabela 2 – Condições dos operadores Bitwise em imagens. . . . .	19
Tabela 3 – Dimensionalidade do vetor de características gerado por cada uma das redes usadas. . . . .	32
Tabela 4 – Parâmetros utilizados nos classificadores. . . . .	34
Tabela 5 – Resultados usando a técnica de redução de dimensionalidade <i>Gain Ratio</i> . . . . .	36
Tabela 6 – Resultados usando a técnica de redução de dimensionalidade <i>PCA-10</i> . . . . .	37
Tabela 7 – Resultados usando a técnica de redução de dimensionalidade <i>Remove Sparse Values</i> . . . . .	37
Tabela 8 – Métricas alcançadas a partir da aplicação de <i>ensemble</i> entre os classificadores Random Forest, Regressão Logística e SVM. . . . .	38
Tabela 9 – Comparação entre métodos propostos . . . . .	38

# Lista de abreviaturas e siglas

CLCM	Matriz de Coocorrência em Nível de Cores
CNN	Rede Neural Convolucional
FN	Falso Negativo
FP	Falso Positivo
GLCM	Matriz de Coocorrência em Nível de Cinza
INCA	Instituto Nacional do Câncer
MCP	McCulloch-Pitts
NCA	Núcleo de Computação Aplicada
OMS	Organização Mundial da Saúde
PDI	Processamento Digital de Imagens
PCA	Principal Component Analysis
ReLU	Rectified Linear Units
SVM	Máquina de Vetores de Suporte
TRV	Teste do Reflexo Vermelho
UFMA	Universidade Federal do Maranhão
VP	Verdadeiro Positivo
VN	Verdadeiro Negativo

# Sumário

<b>1</b>	<b>INTRODUÇÃO</b>	<b>13</b>
<b>1.1</b>	<b>Justificativa</b>	<b>14</b>
<b>1.2</b>	<b>Objetivo</b>	<b>14</b>
1.2.1	Objetivos Específicos	14
<b>1.3</b>	<b>Contribuições</b>	<b>14</b>
<b>1.4</b>	<b>Trabalhos Relacionados</b>	<b>15</b>
<b>1.5</b>	<b>Organização do Trabalho</b>	<b>16</b>
<b>2</b>	<b>FUNDAMENTAÇÃO TEÓRICA</b>	<b>17</b>
<b>2.1</b>	<b>O Teste de Bruckner</b>	<b>17</b>
<b>2.2</b>	<b>Operadores em Imagens</b>	<b>18</b>
<b>2.3</b>	<b>Redes Neurais</b>	<b>19</b>
2.3.1	Redes Neurais Convolucionais	20
<b>2.4</b>	<b>Deep Features</b>	<b>21</b>
2.4.1	VGG-16 e VGG-19	21
2.4.2	Inception v3	22
2.4.3	DeepLoc	23
<b>2.5</b>	<b>Técnicas de Redução de Dimensionalidade</b>	<b>24</b>
2.5.1	Análise de Componentes Principais	24
2.5.2	Gain Ratio	25
<b>2.6</b>	<b>Classificação</b>	<b>25</b>
2.6.1	kNN	26
2.6.2	Máquina de Vetor de Suporte	26
2.6.3	Random Forest	27
2.6.4	Regressão Logística	28
2.6.5	Naïve Bayes	28
<b>3</b>	<b>MATERIAIS E MÉTODO</b>	<b>29</b>
<b>3.1</b>	<b>Base de Imagens</b>	<b>29</b>
<b>3.2</b>	<b>A Plataforma Orange</b>	<b>29</b>
<b>3.3</b>	<b>Método Proposto</b>	<b>30</b>
3.3.1	Segmentação de imagens	31
3.3.2	Extração de características	32
3.3.3	Seleção de Características	33
3.3.4	Treinamento e Classificação	33
<b>3.4</b>	<b>Métricas de Avaliação</b>	<b>34</b>

<b>4</b>	<b>RESULTADOS E DISCUSSÃO</b>	<b>36</b>
4.1	Apresentação dos resultados	36
4.2	Comparação dos resultados	38
4.3	Estudos de caso	38
<b>5</b>	<b>CONCLUSÃO</b>	<b>41</b>
5.1	Trabalhos futuros	42
	<b>REFERÊNCIAS</b>	<b>43</b>

# 1 Introdução

A Atenção Primária à Saúde é uma estratégia governamental que busca melhorar a saúde dos cidadãos por meio de prevenção contínua (CRUZ, 2008), e a ação dos profissionais é essencial para a descoberta de patologias em fases iniciais, aumentando as chances de cura. A presença de uma atenção oftalmológica minuciosa se faz necessária para a manutenção de uma boa qualidade de vida.

O Retinoblastoma é uma das patologias que possuem sintomas identificáveis através de um exame de visualização como o Teste de Bruckner, conhecido popularmente por Teste do Olhinho. É um tipo raro de tumor maligno originário dos retinoblastos imaturos da retina neural. Dentre suas formas de apresentação, pode-se citar estrabismo, olho vermelho e aumento do tamanho do globo ocular por pressão intraocular elevada, além de baixa de visão (AGUIAR, 2010). Há uma discussão importante acerca do cuidado necessário para com as crianças e a importância dos exames de triagem para a identificação de doenças como essa.

Várias crianças não possuem o devido acesso à saúde ocular, e patologias como o Retinoblastoma podem passar despercebidas por falta de profissionais qualificados para a execução de exames de triagem. O diagnóstico precoce do Retinoblastoma é fundamental, já que lesões em estágio inicial tem tendência maior de cura. Além disso, pacientes com histórico familiar da doença precisam de uma avaliação deste o nascimento (AGUIAR, 2010).

A rapidez na identificação destas doenças pode significar muito na eficácia das iniciativas para combate de tais patologias. Manter os profissionais de saúde envolvidos no cuidado infantil para detectar anomalias é um ponto chave para uma melhor taxa de cura (KHEDEKAR et al., 2019).

Deve-se também considerar que o uso de métodos de detecção automatizada podem ser aplicados com o objetivo de propagar este conhecimento, na forma de uma plataforma disponível para o público em geral, além de oferecer uma forma de teste sem o uso de anestesia ou dilatação farmacológica da pupila. Mesmo assim, a aplicação de métodos de detecção automatizada necessitam de validação com confirmação clínica (KHEDEKAR et al., 2019) e não deve ser descartada uma apuração mais específica como a feita por um profissional oftalmológico.

## 1.1 Justificativa

O Estado do Maranhão conta com uma proporção de aproximadamente 18.057 habitantes para cada oftalmologista, abaixo do valor estabelecido pela OMS para países em desenvolvimento, de 1 para 17 mil. Estas informações são pertencentes ao Censo do Conselho Brasileiro de Oftalmologia (2021) e este também aponta que ao menos quinze municípios com população acima de 40 mil habitantes não possuem sequer um profissional da área. Em comparação, a média na Região Nordeste é de 1 profissional para cada 8.408 habitantes e no retrato nacional são cerca de 6.188 brasileiros para cada profissional. Essa disparidade evidencia a carência de profissionais qualificados para servir de apoio para os pacientes e pode significar uma maior incidência de problemas oftalmológicos no futuro e consequentemente diminuição da qualidade de vida desta população.

A triagem executada pelo Teste do Reflexo Vermelho (TRV) é essencial para a detecção de doenças como a catarata, inflamações na retina ou vítreo, que podem levar até a cegueira total. Em relação ao retinoblastoma, há uma taxa de mortalidade relacionada, então a identificação da presença da doença pode significar redução da morbimortalidade associada (TAMURA; TEIXEIRA, 2009).

## 1.2 Objetivo

O objetivo deste trabalho é propôr um novo método para classificação de patologias a partir de imagens de reflexo vermelho na região retiniana usando *deep features*, a fim de tornar o processo de verificação de doenças por meio do Teste do Olhinho mais prático e confiável, possibilitando sua aplicação em plataformas facilmente acessíveis para o público.

### 1.2.1 Objetivos Específicos

- Avaliar a utilização de redes neurais para a caracterização de imagens de reflexo vermelho a fim de detectar características e classificá-las;
- Avaliar estes resultados comparando a classificação com a região ocular completa e apenas a área de interesse;
- Avaliar o uso de deep features com diferentes classificadores.
- Comparar os resultados alcançados com outros trabalhos da literatura;

## 1.3 Contribuições

Dentre as contribuições, pode-se citar a apresentação de um novo método para a detecção de patologias da visão que ofereça mais confiabilidade ao processo, além da

utilização de Redes Neurais para extração de deep features a partir das imagens.

## 1.4 Trabalhos Relacionados

A emergência de ações que ajudem a melhorar a realidade deste grupo de pessoas que não possuem acesso facilitado à saúde oftalmológica e a falta de métodos que permitam maior acurácia e confiabilidade do público alvo deste experimento justificam a necessidade de estudos mais elaborados sobre o tema.

O Núcleo de Computação Aplicada da UFMA possui estudos alinhados com o tópico e isso pode ser confirmado através de seus trabalhos publicados, como o de [Silva et al. \(2019\)](#) e de [Pinto \(2020\)](#). O primeiro engloba uma parte maior do processo, da aquisição a segmentação da região de interesse, enquanto o segundo dá mais enfoque ao processo de classificação, tal qual esta pesquisa.

[Silva et al. \(2019\)](#) traz contribuições para as etapas de localização de olhos, segmentação do reflexo retiniano e classificação do mesmo, a partir de descritores de textura e cor. Com o uso do classificador SVM, estas características são analisadas separadamente e em conjunto e foi verificado que a análise de cores por meio da combinação das matrizes CLCM com as métricas do histograma de cores local grid 5x5 contribuiu de melhor forma com o classificador, trazendo uma taxa de 95,25% de acurácia, 98,9% de especificidade, resultados bem avaliados. No tópico de trabalhos futuros de sua pesquisa, é mencionada a utilização de Deep Features como modelo de aprendizagem profunda, que é um dos pontos-chave desta pesquisa.

[Pinto \(2020\)](#) propõe a utilização de Redes Neurais Convolucionais para o processo de extração de características e classificação do modelo. Resultados de experimentos com o uso de embeddings e meta-classificadores combinados com diferentes espaços de cor como HSV (Hue, Saturation, Value) e Lab entregaram taxas satisfatórias, que podem ser vistas na Tabela 1.

Tabela 1 – Resultados da pesquisa de [Pinto \(2020\)](#).

