



INSTITUTO FEDERAL GOIANO
CAMPUS URUTAÍ
NÚCLEO DE INFORMÁTICA
CURSO DE SISTEMAS DE INFORMAÇÃO

VINÍCIUS BARBOSA DE SOUZA

**SEGMENTAÇÃO DE MICROANEURISMAS EM
IMAGENS DE RETINA USANDO
APRENDIZADO DE MÁQUINA**

Urutaí, 25 de julho de 2023

INSTITUTO FEDERAL GOIANO

NÚCLEO DE INFORMÁTICA
SISTEMAS DE INFORMAÇÃO

VINÍCIUS BARBOSA DE SOUZA

**SEGMENTAÇÃO DE MICROANEURISMAS EM
IMAGENS DE RETINA USANDO
APRENDIZADO DE MÁQUINA**

Monografia apresentada ao Núcleo de Informática, curso de Sistemas de Informação, do Instituto Federal Goiano, como parte das exigências para obtenção do título de Bacharel em Sistemas de Informação.

Orientador(a):
Cristiane de Fátima dos Santos Cardoso

Urutaí, 25 de julho de 2023

Sistema desenvolvido pelo ICMC/USP
Dados Internacionais de Catalogação na Publicação (CIP)
Sistema Integrado de Bibliotecas - Instituto Federal Goiano

S729s Souza, Vinícius Barbosa de
SEGMENTAÇÃO DE MICROANEURISMAS EM IMAGENS DE
RETINA USANDO APRENDIZADO DE MÁQUINA / Vinícius
Barbosa de Souza; orientadora Cristiane de Fátima
dos Santos Cardoso. -- Urutaí, 2023.
42 p.

TCC (Graduação em Sistemas de Informação) --
Instituto Federal Goiano, Campus Urutaí, 2023.

1. Diabetes. 2. Inteligência Artificial. 3.
Microaneurismas. 4. Retina. 5. Deep Learning. I.
Cardoso, Cristiane de Fátima dos Santos , orient.
II. Título.

TERMO DE CIÊNCIA E DE AUTORIZAÇÃO PARA DISPONIBILIZAR PRODUÇÕES TÉCNICO-CIENTÍFICAS NO REPOSITÓRIO INSTITUCIONAL DO IF GOIANO

Com base no disposto na Lei Federal nº 9.610, de 19 de fevereiro de 1998, AUTORIZO o Instituto Federal de Educação, Ciência e Tecnologia Goiano a disponibilizar gratuitamente o documento em formato digital no Repositório Institucional do IF Goiano (RIIF Goiano), sem ressarcimento de direitos autorais, conforme permissão assinada abaixo, para fins de leitura, download e impressão, a título de divulgação da produção técnico-científica no IF Goiano.

IDENTIFICAÇÃO DA PRODUÇÃO TÉCNICO-CIENTÍFICA

- | | |
|--|---|
| <input type="checkbox"/> Tese (doutorado) | <input type="checkbox"/> Artigo científico |
| <input type="checkbox"/> Dissertação (mestrado) | <input type="checkbox"/> Capítulo de livro |
| <input type="checkbox"/> Monografia (especialização) | <input type="checkbox"/> Livro |
| <input checked="" type="checkbox"/> TCC (graduação) | <input type="checkbox"/> Trabalho apresentado em evento |

Produto técnico e educacional - Tipo:

Nome completo do autor:

Vinícius Barbosa de Souza

Matrícula:

2018101202010145

Título do trabalho:

SEGMENTAÇÃO DE MICROANEURISMAS EM IMAGENS DE RETINA USANDO APRENDIZADO DE MÁQUINA

RESTRIÇÕES DE ACESSO AO DOCUMENTO

Documento confidencial: Não Sim, justifique:

Informe a data que poderá ser disponibilizado no RIIF Goiano: 11 /08 /2023

O documento está sujeito a registro de patente? Sim Não

O documento pode vir a ser publicado como livro? Sim Não

DECLARAÇÃO DE DISTRIBUIÇÃO NÃO-EXCLUSIVA

O(a) referido(a) autor(a) declara:

- Que o documento é seu trabalho original, detém os direitos autorais da produção técnico-científica e não infringe os direitos de qualquer outra pessoa ou entidade;
- Que obteve autorização de quaisquer materiais inclusos no documento do qual não detém os direitos de autoria, para conceder ao Instituto Federal de Educação, Ciência e Tecnologia Goiano os direitos requeridos e que este material cujos direitos autorais são de terceiros, estão claramente identificados e reconhecidos no texto ou conteúdo do documento entregue;
- Que cumpriu quaisquer obrigações exigidas por contrato ou acordo, caso o documento entregue seja baseado em trabalho financiado ou apoiado por outra instituição que não o Instituto Federal de Educação, Ciência e Tecnologia Goiano.

Urutaí

24 /07 /2023

gov.br Documento assinado digitalmente
VINICIUS BARBOSA DE SOUZA
Data: 24/07/2023 19:43:37-0300
Verifique em <https://validar.iti.gov.br>

Local

Data

Assinat

gov.br Documento assinado digitalmente
CRISTIANE DE FATIMA DOS SANTOS CARD
Data: 24/07/2023 20:35:10-0300
Verifique em <https://validar.iti.gov.br>

autorais

Ciente e de acordo:

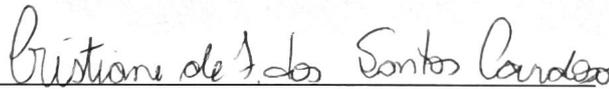
Assinatura do(a) orientador(a)

VINÍCIUS BARBOSA SOUZA

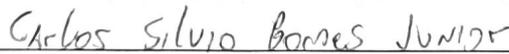
**SEGMENTAÇÃO DE MICROANEURISMAS EM
IMAGENS DE RETINA USANDO APRENDIZADO DE
MÁQUINA**

Monografia, defendida por Vinícius Barbosa Souza, apresentado ao Instituto Federal de Educação, Ciência e Tecnologia Goiano, como parte das exigências para a obtenção do título de Bacharel em Sistemas de Informação, aprovados pela banca examinadora.

COMISSÃO EXAMINADORA



Profa. Dra. Cristiane de Fátima dos Santos Cardoso
Orientadora



Prof. Esp. Carlos Silvio Gomes Júnior
Avaliador



Prof. Me. Gabriel da Silva Vieira
Avaliador

Urutá (GO), 21 de junho de 2023.

AGRADECIMENTOS

Agradeço a todas as pessoas que contribuíram para a conclusão desta monografia e que tornaram esta jornada possível:

Primeiramente, gostaria de expressar minha sincera gratidão a minha orientadora, Cristiane de Fátima Dos Santos Cardoso, por seu apoio, orientação e paciência ao longo deste processo. Seu conhecimento e dedicação foram fundamentais para o desenvolvimento deste trabalho, e sou grato por sua disponibilidade em compartilhar seu tempo e expertise.

Agradeço também aos demais professores e membros do corpo docente do Núcleo de Informática que contribuíram com seus valiosos insights e sugestões durante a apresentação de seminários, discussões em sala de aula e avaliações ao longo dos anos. Seu comprometimento com o ensino e a pesquisa acadêmica me inspirou e moldou minha compreensão sobre o assunto abordado nesta monografia.

Não posso deixar de agradecer aos meus colegas de classe e amigos, que proporcionaram um ambiente de apoio e incentivo. Seus debates e discussões enriqueceram minha compreensão e me ajudaram a aprimorar minhas ideias e argumentos.

Agradeço aos meus familiares pelo constante apoio e encorajamento ao longo desta jornada acadêmica. Sua presença, amor e incentivo foram fundamentais para minha perseverança e sucesso.

Por fim, gostaria de agradecer a todos aqueles que direta ou indiretamente contribuíram para a conclusão desta monografia, seja por meio de discussões acadêmicas, indicação de referências bibliográficas, suporte técnico ou qualquer outra forma de ajuda. Seu apoio foi inestimável e sou grato por cada contribuição.