Espaço de Cor	Acurácia	Sensibilidade	F1	Especificidade	Característica
HSV	82.72%	80%	71.47%	83.47%	Embeddings
RGB	83.25%	80%	71.30%	84.16%	Histograma
Lab	90.60%	70%	71.66%	97.77%	Meta

Fonte: ([PINTO, 2020](#))

Apesar dos resultados não serem suficientes para superar os encontrados na pesquisa de [Silva et al. \(2019\)](#), [Pinto \(2020\)](#) validou a utilização de CNNs como classificadores no processo de detecção de patologias da visão. Este é um ponto importante para demonstrar

a necessidade da continuidade dos estudos nesta temática, e é de onde esta pesquisa pretende continuar.

## 1.5 Organização do Trabalho

Neste capítulo, é apresentada uma contextualização do problema a fim de mostrar onde se baseiam os objetivos do trabalho. A seguir, no Capítulo 2, é apresentada a fundamentação teórica com tópicos necessários à execução da pesquisa, além da introdução dos métodos utilizados na pesquisa, que está disponível mais abaixo, no Capítulo 3. Dentre estes, as Redes Neurais que baseiam o processo de extração de características, os classificadores e uma apresentação dos softwares utilizados neste processo. Mais à frente, no Capítulo 4, a apresentação dos resultados obtidos, discussão e comparação com outros trabalhos, e no último capítulo (Capítulo 5) as considerações finais deste trabalho.

## 2 Fundamentação Teórica

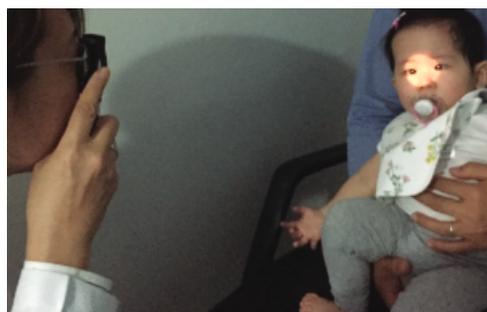
Neste capítulo, serão apresentados conceitos utilizados de base para esta pesquisa, abordando pontos-chave para a execução da mesma, como o de processamento de imagens, redes neurais, aprendizado profundo, extração de *deep features*, além de ir mais a fundo nos conceitos das principais redes neurais utilizadas na extração das características e classificadores usados na detecção de parâmetros que ajudem na identificação de patologias da visão nas imagens de reflexo vermelho.

### 2.1 O Teste de Bruckner

O Teste de Bruckner é um procedimento simples, indolor e de baixo custo para detecção de patologias por meio de anomalias de refração, especialmente em crianças. A aplicação do Teste do Olhinho é simples e de baixo custo, tornando-a eficiente para a detecção de sinais de doenças relacionadas à visão, sendo assim uma significativa estratégia de prevenção da cegueira na infância (AGUIAR, 2010).

Usado por profissionais da atenção primária, este exame utiliza a transmissão de luz a partir de um equipamento conhecido como Oftalmoscópio Direto, que imediatamente é refletida no fundo do olho e apresentada ao examinador na forma de reflexo vermelho, como apresentado na Figura 1.

Figura 1 – Aplicação do Teste do Olhinho realizada em um ambulatório.



Fonte: Hopker et al. (2018)

De acordo com Hopker et al. (2018),

"O Teste de Reflexo Vermelho é uma ferramenta de alta sensibilidade para o rastreamento de alterações oculares com risco de causar ambliopia ou deficiência visual (cegueira ou baixa visão)"(2018, p.1).

Um dos principais sintomas que identificam a presença de patologias é a leucocoria, quando há um reflexo de aparência esbranquiçada na pupila. A alteração ocular no resultado

pode significar a presença de doenças como catarata congênita ou retinoblastoma, que impactam na qualidade de vida e podem oferecer até risco de vida ao paciente (AGUIAR, 2010).

Segundo Hopker et al. (2018), existem algumas recomendações para a realização do teste, como a execução em uma sala escurecida, que por deixar a visualização do resultado mais fácil elimina a necessidade do uso de colírios para dilatação das pupilas e o posicionamento do equipamento entre 50 centímetros e 1 metro de distância do paciente.

Erros refrativos podem ser identificados por meio deste teste e são indicativos de possível ocorrência de problemas como dificuldade de atenção no âmbito escolar e conseqüentemente menor aprendizado. Por isso, há uma certa urgência para que esta testagem seja feita o quanto antes e possibilite a identificação enquanto há tempo hábil para correção.

Estudos como o de Jalis, Ashfaq e Imdad (2015) corroboram o potencial que o Teste de Bruckner possui para este fim e encorajam a aplicação em larga escala deste método para identificação de problemas oculares em estágios iniciais da vida, em auxílio com outros procedimentos minuciosos que podem fornecer uma maior confiabilidade para o paciente.

Neste estudo, há uma simulação do procedimento realizado no Teste de Bruckner. A base de imagens (SILVA et al., 2019) usada para extração de características e classificação das informações foi adquirida por meio de fotos tiradas por um smartphone com o flash ligado, obtendo assim o reflexo vermelho necessário para a identificação de características como a leucocoria.

## 2.2 Operadores em Imagens

Para que o processamento de uma imagem seja feito de forma ideal, há a possibilidade de otimização dos conteúdos da imagem, a fim de facilitar a detecção e/ou segmentação de pontos de interesse nas imagens. Para isso, diferentes estratégias no campo da morfologia matemática podem ser aplicadas por meio destes operadores em imagens, como as que serão apresentadas a seguir:

Gonzalez e Woods (2000) afirmam que o conceito da morfologia matemática é baseado na teoria dos conjuntos e aqui estes conjuntos são representados pelos objetos encontrados nas imagens, que fazem parte de um espaço 2D de números inteiros. Ainda, cada elemento de um conjunto pode ser representado como as coordenadas de um pixel na imagem.

Dentre as operações fundamentais, pode-se citar dilatação, erosão, abertura e fechamento. Todas essas são baseadas na morfologia matemática e são bastante usadas no

contexto do processamento de imagens. Pela natureza da representação destes conjuntos de imagens digitalmente operações binárias podem ser aplicadas, aqui conhecidas pelos operadores bitwise. Estas operações estão relacionadas com a segmentação de região de interesse por meio de máscaras, e se baseiam nos operadores lógicos NOT, AND, OR e XOR.

Rosebrook (2016) afirma também que estas operações podem ser representadas como imagens em escala de cinza. A imagem é comparada com a sua máscara pixel a pixel, determinando a presença de uma região de interesse, e possui comportamentos distintos conforme o tipo de operação que está sendo utilizada:

Tabela 2 – Condições dos operadores Bitwise em imagens.

Operadores	Condição Verdadeira
AND	Os pixels de mesma coordenada nas duas imagens possuem valor maior que zero.
OR	Um dos pixels de mesma coordenada possui valor maior que zero.
XOR	Se e somente se um dos pixels de mesma coordenada possui valor maior que zero, mas não os dois.
NOT	Inversão do valor contido no pixel da coordenada dada.

Fonte: Adaptado de Rosebrook (2016)

É importante ressaltar que possuir condição verdadeira nestes casos significa que o atual pixel da imagem que está sofrendo alteração terá seu valor mantido. Desligar o valor ou possuir uma condição falsa significa que seu valor será alterado para zero, indicando a ausência de luz e assim passará a perder sua cor.

As operações bitwise podem ser utilizadas em bibliotecas de processamento de imagem como o OpenCV em conjunto com a biblioteca NumPy, e permitem que haja foco apenas nas regiões de interesse das imagens. No exemplo deste trabalho, a segmentação da região de interesse é executada por meio destes operadores, gerando novas imagens com foco máximo na região retiniana, a partir dos valores recebidos pela imagem de máscara.

## 2.3 Redes Neurais

O conceito de Rede Neural foi introduzido primeiramente por Warren McCulloch e Walter Pitts em 1943, em um modelo matemático do neurônio biológico que contém um conjunto de  $n$  entradas multiplicadas por pesos que possuem seus valores somados e comparados a determinado limiar. Este modelo é chamado de neurônio MCP (McCulloch-Pitts) e é a primeira forma conhecida e adotada de Rede Neural. Este conceito inicial foi sendo desenvolvido e diversas arquiteturas foram introduzidas deste então, mas não se deixou de lado o seu objetivo inicial, que é de simular a forma que o cérebro usa para realizar suas tarefas (FLECK et al., 2016).

### 2.3.1 Redes Neurais Convolucionais

Redes Neurais Convolucionais são derivadas de Redes Neurais Artificiais (ANNs), contendo neurônios artificiais que se auto-otimizam a partir do aprendizado. Elas são muito populares na literatura e em desafios como o da *ImageNet* pela sua excelente performance em problemas que envolvem aprendizado profundo e são denominadas assim pelo uso da operação linear entre matrizes chamada de convolução (ALBAWI; MOHAMMED; AL-ZAWI, 2017). A sua estrutura foca em explorar conteúdo sobre o tipo de entrada específico, e CNNs são primariamente usadas no campo de reconhecimento de padrões em imagens, permitindo a adição de funcionalidades relacionadas especificamente a imagens na sua arquitetura, tornando-a adequada para tarefas neste domínio. Camadas com estruturas diferentes podem ser aplicadas à arquitetura da rede, e destas pode-se citar três tipos: camadas convolucionais, camadas de pooling e camadas completamente conectadas (O'SHEA; NASH, 2015).

O'Shea e Nash (2015) também afirmam que as funcionalidades básicas de uma CNN podem ser descritas em 4 momentos:

- A camada de entrada recebe os valores dos pixels da imagem. Imagens em RGB são representadas em matrizes de três dimensões, onde altura e largura são informações da imagem e cada dimensão especifica um canal de cor.
- Camadas convolucionais utilizam *kernels* (ou filtros) para a construção de mapas 2D de ativação. A aplicação do produto escalar entre o conjunto de pixels da matriz de entrada e o *kernel* forma as convoluções e a rede aprende *kernels* que ativam quando são apresentados a traços marcantes em posições específicas da entrada.
- Camadas de *Pooling* são usadas com o objetivo de reduzir a dimensionalidade da representação, e assim simplificar o número de parâmetros e a complexidade computacional do modelo.
- Camadas totalmente conectadas (*fully-connected*) recebem a saída da camada anterior como entrada de dados e possuem como saída a quantidade  $n$  de neurônios sendo a mesma de classes do modelo para a classificação.

CNNs, em suma, podem ser estruturadas para as necessidades que problemas relacionados a processamento de imagem exigem e sem dúvida são consideradas ferramentas de aprendizado profundo com alta performance para aplicações com este fim.

## 2.4 Deep Features

O processo de tomada de decisão no que tange a processamento e classificação de imagens exige que se tenha uma quantidade razoável de informações para auxiliar no seu viés de decisão (MA et al., 2021). Para isso, representações eficientes das imagens de entrada são fundamentais. Para isso, diferentes métodos podem ser utilizados para descrever tais características, como LBP e SFTA. Duas propriedades são importantes para definir um bom descritor de características em uma imagem: É preciso descobrir informações discriminativas nas imagens e detectar padrões repetidamente conforme as cenas sofrem alteração (ZHANG et al., 2017).

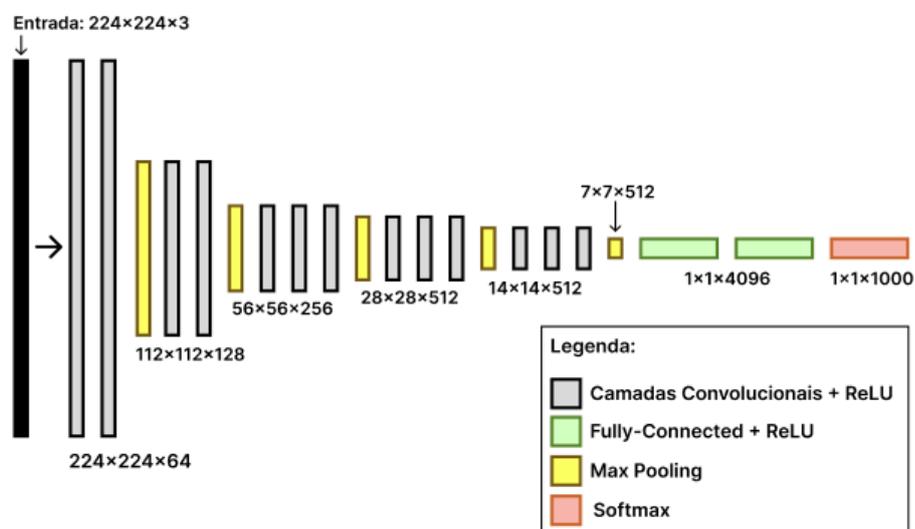
Descritores baseados em Redes Neurais Convolucionais são uma opção para se obter um mapa de informações das regiões de interesse da imagem de entrada, e possuem um desempenho satisfatório nesta tarefa devido a sua grande capacidade em reconhecimento de padrões e extração de informações (MA et al., 2021).

Dentre os objetivos desta pesquisa, está o de avaliar a utilização de aprendizado profundo com CNNs para o processo de extração e descrição de características das imagens na base de dados de imagens de reflexo vermelho. A seguir, serão apresentadas algumas das redes neurais que baseiam os descritores usados neste procedimento.

### 2.4.1 VGG-16 e VGG-19

VGGNet é um acrônimo para *Visual Geometry Group Nets* e a rede VGG-16 e sua variante VGG-19, que difere da irmã por ter dezesseis camadas convolucionais ao invés de treze, são redes neurais convolucionais introduzidas por Simonyan e Zisserman (2014), pré-treinadas e disponíveis de forma aberta que ficaram conhecidas por se destacarem na competição *ImageNet Large Scale Visual Recognition Competition* (ILSVRC) em 2014 com sua abordagem simples (BELO, 2021). A Figura 2 apresenta uma representação da arquitetura usada para a construção da rede neural VGG-16.

Figura 2 – Ilustração da arquitetura da rede VGG-16.



Fonte: Adaptado de [Simonyan e Zisserman \(2014\)](#)

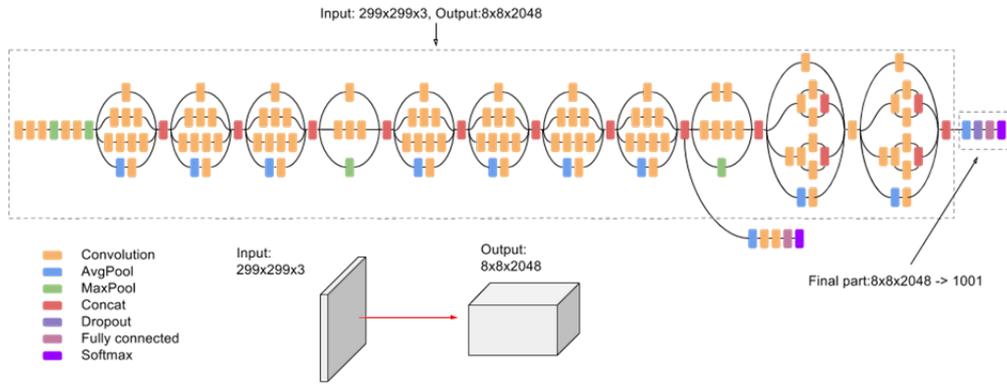
As redes foram treinadas com imagens de resolução 224x224, e as entradas passam por camadas convolucionais empilhadas com filtros de 3x3 e uso da função de ativação ReLU (*Rectified Linear Units*). Camadas de *max-pooling* são aplicadas entre as pilhas de camadas convolucionais. Ao final, são aplicadas três camadas *fully-connected*. As duas primeiras possuem 4096 canais, e a última possui 1000 camadas, no formato de classificação do desafio ILSVRC ([SIMONYAN; ZISSERMAN, 2014](#)). Para o processo de extração de *deep features*, esta última camada é descartada, sem a aplicação de pooling nas características.

No processo de configuração das redes, algumas variações foram criadas, com 11, 13, 16 e 19 camadas. As arquiteturas não só alcançaram resultados *state-of-the-art* nas tarefas de classificação do ILSVRC como foram generalizadas para aplicação em outras tarefas de reconhecimento de imagem como a extração de *deep features* a serem classificadas posteriormente ([SIMONYAN; ZISSERMAN, 2014](#)).

### 2.4.2 Inception v3

Inception v3 é uma arquitetura de Rede Neural Convolucional introduzida por [Szegedy et al. \(2016\)](#) que nasceu da necessidade de busca por redes convolucionais mais eficientes em capacidade de processamento e gasto de memória. O custo computacional da rede Inception é menor que o usado pelas redes VGGNet. A Figura 3 ilustra a arquitetura da rede Inception v3.

Figura 3 – Ilustração da arquitetura da rede Inception v3.



Fonte: (HONGKUN et al., 2020)

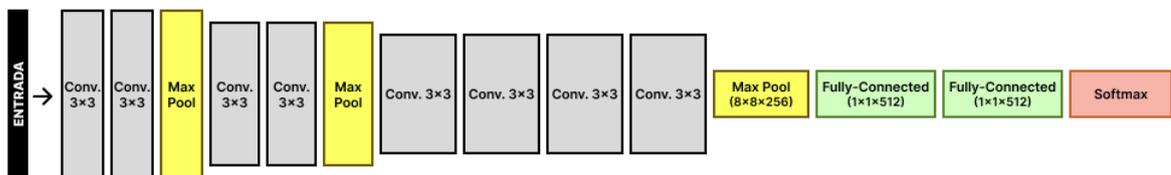
De acordo com Chagas et al. (2019), sua arquitetura contém componentes que incluem convoluções, camadas de *pooling*, concatenações, *dropouts* e camadas totalmente conectadas, além da função *softmax*, que é aplicada na camada de saída da rede e determina a probabilidade de uma imagem pertencer a uma classe e está descrita na Equação 2.1:

$$softmax(y_i) = \frac{e^{y_i}}{\sum_{j=1}^n e^{y_j}} \quad (2.1)$$

### 2.4.3 DeepLoc

*DeepLoc* é uma rede neural convolucional que foi desenvolvida por Kraus et al. (2017) a partir do objetivo de criar uma plataforma para análise automatizada de dados microscópicos de alto conteúdo, implementada por meio do *TensorFlow* (ABADI et al., 2016), uma ferramenta para desenvolvimento e treinamento de algoritmos em aprendizado de máquina. A Figura 4 apresenta uma ilustração da arquitetura da rede.

Figura 4 – Ilustração da arquitetura da rede *DeepLoc*.



Fonte: Adaptado de (KRAUS et al., 2017).

Na *DeepLoc*, a entrada é processada por meio de oito blocos convolucionais nos quais filtros são aplicados ao longo do espaço de informações de cada uma das imagens, buscando padrões invariáveis associados a uma classe específica. Três camadas de *max pooling* entre os blocos convolucionais buscam simplificar a complexidade do método. Por

fim, duas camadas completamente conectadas conectam elementos das suas camadas a todos na camada anterior, formando mais de dez milhões de parâmetros treináveis (KRAUS et al., 2017).

Estudos realizados por Kraus et al. (2017) comprovaram o desempenho da rede aplicada a problemas relacionados a microscopia de alto conteúdo, como a classificação de proteínas subcelulares, e a rede também foi usada como extrator de características, tendo a sua última camada removida assim permitindo a classificação dos 512 parâmetros gerados pela penúltima camada por outros classificadores lineares como SVM, Random Forest ou kNN, por exemplo.

## 2.5 Técnicas de Redução de Dimensionalidade

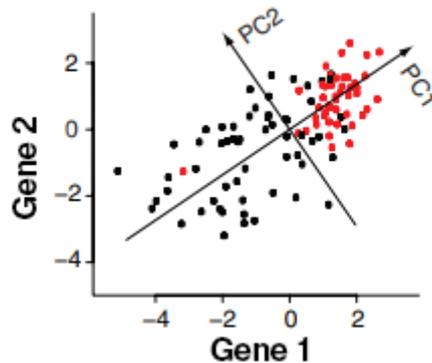
Tarefas de aprendizado de máquina requerem técnicas eficientes para fornecer bons resultados. A aplicação de métodos de redução de dimensionalidade é uma parte importante da seleção de características, que busca remover informações irrelevantes e construir um subconjunto de informações que mantenham padrões pertencentes à coleção inicial e sejam apropriadas para o treinamento adequado do modelo a ser usado (ZEBARI et al., 2020). Esta seção busca introduzir técnicas de redução aplicadas nesta pesquisa.

### 2.5.1 Análise de Componentes Principais

Análise de componentes principais é uma técnica de análise multivariada que trata a compressão dos dados em função da existência de correlação entre as variáveis medidas. A aplicação do PCA cria um novo conjunto de variáveis chamado de Componentes Principais, construído a partir da combinação linear das variáveis do conjunto original (SABIN; FERRÃO; FURTADO, 2004).