A todos vocês, meu sincero agradecimento por serem parte fundamental deste trabalho e por terem acreditado em mim ao longo dessa jornada.

Atualmente, os oftalmologistas utilizam imagens como uma etapa de triagem, capaz de identificar várias doenças oculares degenerativas, por exemplo, o glaucoma e a retinopatia diabética. Alteração das estruturas fisiológicas como disco óptico, mácula e vasos sanguíneos podem indicar o início de uma enfermidade ocular. Neste trabalho o objetivo é realizar a segmentação dos microaneurismas em imagens da retina extraídas do banco de dados publicamente disponível, chamado DIARETDB1. Para atingir este objetivo, são utilizados Algoritmos de Aprendizado de Máquina e como resultado observa-se uma accuracy de 96,2% na segmentação dos microaneurismas. A segmentação pode ser entendida como a separação de um determinado elemento do restante da imagem.

Palavras-chave: Diabetes. Inteligência Artificial. Microaneurismas. Retina.

LISTA DE FIGURAS

1.1	(a) Reflexo no centro do vaso sanguíneo e irregularidades no fundo. (b) Hemorragias e manchas algodinosas. (c) Exsudatos duros e diversos microaneurismas - pontos vermelhos. (d) Retina com diversos sinais de retinopatia diabética . . .	9
2.1	Retinopatia diabética proliferativa e não proliferativa.	13
2.2	Retinógrafo Modelo TRC-NW8 Non-Mydriatic Retinal Camera.	14
2.3	Níveis de gravidade da retinopatia e alterações observáveis ao exame de fundo de olho.	14
2.4	Microaneurisma, hemorragia e exsudato duro.	16
2.5	Redes Neurais artificiais.	17
2.6	Neurônio artificial.	18
2.7	Convolução.	19
2.8	Arquitetura da U-Net.	21
3.1	Funções não lineares que são frequentemente utilizadas em arquiteturas de aprendizado profundo.	25
3.2	<i>Pooling</i>	25
3.3	Max Pooling.	26
3.4	Max Pooling.	26
3.5	<i>Dropout</i>	27
3.6	<i>Softmax</i>	27
3.7	Funcionamento do <i>Keras</i>	30
3.8	Arquitetura de rede proposta	33
4.1	Etapas.	35
4.2	Resultados obtidos pela rede.	35

1	INTRODUÇÃO	8
2	Fundamentos Teóricos	12
2.1	Aspectos em oftalmologia	12
2.1.1	Retinopatia diabética	12
2.1.2	Avaliação da retinografia	13
2.1.3	Microaneurisma	15
2.2	Redes neurais artificiais	16
2.2.1	Redes neurais convolucionais	17
2.2.2	Arquiteturas de redes neurais convolucionais para segmentação e reconhecimento de imagens	19
3	Materiais e métodos	22
3.1	Métricas	22
3.2	Características no projeto de arquitetura de rede	24
3.3	Ferramentas	28
3.3.1	Tensor Flow	29
3.3.2	Keras	30
3.3.3	Spyder	31
3.4	Base de dados	31
3.5	Arquitetura da rede proposta	32
4	Resultados e discussões	34
4.1	Etapas do processo de detecção de microaneurismas	34
4.2	Resultado	35
	CONCLUSÃO	37
4.3	Trabalhos Futuros	37