Na Figura 5, é apresentada uma ilustração do funcionamento do método. O gráfico expressa uma base dados de genes e seus níveis de expressão a partir de uma amostra de célula de câncer de mama (RINGNÉR, 2008). A Análise de Componentes Principais identifica regiões que possuem grandes variações, e o primeiro componente principal (PC1) alcança a área com a maior variação no gráfico. O segundo componente principal (PC2) alcança a segunda área com maior variação, e assim sucessivamente.

Figura 5 – Ilustração da aplicação da técnica de Análise de Componentes Principais.



Fonte: Adaptado de Ringnér (2008)

O método é considerado versátil e o mais popular em sua área, e dentre seus objetivos, está o de extrair as informações mais importantes, comprimir as informações originais mantendo informações que sejam importantes para o escopo do problema e analisar a estruturação das variáveis (ABDI; WILLIAMS, 2010).

### 2.5.2 Gain Ratio

Gain Ratio é uma técnica para seleção de atributos que leva em consideração a entropia de um determinado atributo para medir o ganho de informação do mesmo. A Equação 2.2 representa o cálculo que é feito para medir esta informação. Nela,  $H(C)$  é a entropia da classe  $C$ ,  $H(C|F_i)$  é a entropia da classe  $C$  a partir de  $F_i$  e  $H(F_i)$  é a entropia de  $F_i$ , o atributo analisado (SILVA, 2014).

$$GainRatio(C, F_i) = \frac{H(C) - H(C|F_i)}{H(F_i)} \quad (2.2)$$

## 2.6 Classificação

Um dos objetivos desta pesquisa é avaliar o uso das características extraídas com diferentes classificadores. O objetivo da classificação é a produção de um algoritmo que generalize informações contidas em treinamento para identificá-las e posteriormente classificar conjuntos de dados com rótulos desconhecidos (PADILHA; CARVALHO, 2017). Nesta seção, é feita uma breve introdução aos classificadores que foram usados para a apuração das *deep features* entregues pelos descritores.

### 2.6.1 kNN

kNN, acrônimo para *k-Nearest-Neighbors* é um algoritmo de aprendizado de máquina supervisionado popularmente aplicado em mineração de dados aprendizado de máquina (PADILHA; CARVALHO, 2017). De acordo com Zhang (2016), assume que elementos similares estão localizados próximos uns aos outros e assim classifica dados calculando a distância entre as amostras de treinamento e teste, com o objetivo de obter seus vizinhos mais próximos, delimitados por K, em um processo iterativo. Uma ilustração da técnica em formato de pseudo-códigos é apresentada na Figura 6.

Figura 6 – Pseudo-códigos para treinamento e teste do algoritmo kNN.

<p><b>Algoritmo 1: Treinamento kNN.</b></p> <p><b>Entrada:</b> conjunto de treinamento <math>T = \{x_i, y_i\}_{i=1}^n</math>, valor de <math>k</math> e uma medida de distância <math>d(\cdot, \cdot)</math></p> <p><b>Saída</b> : classificador kNN</p> <p>1 armazenar o conjunto de treinamento e o valor de <math>k</math>;</p>
<p><b>Algoritmo 2: Teste kNN.</b></p> <p><b>Entrada:</b> classificador kNN e um objeto <math>x</math> cuja classe é desconhecida</p> <p><b>Saída</b> : classe <math>y</math> atribuída a <math>x</math></p> <p>1 buscar pelos <math>k</math> objetos mais próximos a <math>x</math> no conjunto de dados de treinamento do classificador kNN informado;</p> <p>2 dentre os <math>k</math> vizinhos, determinar <math>y</math> como a classe mais frequente entre eles, resolvendo possíveis empates de maneira arbitrária;</p>

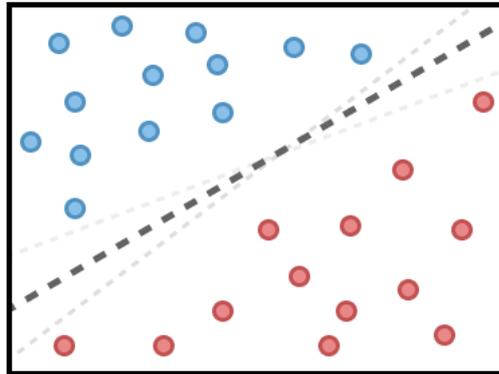
Fonte: (PADILHA; CARVALHO, 2017)

É considerado um modelo simples e versátil, relativamente fácil de implementar e pode ser usado para classificação, e regressão, porém exige uma alta quantidade de memória conforme suas variáveis aumentam, impactando no seu desempenho (ZHANG, 2016).

### 2.6.2 Máquina de Vetor de Suporte

De acordo com Carvalho (2016), Máquinas de Vetor de Suporte (SVM - Support Vector Machines) são métodos de aprendizado de máquina utilizados no reconhecimento de padrões e que buscam fornecer classificadores com um alto nível de generalização, sendo capaz de fornecer um desempenho próximo ao ótimo possível. A ideia principal é criar um hiperplano de separação que maximize a margem de distância do classificador e da amostra mais próxima de cada uma das classes. Este método pode ser visualizado na Figura 7. Das funções lineares, a escolhida é a que oferece a margem máxima entre as classes pertencentes ao plano, chamada de hiperplano ótimo.

Figura 7 – Separação de classes por meio de hiperplanos, método usado pelo algoritmo SVM.



Fonte: Adaptado de [Silva \(2020\)](#).

As SVMs são fundamentadas pela teoria do aprendizado estatístico, introduzida por [Vapnik \(1999\)](#), e têm entre suas vantagens a estabilidade da sua performance em altas dimensionalidades, evitando a modelagem de classificadores super ou sub ajustados, vertendo a um único mínimo global. Porém, há limitações intrínsecas a seu uso, como a dificuldade da interpretação a seu modelo e a necessidade do uso de estratégias decompositivas para reduzir o problema de otimização do algoritmo em larga escala, que pode ser contornado a partir de técnicas matemáticas ([LORENA; CARVALHO, 2007](#)).

### 2.6.3 Random Forest

Random Forests também são classificadores utilizados na área de mineração de dados. Neste processo, um conjunto de árvores de decisão é gerado, com informações baseadas nas respostas dos procedimentos anteriores executados em todas as árvores de decisão. Na definição proposta por [Breiman \(2001\)](#), contém uma coleção de classificadores estruturados em formato de árvore  $\{h(x, \theta_k), k = 1, \dots\}$ , onde  $\{\theta_k\}$  são vetores de características distribuídos independentemente e identicamente e cada árvore vota para a classe mais popular presente na entrada  $x$ .

Este método pode ser utilizado tanto para procedimentos de classificação quanto de regressão. Dentre as vantagens em seu uso, pode-se mencionar a sua eficiência e rapidez e a dependência de poucos parâmetros de afinamento ([CUTLER; CUTLER; STEVENS, 2012](#)).

Além disso, também é utilizado em problemas que envolvem classificação de imagens, como mencionado por [Bosch, Zisserman e Munoz \(2007\)](#). Na pesquisa, o classificador foi capaz de oferecer algumas vantagens em comparação com um classificador SVM, como a facilidade e rapidez nos processos de treino e teste, com resultados iguais ou superiores aos seus concorrentes.

### 2.6.4 Regressão Logística

Regressão Logística é um modelo de classificação de dados que se baseia na relação entre fatores de dados. A técnica permite obter probabilidades associadas a ocorrência de certo tipo de evento. O seu conceito vem da ideia de Modelos Generalizados Lineares, introduzidos por [Nelder e Wedderburn \(1972\)](#), que afirmam que uma variável dependente possui uma distribuição de probabilidade, em conjunto com uma função que relaciona a mesma ao modelo ([SOUSA, 2007](#)).

Em problemas de natureza binária como o abordado neste estudo, modelos de regressão logística possuem uma boa performance, atribuindo um índice de probabilidade associado a uma variável de interesse. Aqui, o modelo de regressão descreve relação entre variáveis associadas  $X_1, X_2, \dots, X_n$  a uma variável dependente  $Y$  que pode ser codificada como 1 ou 0 para as suas duas possíveis categorias ([SOUSA, 2007](#)).

### 2.6.5 Naïve Bayes

*Naïve Bayes* é um classificador amplamente usado para indução probabilística e busca classificar elementos associando distribuições de probabilidade a cada um dos atributos individualmente ([LANGLEY; SAGE, 1994](#)).

Baseado na teoria de Bayes, o classificador considera que todos os atributos são independentes e possuem a mesma importância, daí o termo ingênuo, do inglês *naïve*. O método classifica elementos a partir de uma taxa de probabilidade baseada na Equação 2.3 ([NEIVA, 2016](#)).

$$P[A|B] = \frac{P[B|A] * P[A]}{P[B]} \quad (2.3)$$

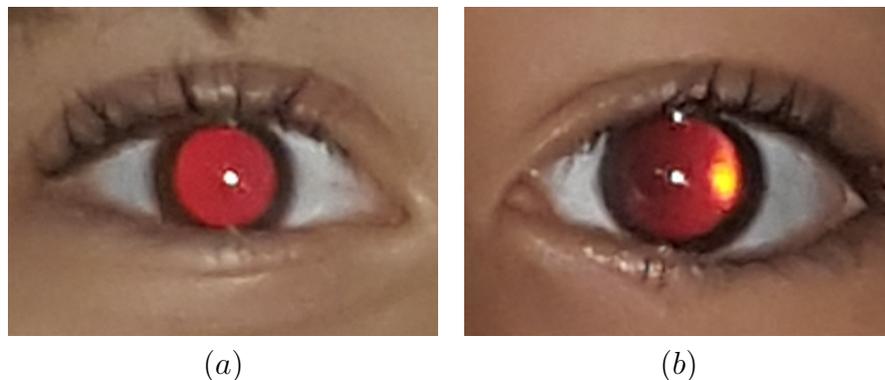
## 3 Materiais e Método

Neste capítulo, é descrita a metodologia usada nesta pesquisa e seus passos, além de uma breve apresentação sobre a base de imagens utilizada, a ferramenta em que os estudos foram executados e as métricas para avaliação dos resultados.

### 3.1 Base de Imagens

As ações deste projeto foram executadas em cima de uma base de dados obtida a partir do projeto de [Silva et al. \(2019\)](#), consistida de imagens de pacientes de uma clínica oftalmológica localizada em São Luís - MA, seguindo recomendações padrão propostas para o Teste de Bruckner, já que os pacientes não faziam uso de óculos e estavam com as pupilas dilatadas. Por meio da captura de imagens de um smartphone com a função flash ativada, simulando o procedimento com o oftalmoscópio direto, foi possível ter um resultado armazenado semelhante ao do exame para posterior estudo e classificação. Duas imagens utilizadas na base são apresentadas na Figura 8.

Figura 8 – (a) Imagem classificada com patologias e (b) imagem classificada sem patologias.



Dentre as 108 imagens da base, 88 foram separadas como de pacientes em condição normal e 30 com condições patológicas.

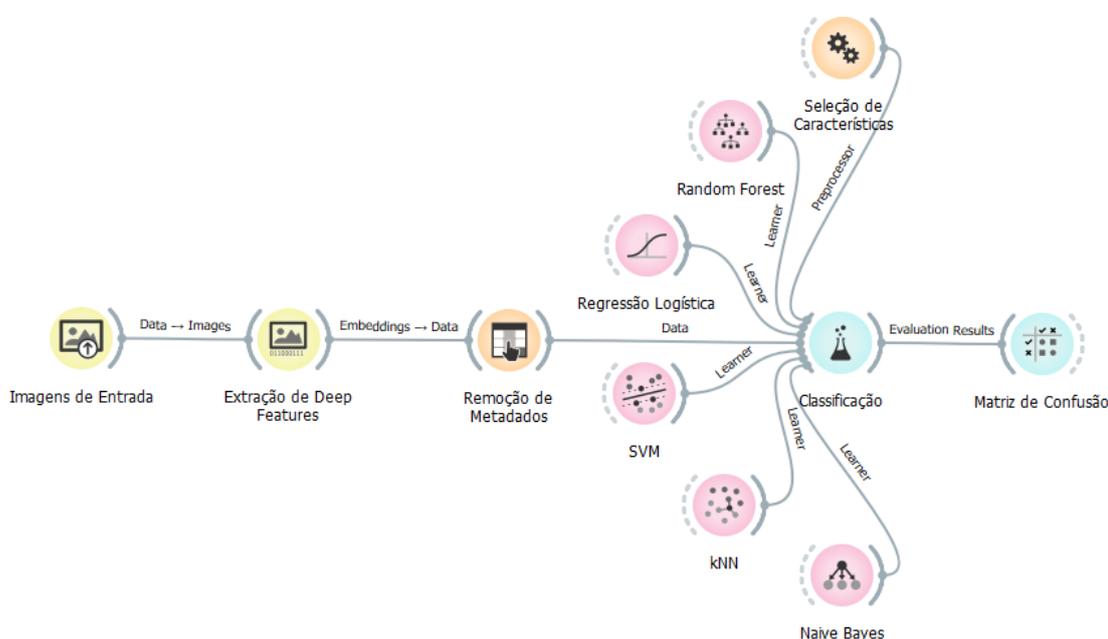
### 3.2 A Plataforma Orange

Orange é um projeto open-source desenvolvido na Universidade de Ljubljana, na Eslovênia, e é uma ferramenta de programação visual desenhada para simplificar a montagem de fluxos de análise de dados e construção de ambientes de mineração de dados a partir de uma combinação de componentes existentes ([DEMŠAR et al., 2013](#)). Oferecendo

diferentes funcionalidades de pré e pós-processamento, o programa oferece uma forma intuitiva de trabalhar com a análise de dados.

Neste trabalho, seus módulos foram utilizados para a extração de *deep features*, treinamento e validação. A versão do programa utilizada foi a 3.31. A extração das características foi realizada com diferentes redes neurais, dentre elas VGG-16, VGG-19, Inception, SqueezeNet e DeepLoc, que inferem seus dados de diferentes formas, gerando distintos vetores de características. Uma explicação detalhada sobre o funcionamento das mesmas pode ser acompanhado no Capítulo 2. A Figura 9 apresenta uma ilustração do fluxograma utilizado para este projeto.

Figura 9 – Apresentação do fluxograma criado na plataforma Orange.



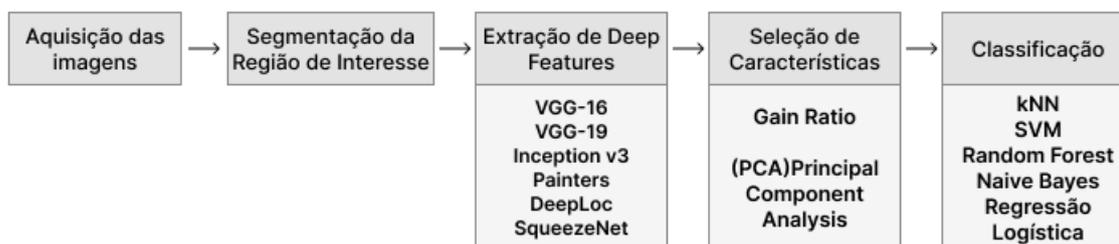
Fonte: O próprio autor.

A ferramenta abstrai os passos do tratamento dos dados de entrada em módulos chamados de *widgets*, conectados entre si, formando fluxogramas. Em cada círculo, parâmetros específicos a cada uma das técnicas são configuráveis e informações podem ser vistas a qualquer momento. Cada um desses passos é apresentado nos módulos a seguir.

### 3.3 Método Proposto

A seguir, é apresentada a metodologia usada para o desenvolvimento desta pesquisa. Em linhas gerais, foram executados os passos de pré-processamento, extração e seleção de características, além da classificação das imagens nas características aplicadas sobre a base de dados, como é apresentado na Figura 10. Além disso, técnicas, redes e classificadores usados são apresentados abaixo das caixas contendo os passos.

Figura 10 – Diagrama ilustrando os passos desta pesquisa.



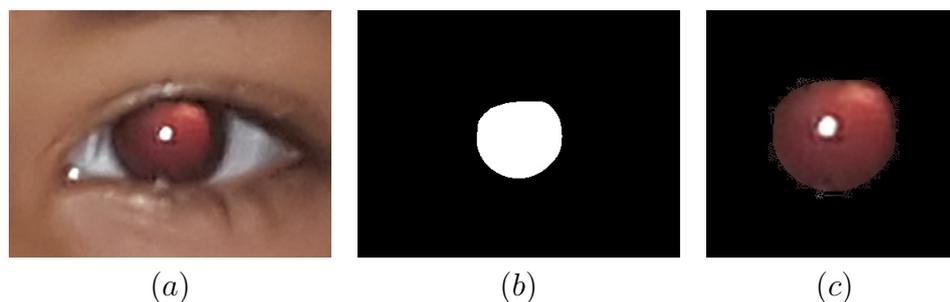
Fonte: O próprio autor.

Ao longo dos tópicos, serão mostrados o passo-a-passo, os algoritmos usados para executar o pré-processamento, as operações morfológicas entre as imagens e as máscaras delimitando regiões de interesse a serem analisadas pelos softwares, além do pipeline utilizado pelo Orange para pré-processamento, extração de deep features, treinamento e validação da base de imagens utilizada. O Orange foi escolhido como plataforma para realização dos estudos por ser uma ferramenta de programação visual completa que permitiu a modularização dos passos e por abranger todas as etapas necessárias a partir da seleção de características de forma clara e objetiva, fornecendo métricas que serão apresentadas posteriormente. A seguir, uma breve explicação sobre cada uma das fases do método aplicado.

### 3.3.1 Segmentação de imagens

Além da base original de imagens com suas características, foram usadas também outras 108 imagens contendo marcações da região de interesse para posterior tratamento. Estas imagens estão representadas em escala de preto e branco e foram extraídas a partir do método proposto por [Silva et al. \(2019\)](#), que segundo o autor, utiliza uma arquitetura de rede completamente convolucional chamada U-net. Abaixo, há um exemplo das imagens e seus formatos antes e depois da aplicação da máscara. Esta imagem pertence a classe "normal", sem a presença de patologias.

Figura 11 – (a) Imagem original com reflexo vermelho, (b) imagem de máscara contendo a região de interesse extraída a partir de (a) e (c) imagem pós-aplicação da máscara.



Operações de natureza binária, mencionadas na Seção 2.2, foram aplicadas entre as imagens originais e referências em máscara, resultando em novas imagens com a região de interesse segmentada. Além disso, as imagens passaram por um processo de normalização, mantendo um tamanho uniforme de 220x213 pixels entre si. A Figura 11 representa a imagem resultante (c) da combinação entre as imagens (a) e (b) com o uso do operador AND.

### 3.3.2 Extração de características

O processo de extração de características busca representar de forma semântica estruturas específicas em uma imagem, de forma que haja distinção entre pontos e regiões de interesse e outras áreas planas sem características que possam ser relevantes. Métodos que aplicam técnicas de aprendizado profundo possuem comprovadamente maior performance de correspondência e potencial comparados a métodos tradicionais (MA et al., 2021).

As técnicas usadas para tal são conhecidas pelo seu poder de generalização para tarefas de uso geral. Este é o caso da Rede Neural Convolutiva VGG-16 (SIMONYAN; ZISSERMAN, 2014), que para a tarefa de classificação de imagens a partir de outras bases de dados, tem a sua última camada totalmente conectada removida e suas 4096 ativações geradas pela penúltima camada descritas em forma de vetor de características, que podem ser combinadas a classificadores como SVM.

Assim, foram testadas diferentes CNNs que utilizam o mesmo método para extração de *deep features* a partir das imagens de entrada. As redes utilizadas estão disponíveis a partir da plataforma Orange, ferramenta utilizada para extração de características, seleção e classificação da base a partir das classes definidas: com e sem patologias. A partir da ferramenta, também foi possível avaliar as métricas de avaliação de desempenho do método proposto. A Tabela 3 apresenta a quantidade de características geradas por cada uma das redes usadas neste passo.

Tabela 3 – Dimensionalidade do vetor de características gerado por cada uma das redes usadas.

Rede Usada	Número de Características
VGG-16	4096
VGG-19	4096
Inception v3	2048
SqueezeNet	1000
DeepLoc	512

Fonte: O próprio autor.

### 3.3.3 Seleção de Características

O conhecimento extraído da informação de entrada é formalizado em vetores de características, que podem armazenar grandes quantidades de dados. Estes conjuntos podem apresentar alta dimensionalidade, porém nem todas as informações são realmente intrínsecas à tarefa demandada. Para isso, se faz necessária a aplicação de técnicas de redução de dimensionalidade, que buscam remover características redundantes ou irrelevantes ao problema, sem se desfazer de informações cruciais. Isso significa que mesmo com menos características que a entrada original, padrões e variância de informações essenciais para a descrição correta tendem a ser mantidas, diminuindo a capacidade de memória e tempo necessárias para seu processamento (ZEBARI et al., 2020).