CAPÍTULO 1

INTRODUÇÃO

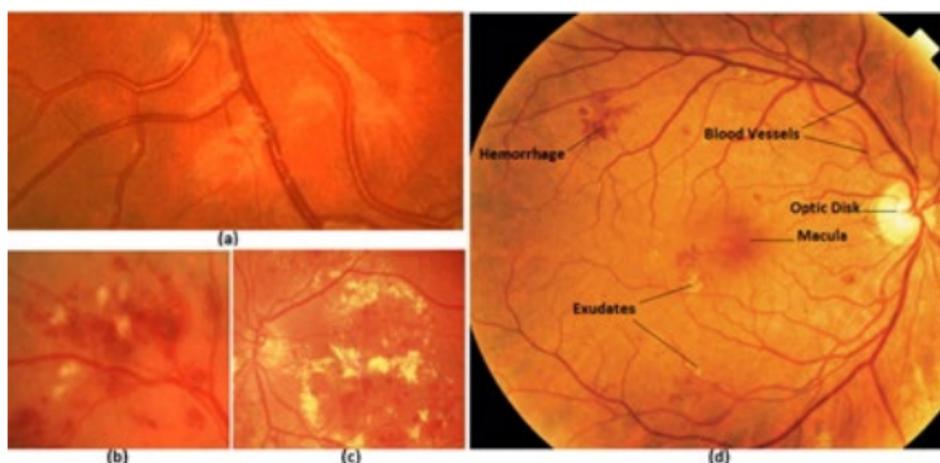
De acordo com a *International Agency for the Prevention of Blindness* (IAPB, 2013), a diabetes *mellitus* está se tornando uma epidemia global. Em 2014, havia aproximadamente 422 milhões de pessoas (8,5% da população adulta mundial) vivendo com diabetes e, portanto, sujeitos ao desenvolvimento da retinopatia diabética. A Organização Mundial da Saúde (OMS, 2021), cita que pelo menos 2,2 bilhões de pessoas têm deficiência visual ou cegueira, das quais pelo menos 1 bilhão têm uma deficiência visual que poderia ter sido evitada ou tratada e uma das principais causas de cegueira em todo o mundo é a retinopatia diabética decorrente da diabetes. Mesmo pacientes com bom controle da doença acabam desenvolvendo algum grau de retinopatia ao longo dos anos e por isso devem estar em constante monitoramento, recomenda-se pelo menos um exame anual (HENDRICK et al., 2015), (MOOKIAH et al., 2013a), (MOOKIAH et al., 2013b), (DORION, 2002).

A cegueira pode ocorrer em pacientes diabéticos devido a danos nos vasos sanguíneos dentro da retina, na parte de trás do olho, tal condição geralmente afeta ambos os olhos, podendo levar à perda de visão se não for tratada. De acordo com a *International Agency for the Prevention of Blindness* (IAPB, 2013), cerca de 1 em cada 3 pessoas que vivem com diabetes têm algum grau de retinopatia diabética e 1 em cada 10 desenvolverá uma forma da doença que ameaça a visão.

Assim, é importante um diagnóstico preciso, confiável e ao mesmo tempo de baixo custo. Uma alternativa altamente viável e eficiente é o diagnóstico por meio de processamento de imagens digitais. A imagem de fundo de olho usada por especialistas em retina é a retinografia e como o número de diabéticos tem crescido consideravelmente, cresce também a demanda por

este tipo de exame. A Figura 1.1 (d) apresenta uma retinografia, com destaque para estruturas fisiológicas como o disco óptico, a mácula e os vasos sanguíneos e também alguns sinais patológicos como o exsudatos e hemorragia. Figura 1.1 também mostra alguns fragmentos contendo sinais oculares perceptíveis, como variações na aparência, na forma, e na direção dos vasos sanguíneos. Os oftalmologistas têm usado as imagens da retina para monitorar enfermidades como: pequenos sangramentos, presença de microaneurismas, exsudatos, descolamento da retina e tumores. Muitos destes sinais podem indicar o início da retinopatia diabética e glaucoma (NIEMEIJER et al., 2004)(MOOKIAH et al., 2013b).

Figura 1.1: (a) Reflexo no centro do vaso sanguíneo e irregularidades no fundo. (b) Hemorragias e manchas algodinosas. (c) Exsudatos duros e diversos microaneurismas - pontos vermelhos. (d) Retina com diversos sinais de retinopatia diabética



Fonte: (FRAZ et al., 2012)

Um microaneurisma é um sinal clássico de retinopatia diabética e pode ser visto na forma de um pequeno ponto vermelho em imagens de retinografias, trata-se de lesões em forma sacular que surgem na parede de vasos capilares, assim como pode ser observado na Figura 1.1 (c). A identificação de microaneurismas é uma tarefa fundamental na análise de imagens de fundo de olho. No entanto, a identificação manual de microaneurismas pode ser uma tarefa árdua e sujeita a erros, devido à sua natureza minuciosa e à possível presença de artefatos e ruídos nas imagens (JIMÉNEZ et al., 2011).