Neste passo, foram extraídas características das mesmas imagens com e sem as operações de segmentação da região de interesse. As 108 imagens estavam classificadas nas classes *normal* e *patologia* e suas características foram extraídas a partir de diferentes redes neurais convolucionais.

Além destes vetores de características, alguns metadados foram gerados pelo Orange, como nome da imagem, tamanho, largura e altura. Esses dados foram ignorados para critério de avaliação. Posteriormente, técnicas de redução de características foram aplicadas neste passo, como PCA e *Gain Ratio*, apresentados na Seção 2.5. Também foi aplicado um método para remoção de características com muita repetição de valores zero, que tendem a indicar áreas fora da região de interesse da classificação. Estas técnicas foram combinadas com o uso dos classificadores e dos extratores de características e uma introdução teórica sobre as mesmas está disponível na Seção 2.5.

### 3.3.4 Treinamento e Classificação

Este segmento trata de detalhar os passos utilizados para treinar, classificar e avaliar a utilização da técnica de extração de *deep features* em conjunto com os métodos de redução de características. Cada rede usada, como descrito a seguir, foi testada com diferentes passos e estratégias de treinamento.

Foram usados diferentes métodos para o processo de classificação das características adquiridas e pré-processadas. Dentre os modelos, pode-se listar kNN, Máquina de Vetor de Suporte (SVM), Random Forest, Regressão Logística e Naive Bayes, disponíveis por meio da plataforma Orange. Todos esses classificadores estão descritos na Seção 2.6.

Além disso, cada uma das redes e métodos contou com três formas de amostragem entre treino e validação. Foram usadas as estratégias de *Cross-validation* com 5 e 10 *folds* e *Leave one out*.

Alguns dos classificadores permitem a utilização de hiper-parâmetros que auxiliam a generalização de diferentes tipos de características. A Tabela 4 indica algumas das

configurações aplicadas para os classificadores por meio da ferramenta Orange.

Tabela 4 – Parâmetros utilizados nos classificadores.

Classificador	Parâmetros utilizados
SVM	Kernel RBF, Tolerância numérica de 0,001; Limite de 100 iterações
Regressão Logística	Regularização Ridge (L2)
Random Forest	Número de árvores: 10
kNN	K = 5; Distância Euclidiana; Peso uniforme

Fonte: O próprio autor.

### 3.4 Métricas de Avaliação

Os resultados obtidos pela classificação das imagens podem ser avaliados a partir de uma Matriz de Confusão, que separa as classificações efetuadas em quatro casos:

- Verdadeiro Positivo (VP): Não há patologia e a classificação foi correta.
- Falso Positivo (FP): Não há patologia, mas a classificação foi incorreta.
- Verdadeiro Negativo (VN): Há patologia e a classificação foi correta.
- Falso Negativo (FN): Há patologia, mas a classificação foi incorreta.

Estas informações são usadas em equações que usadas como métodos estatísticos comuns em problemas de classificação e são disponibilizadas na ferramenta Orange. Tais métricas são utilizadas neste trabalho e são AUC (Área sobre a curva), CA (Acurácia de Classificação), F1-Score, Precisão, Sensibilidade, Especificidade e índice Kappa.

Sensibilidade ou *recall* (Equação 3.1) se refere a razão de imagens classificadas corretamente com patologias, avaliando a efetividade do método em uma única classe. Já a especificidade (Equação 3.2) se refere à probabilidade de imagens serem classificadas como saudáveis corretamente.

$$\text{Sensibilidade} = \frac{VP}{VP + FN} \quad (3.1)$$

$$\text{Especificidade} = \frac{VN}{VN + FP} \quad (3.2)$$

A métrica de precisão (Equação 3.3) fornece uma relação entre verdadeiros positivos e falsos positivos e a acurácia (Equação 3.4) se refere a proporção total das amostras que foram classificadas corretamente, apresentando a efetividade geral do modelo.

$$\text{Precisão} = \frac{VP}{VP + FP} \quad (3.3)$$

$$\text{Acurácia} = \frac{VP + VN}{VP + FP + VN + FN} \quad (3.4)$$

Em problemas de classificação onde o desbalanceamento entre as classes se faz presente, a acurácia pode não oferecer um retrato real do desempenho do classificador. Assim, uma métrica que serve de apoio para a avaliação é o F1-Score (Equação 3.5), que indica a relação entre precisão e sensibilidade.

$$F1 - Score = 2 * \frac{\text{precisão} * \text{sensibilidade}}{\text{precisão} + \text{sensibilidade}} \quad (3.5)$$

Além disso, a curva ROC (*Receiver Operations Characteristic*) é uma boa forma de visualizar performances de classificadores e seus limiares de decisão, sendo organizada da classificação mais positiva à mais negativa. Desta, pode ser extraída a área sobre a curva (AUC, de *Area Under Curve*), que dentre as suas vantagens, permite uma sensibilidade maior de análise de variância em relação a métrica de acurácia padrão (BRADLEY, 1997). A AUC agrega os limiares da curva ROC em um único valor. Quanto maior o valor, melhor.

## 4 Resultados e Discussão

Este capítulo apresenta seus resultados na forma de métricas de precisão dos experimentos efetuados nesta pesquisa e realiza uma análise dos dados obtidos, comparando sua performance com outros trabalhos relacionados na literatura, além de apresentar estudos de caso.

### 4.1 Apresentação dos resultados

Para avaliar comparativamente a aplicação da metodologia apresentada no Capítulo 3 e propôr um método que combine as melhores performances em relação a extração, seleção de características e classificação, foram testados os cinco classificadores propostos pela pesquisa, disponíveis no *Orange* e apresentados na Seção 2.6.

As tabelas abaixo indicam o desempenho dos classificadores combinadas aos extratores de características baseados em aprendizado profundo, separadas pelas técnicas de redução de dimensionalidade de validação cruzada com 5 e 10 folds e *Leave-one-out*. A Seção 2.5 contém uma introdução a estes métodos. Nos resultados das Tabelas 5, 6 e 7, a classificação foi realizada a partir das características extraídas nas imagens resultantes das operações para delimitação da região de interesse, ordenadas a partir da métrica de sensibilidade. Os dez melhores resultados de cada combinação estão presentes <sup>1</sup>.

Tabela 5 – Resultados usando a técnica de redução de dimensionalidade *Gain Ratio*.

Embed.	Classificador	Amostra	AUC	A	F1	P	S	E
DeepLoc	Random F.	CV-10	0.878	88.10	87.80	87.80	88.10	76.20
DeepLoc	Random F.	LOO	0.912	85.60	84.40	85.40	85.60	64.30
Inception v3	SVM	CV-5	0.904	84.70	83.90	84.10	84.70	66.20
DeepLoc	Random F.	CV-5	0.889	84.70	83.90	84.10	84.70	66.20
DeepLoc	SVM	LOO	0.897	84.70	84.40	84.20	84.70	70.60
DeepLoc	Naive Bayes	LOO	0.833	83.90	84.00	84.10	83.90	74.70
VGG-19	Naive Bayes	LOO	0.588	83.90	82.20	83.70	83.90	59.40
VGG-19	Naive Bayes	CV-10	0.716	83.10	80.20	84.50	83.10	52.50
VGG-19	Random F.	LOO	0.662	83.10	81.50	82.40	83.10	59.10
DeepLoc	Naive Bayes	CV-10	0.819	82.20	82.30	82.40	82.20	72.00

Fonte: O próprio autor.

<sup>1</sup> Legendas: LOO - Leave one out; CV-10 - Cross-validation em 10 folds; CV-5 - Cross-validation em 5 folds; A - Acurácia; P - Precisão; S - Sensibilidade; E - Especificidade.

Tabela 6 – Resultados usando a técnica de redução de dimensionalidade *PCA-10*.

<b>Embed.</b>	<b>Classificador</b>	<b>Amostra</b>	<b>AUC</b>	<b>A</b>	<b>F1</b>	<b>P</b>	<b>S</b>	<b>E</b>
DeepLoc	Regressão L.	LOO	0.899	90.70	90.50	90.50	90.70	81.40
VGG-16	Regressão L.	LOO	0.899	90.70	90.50	90.50	90.70	81.40
DeepLoc	SVM	LOO	0.944	89.00	88.30	89.10	89.00	72.10
DeepLoc	Random F.	CV-10	0.896	82.20	81.60	81.40	88.20	65.40
DeepLoc	SVM	CV-10	0.931	88.10	88.70	87.80	88.10	74.00
DeepLoc	Regressão L.	CV-10	0.921	88.10	88.10	88.10	88.10	80.60
DeepLoc	kNN	LOO	0.852	84.70	83.60	84.30	84.70	64.00
DeepLoc	Random F.	LOO	0.897	84.70	83.90	84.10	84.70	61.10
VGG-16	Random F.	LOO	0.897	84.70	83.90	84.10	84.70	61.10
DeepLoc	kNN	CV-10	0.877	83.90	82.60	83.30	83.90	61.60

Fonte: O próprio autor.

Tabela 7 – Resultados usando a técnica de redução de dimensionalidade *Remove Sparse Values*.

<b>Embed.</b>	<b>Classificador</b>	<b>Amostra</b>	<b>AUC</b>	<b>A</b>	<b>F1</b>	<b>P</b>	<b>S</b>	<b>E</b>
DeepLoc	Regressão L.	CV-5	0.898	91.50	91.30	91.40	91.50	83.90
DeepLoc	Regressão L.	LOO	0.905	90.70	90.50	90.50	90.70	83.90
DeepLoc	Regressão L.	CV-10	0.915	89.00	88.80	88.70	89.00	78.70
VGG-16	Regressão L.	LOO	0.875	88.10	87.30	88.30	88.10	69.60
DeepLoc	Random F.	LOO	0.91	87.30	86.90	86.90	87.30	73.70
DeepLoc	SVM	LOO	0.924	86.40	85.20	86.70	86.40	64.60
VGG-19	Regressão L.	LOO	0.902	86.40	85.40	86.30	86.40	66.80
DeepLoc	SVM	CV-5	0.903	85.60	84.70	85.20	85.60	66.50
DeepLoc	Random F.	CV-10	0.91	85.60	84.90	85.10	85.60	68.70
VGG-19	Regressão L.	CV-10	0.902	85.60	84.70	85.20	85.60	66.50

Fonte: O próprio autor.

Dentre os resultados com melhor média de classificação, percebe-se a forte presença do classificador Regressão Logística combinado com o descritor de características baseado na rede *DeepLoc*. Para avaliação do método e obtenção de métricas, foi usada a técnica de validação cruzada em cinco *folds*, e os resultados mostram o desempenho da combinação: foram alcançados 91,5% de acurácia, 91,4% de precisão, 91,5% de sensibilidade e 83,90% de especificidade.

A partir da presença constante dos classificadores Random Forest, Regressão Logística e SVM, foi testada também a aplicação de um *ensemble* entre os três, com o objetivo de alcançar métricas mais eficientes. Esta combinação foi utilizada com as técnicas aplicadas ao modelo de maior sensibilidade, com a extração de características baseada na rede *DeepLoc* e validação cruzada em cinco *folds*. Com isso, índices de sensibilidade, especificidade e precisão foram aumentados, e podem ser verificados na Tabela 8:

Tabela 8 – Métricas alcançadas a partir da aplicação de *ensemble* entre os classificadores Random Forest, Regressão Logística e SVM.

Embed.	Classificador	AUC	A	F1	P	S	E
DeepLoc	Regressão L./R. Forest/SVM	0.925	93.20	93.10	93.20	93.20	84.50

Fonte: O próprio autor.

## 4.2 Comparação dos resultados

Aqui, é feita uma comparação entre os dados obtidos a partir de pesquisas anteriores e o modelo proposto nesta pesquisa. Estas informações são apresentadas na tabela a seguir.

Tabela 9 – Comparação entre métodos propostos

Autor	Técnicas aplicadas	P (%)	S (%)	E (%)
( <a href="#">SILVA et al., 2019</a> )	CLCM + MHLC 5x5 e SVM	95,25	84,66	98,90
( <a href="#">PINTO, 2020</a> )	CNNs e Meta-classificadores.	90,60	70	97,77
Método atual	DeepLoc e Reg. Logística	93,20	93,20	84,50

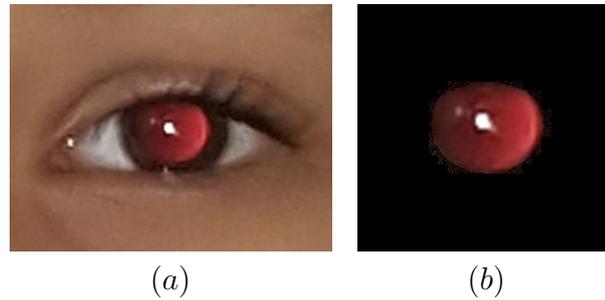
A partir das métricas avaliadas, os resultados apresentados pela pesquisa de [Silva et al. \(2019\)](#) possuem uma precisão maior, porém comparada à pesquisa de [Pinto \(2020\)](#) que também utiliza uma técnica baseada em CNNs, os resultados apresentam um índice de precisão maior. Além disso, as métricas de precisão e sensibilidade encontram-se balanceadas, mesmo com o desbalanceamento entre classes nas imagens de entrada.

## 4.3 Estudos de caso

Na seção de estudos de caso, imagens da base e suas respectivas classificações são apresentadas para uma análise guiada sobre o método de classificação proposto por esta pesquisa. As figuras abaixo representam imagens que foram classificadas respectivamente como saudável e com patologias.

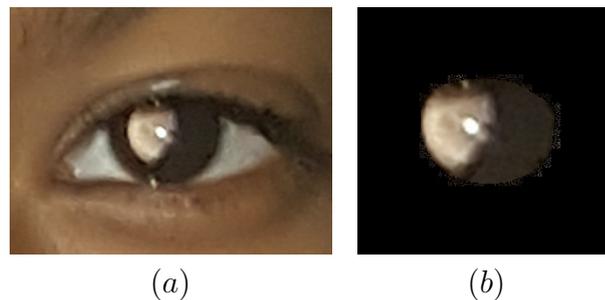
Na Figura 12, é ilustrada uma imagem que foi classificada corretamente por todos os classificadores como não patológica. Imagens de olhos saudáveis tendem a apresentar um reflexo em cor avermelhada mais uniforme e intensa ([SILVA et al., 2019](#)).

Figura 12 – Imagens (a) original e (b) com a aplicação da máscara com a região de interesse, da classe sem patologias.



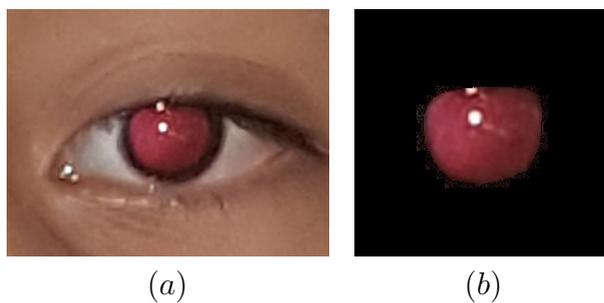
Já a Figura 13 conta com uma imagem que foi classificada corretamente como patológica por todos os classificadores. Percebe-se um reflexo esbranquiçado em parte da retina, diferindo do reflexo vermelho comum. Isso caracteriza a leucocoria. O diagnóstico de leucocoria pode ser sinal da presença de doenças como catarata, retinoblastoma e doenças da retina e vítreo (TAMURA; TEIXEIRA, 2009).

Figura 13 – Imagens (a) original e (b) com a aplicação da máscara com a região de interesse, da classe com patologias.



A Figura 14 ilustra um caso que oferece dificuldade para os classificadores por apresentar reflexo vermelho, porém em menor intensidade que na Figura 12. A baixa intensidade do reflexo também pode ser um sinal da presença de patologias. Usando o descritor de características baseado na rede *DeepLoc* e a técnica de validação cruzada de 5 *folds*, três dos cinco classificadores propostos definiu a imagem como pertencendo à classe normal. O classificador baseado em Regressão Logística classificou corretamente a presença de patologias.

Figura 14 – Imagens (a) original e (b) com a aplicação da máscara com a região de interesse, da classe com patologia e que gerou classificações falhas.



## 5 Conclusão

Esta pesquisa se propôs a desenvolver uma metodologia para a detecção de patologias da visão a partir de imagens de reflexo vermelho, utilizando deep features. A base de imagens utilizada para a execução da pesquisa foi obtida em São Luís - MA e é dividida em classes *normal* e *problema*. Junto com a base, foram disponibilizadas imagens contendo máscaras delimitando regiões de interesse das respectivas imagens contidas na base original. Tanto as imagens originais quanto as com máscaras delimitando suas respectivas regiões de interesse foram obtidas a partir da pesquisa de [Silva et al. \(2019\)](#). Pode-se esclarecer o procedimento adotado em três grandes etapas, descritas a seguir.

A primeira foi o pré-processamento, onde foram utilizadas técnicas para oclusão de elementos fora da região de interesse nas imagens obtidas anteriormente. Foram aplicadas operações binárias para a remoção de pixels indiferentes à incorporação de características, que delimitam por exemplo regiões periféricas ao olho do paciente.

Já para a extração de deep features, foram utilizados descritores de características baseados em Redes Neurais Convolucionais. Descritores de características baseados em CNNs são uma opção para se obter um mapa de informações das regiões de interesse da imagem de entrada, e possuem um desempenho satisfatório nesta tarefa devido a sua grande capacidade em reconhecimento de padrões e extração de informações ([MA et al., 2021](#)). Também foram aplicadas técnicas para redução de dimensionalidade, que buscam remover características redundantes ou irrelevantes ao contexto do problema sem perder características essenciais, tornando seu processamento mais rápido e eficiente ([ZEBARI et al., 2020](#)).

Para a classificação, foram executados e testados diferentes métodos como Regressão Logística e Random Forest em experimentos realizados usando validação cruzada e *Leave one out*, obtendo cerca de 91% de sensibilidade combinando o descritor de características baseado na rede neural *DeepLoc* e o classificador Regressão Logística usando validação cruzada. Posteriormente, com a aplicação de um *ensemble* entre os classificadores Random Forest, Regressão Logística e SVM índices maiores foram alcançados, com 93% de sensibilidade. Também pode ser encontrada na Seção [3.3.3](#) uma abordagem sobre as técnicas utilizadas.

Estes resultados validam a eficácia do uso de *deep features* para descrição de características em imagens e corroboram a importância do uso de CNNs para tal. Os resultados são comparáveis aos obtidos nas pesquisas anteriores relacionadas a esta base de imagens, mencionadas na Seção [1.4](#), e evidenciam a importância de estudos mais detalhados nesta área de atuação. Há margem para melhora e mais técnicas podem ser

empregadas a fim de oferecer uma melhor acurácia na detecção, mas os resultados já são promissores e indicam também como este modelo pode contribuir para a produção de uma plataforma auxiliar para a detecção de patologias da visão, servindo de apoio ao profissional oftalmológico que é quem pode validar o resultado e oferecer um diagnóstico mais apurado.

Além disso, é importante mencionar como a ferramenta *Orange* foi de grande valia para a aplicação dos estudos, servindo como uma plataforma intuitiva e com alta parametrização, abordando grande parte dos conceitos apresentados nesta pesquisa. Esta ferramenta, além de ser *open-source*, democratiza o acesso a métodos de processamento e mineração de dados e pode ser um bom ponto de partida para aqueles que não possuem afinidade com o código ou que desenvolvem tarefas que podem ser posteriormente apresentadas de uma forma visual e didática a todos.

## 5.1 Trabalhos futuros

Concluindo este trabalho e reiterando o que foi descrito anteriormente, há espaço para melhora e alguns aspectos podem ser levados em consideração para seguir com os objetivos desta pesquisa mais adiante. Dentre pontos importantes, pode-se considerar:

- Buscar formas de aplicar o uso de data augmentation a fim de tornar a base maior. A curta dimensionalidade da base foi um dos fatores limitantes e incrementá-la pode significar um melhor espaço de amostragem para treinamento do modelo;
- Utilizar outras Redes Neurais para extração de características, já que o trabalho se limitou a usar as disponíveis na plataforma Orange e mesmo assim apresentou resultados satisfatórios. É importante considerar que as Redes Neurais utilizadas não foram criadas para este fim e mesmo assim auxiliaram no processo de reconhecimento de padrões e descrição de características;
- Avaliar a utilização de mais técnicas de seleção de características;
- Analisar o uso de diferentes classificadores para os dados obtidos na pesquisa, além de alternativas que ofereçam mais sensibilidade aos dados, como otimização de hiperparâmetros.
- Aplicar técnicas de pós-processamento de dados, com o objetivo de recuperar possíveis classificações em que resultados incorretos tenham sido alcançados inicialmente, porém com um limiar tênue de decisão.

## Referências

- ABADI, M.; BARHAM, P.; CHEN, J.; CHEN, Z.; DAVIS, A.; DEAN, J.; DEVIN, M.; GHEMAWAT, S.; IRVING, G.; ISARD, M. et al. {TensorFlow}: a system for {Large-Scale} machine learning. In: *12th USENIX symposium on operating systems design and implementation (OSDI 16)*. [S.l.: s.n.], 2016. p. 265–283. Citado na página 23.
- ABDI, H.; WILLIAMS, L. J. Principal component analysis. *Wiley interdisciplinary reviews: computational statistics*, Wiley Online Library, v. 2, n. 4, p. 433–459, 2010. Citado na página 25.
- AGUIAR, A. S. C. d. Validação de tecnologia para avaliação do teste do reflexo vermelho. 2010. Citado 3 vezes nas páginas 13, 17 e 18.
- ALBAWI, S.; MOHAMMED, T. A.; AL-ZAWI, S. Understanding of a convolutional neural network. In: IEEE. *2017 international conference on engineering and technology (ICET)*. [S.l.], 2017. p. 1–6. Citado na página 20.
- BELO, L. M. *Reconhecimento de patologias em guias de solicitações médicas utilizando aprendizado de máquina*. Tese (Doutorado) — Mestrado em Sistemas de Informação e Gestão do Conhecimento, 2021. Citado na página 21.
- BOSCH, A.; ZISSERMAN, A.; MUNOZ, X. Image classification using random forests and ferns. In: IEEE. *2007 IEEE 11th international conference on computer vision*. [S.l.], 2007. p. 1–8. Citado na página 27.
- BRADLEY, A. P. The use of the area under the roc curve in the evaluation of machine learning algorithms. *Pattern recognition*, Elsevier, v. 30, n. 7, p. 1145–1159, 1997. Citado na página 35.
- BREIMAN, L. Random forests. *Machine learning*, Springer, v. 45, n. 1, p. 5–32, 2001. Citado na página 27.
- CARVALHO, V. P. d. Discriminação de população por meio de inteligência computacional. Universidade Federal de Viçosa, 2016. Citado na página 26.
- CHAGAS, J. E. L. P. et al. Estudo para a avaliação em tempo real do grau de retinopatia diabética e risco de edema macular. 2019. Citado na página 23.
- CRUZ, A. *Atenção Primária à Saúde no Brasil*. [S.l.]: Conselho Nacional de Secretários de Saúde, 2008. 4-9 p. <<https://bvsmms.saude.gov.br/bvs/periodicos/consensus35.pdf>>. Citado na página 13.
- CUTLER, A.; CUTLER, D. R.; STEVENS, J. R. Random forests. In: *Ensemble machine learning*. [S.l.]: Springer, 2012. p. 157–175. Citado na página 27.
- DEMŠAR, J.; CURK, T.; ERJAVEC, A.; GORUP, Č.; HOČEVAR, T.; MILUTINOVIČ, M.; MOŽINA, M.; POLAJNAR, M.; TOPLAK, M.; STARIČ, A. et al. Orange: data mining toolbox in python. *the Journal of machine Learning research*, JMLR. org, v. 14, n. 1, p. 2349–2353, 2013. Citado na página 29.

- FLECK, L.; TAVARES, M. H. F.; EYNG, E.; HELMANN, A. C.; ANDRADE, M. d. M. Redes neurais artificiais: Princípios básicos. *Revista Eletrônica Científica Inovação e Tecnologia*, v. 1, n. 13, p. 47–57, 2016. Citado na página 19.
- GONZALEZ, R. C.; WOODS, R. E. *Processamento de imagens digitais*. [S.l.]: Editora Blucher, 2000. Citado na página 18.
- HONGKUN, Y.; CHEN, C.; XIANZHI, D.; YEQING, L.; ABDULLAH, R.; LE, H.; PENGCHONG, J.; FAN, Y.; FREDERICK, L.; JAEYOUN, K.; JING, L. *TensorFlow Model Garden*. 2020. <<https://github.com/tensorflow/models/tree/archive/research/inception>>. Citado na página 23.
- HOPKER, L. M.; ROSSETO, J. D.; CARVALHO, V. G.; NAKANAMI, C. R. Teste do reflexo vermelho. *Grupo de Trabalho em Oftalmologia Pediátrica*, Sociedade Brasileira de Pediatria, 2018. Citado 2 vezes nas páginas 17 e 18.
- JALIS, M.; ASHFAQ, M. W.; IMDAD, A. Use of bruckner test for the detection of significant refractive errors in children. *Journal of Rawalpindi Medical College*, v. 19, n. 3, 2015. Citado na página 18.
- KHEDEKAR, A.; DEVARAJAN, B.; RAMASAMY, K.; MUTHUKKARUPPAN, V.; KIM, U. Smartphone-based application improves the detection of retinoblastoma. *Eye*, v. 33, p. 1, 01 2019. Citado na página 13.
- KRAUS, O. Z.; GRYS, B. T.; BA, J.; CHONG, Y.; FREY, B. J.; BOONE, C.; ANDREWS, B. J. Automated analysis of high-content microscopy data with deep learning. *Molecular systems biology*, v. 13, n. 4, p. 924, 2017. Citado 2 vezes nas páginas 23 e 24.
- LANGLEY, P.; SAGE, S. Induction of selective bayesian classifiers. In: *Uncertainty Proceedings 1994*. [S.l.]: Elsevier, 1994. p. 399–406. Citado na página 28.
- LORENA, A. C.; CARVALHO, A. C. D. Uma introdução às support vector machines. *Revista de Informática Teórica e Aplicada*, v. 14, n. 2, p. 43–67, 2007. Citado na página 27.
- MA, J.; JIANG, X.; FAN, A.; JIANG, J.; YAN, J. Image matching from handcrafted to deep features: A survey. *International Journal of Computer Vision*, Springer, v. 129, n. 1, p. 23–79, 2021. Citado 3 vezes nas páginas 21, 32 e 41.
- NEIVA, M. B. *Métodos de pré-processamento de texturas para otimizar o reconhecimento de padrões*. Tese (Doutorado) — Universidade de São Paulo, 2016. Citado na página 28.
- NELDER, J. A.; WEDDERBURN, R. W. Generalized linear models. *Journal of the Royal Statistical Society: Series A (General)*, Wiley Online Library, v. 135, n. 3, p. 370–384, 1972. Citado na página 28.
- NETOR, J. B.; UMBELINO, C. C. Censo do conselho brasileiro de oftalmologia 2021. *Conselho Brasileiro de Oftalmologia*, 2021. Disponível em: <[https://cbo.net.br/2020/admin/docs\\_upload/034327CensoCbo2021.pdf](https://cbo.net.br/2020/admin/docs_upload/034327CensoCbo2021.pdf)>. Citado na página 14.
- O'SHEA, K.; NASH, R. An introduction to convolutional neural networks. *arXiv preprint arXiv:1511.08458*, 2015. Citado na página 20.

- PADILHA, V. A.; CARVALHO, A. Mineração de dados em python. *Instituto de Ciências Matemáticas e de Computação da Universidade de São Paulo*, 2017. Citado 2 vezes nas páginas 25 e 26.
- PINTO, A. M. Detecção de patologias da visão em imagens de reflexo vermelho utilizando deep learning. Universidade Federal do Maranhão, 2020. Departamento de Informática/CCET. Citado 3 vezes nas páginas 9, 15 e 38.
- RINGNÉR, M. What is principal component analysis? *Nature biotechnology*, Nature Publishing Group, v. 26, n. 3, p. 303–304, 2008. Citado 2 vezes nas páginas 24 e 25.
- ROSEBROOK, A. *Practical Python and OpenCV+ Case Studies: An Introductory, Example Driven Guide to Image Processing and Computer Vision*. [S.l.]: PyImageSearch, 2016. Citado na página 19.
- SABIN, J. G.; FERRÃO, M. F.; FURTADO, J. C. Análise multivariada aplicada na identificação de fármacos antidepressivos. parte ii: Análise por componentes principais (pca) e o método de classificação simca. *Revista Brasileira de Ciências Farmacêuticas*, SciELO Brasil, v. 40, p. 387–396, 2004. Citado na página 24.
- SILVA, I. F. S. d. et al. Detecção automática da presença de patologia na visão baseada em imagens do teste de brückner. Universidade Federal do Maranhão, 2019. Departamento de Informática/CCET. Disponível em: <<https://tedebc.ufma.br/jspui/handle/tede/tede/2592>>. Citado 6 vezes nas páginas 15, 18, 29, 31, 38 e 41.
- SILVA, L. O. L. A. *Classificação visual de mudas de plantas ornamentais: análise da eficácia de técnicas de seleção de atributos*. Tese (Doutorado) — Universidade de São Paulo, 2014. Citado na página 25.
- SILVA, M. G. Classificação automática de patologias foliares na cultura do milho usando matriz de coocorrência e máquina de vetor de suporte. Universidade Federal do Maranhão, 2020. Departamento de Informática/CCET. Citado na página 27.
- SIMONYAN, K.; ZISSERMAN, A. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014. Citado 3 vezes nas páginas 21, 22 e 32.
- SOUSA, C. A. d. *Teoria de conjuntos fuzzy e regressão logística na tomada de decisão para realização de cintilografia das paratiróides*. Tese (Doutorado) — Universidade de São Paulo, 2007. Citado na página 28.
- SZEGEDY, C.; VANHOUCKE, V.; IOFFE, S.; SHLENS, J.; WOJNA, Z. Rethinking the inception architecture for computer vision. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. [S.l.: s.n.], 2016. p. 2818–2826. Citado na página 22.
- TAMURA, M.; TEIXEIRA, L. F. Leukocoria and the red reflex test. *Einstein (São Paulo)*, v. 7, 2009. Citado 2 vezes nas páginas 14 e 39.
- VAPNIK, V. *The nature of statistical learning theory*. [S.l.]: Springer science & business media, 1999. Citado na página 27.

ZEBARI, R.; ABDULAZEEZ, A.; ZEEBAREE, D.; ZEBARI, D.; SAEED, J. A comprehensive review of dimensionality reduction techniques for feature selection and feature extraction. *Journal of Applied Science and Technology Trends*, v. 1, n. 2, p. 56–70, 2020. Citado 3 vezes nas páginas [24](#), [33](#) e [41](#).

ZHANG, X.; YU, F. X.; KARAMAN, S.; CHANG, S.-F. Learning discriminative and transformation covariant local feature detectors. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. [S.l.: s.n.], 2017. p. 6818–6826. Citado na página [21](#).

ZHANG, Z. Introduction to machine learning: k-nearest neighbors. *Annals of translational medicine*, AME Publications, v. 4, n. 11, 2016. Citado na página [26](#).