Por isso, é importante desenvolver sistemas automáticos de detecção e classificação da RD (*Retinopatia Diabética*) de forma a identificar a doença em seus diferentes estágios e permitir um tratamento precoce e eficaz, uma vez que a detecção precoce e precisa de microaneurismas é essencial para um diagnóstico adequado e um plano de tratamento efetivo.

Alguns métodos propostos na literatura científica como o aprendizado de máquina e

reconhecimento de padrões lidam essencialmente com o problema de encontrar automaticamente uma decisão, por exemplo, segmentar as estruturas como o disco óptico, mácula e vasos sanguíneos (CARDOSO, 2019)(CARDOSO et al., 2019) e trabalhos que fizeram melhorias nas imagens para serem segmentadas, como (SANTOS et al., 2020).

Este trabalho utiliza o aprendizado profundo, a diferença é que, o aprendizado de máquina refere-se a um conjunto de algoritmos e técnicas que permitem que os computadores aprendam a partir dos dados. Ele se concentra em projetar algoritmos que podem aprender com os dados, identificar padrões e tomar decisões ou fazer previsões com base nessas informações. O aprendizado de máquina geralmente requer que os dados sejam pré-processados e que características relevantes sejam extraídas antes que o modelo de aprendizado seja aplicado. Os algoritmos de aprendizado de máquina podem ser divididos em três categorias principais: aprendizado supervisionado, aprendizado não supervisionado e aprendizado por reforço.

Por outro lado, o aprendizado profundo é uma subárea específica do aprendizado de máquina que se concentra no treinamento de redes neurais profundas, também conhecidas como redes neurais artificiais profundas. As redes neurais profundas são modelos computacionais compostos por várias camadas, que podem ser treinadas para aprender representações hierárquicas dos dados. Essas redes são inspiradas no funcionamento do cérebro humano e são capazes de aprender automaticamente características relevantes dos dados sem a necessidade de extração manual de características. O aprendizado profundo tem se destacado em tarefas como reconhecimento de imagens, processamento de linguagem natural, tradução automática e muitas outras.

Na literatura, existem outros estudos sobre segmentação de microaneurismas baseados em aprendizado de máquina. Dentre eles, destacam-se o modelo bayesiano hierárquico de três níveis CNN-LDA (CNN - Convolutional Neural Network) (LDA - Latent Dirichlet Allocation), que realiza inferência de tópicos.(MCLACHLAN, 2004) propõe o uso do aprendizado de transferência de conhecimento LDA para DNN (Deep Neural Network), treinando um modelo DNN capaz de simular o comportamento da inferência LDA com menor consumo computacional.

(COVER; HART, 1967) apresentam a CNN-KNN (KNN - K-Nearest Neighbors), que utiliza a estrutura não paramétrica clássica baseada no KNN. Nesse método, a rede neural convolucional paramétrica (M CNN) é utilizada para prever a confiança e os deslocamentos correspondentes da região mais compatível em uma imagem KNN, para uma determinada região semântica na imagem de teste.

Outra abordagem é a CNN-SVM (SVM - Support Vector Machine), introduzida por (ALALSHEKMUBARAK; SMITH, 2013); (AGARAP, 2017); (TANG, 2013), que combina a máquina vetorial de suporte linear (SVM) em uma arquitetura de rede neural artificial. A arquitetura proposta neste trabalho é uma modificação da U-NET básica, composta por seis camadas e a ativação "relu". Detalhes adicionais serão descritos no capítulo 3.

A estrutura deste trabalho é a seguinte: o capítulo 2 aborda os fundamentos básicos para a compreensão do estudo, o capítulo 3 apresenta os materiais e métodos utilizados na pesquisa, o capítulo 4 descreve os resultados obtidos e a discussão sobre eles, e, por fim, a conclusão é apresentada.

CAPÍTULO 2

FUNDAMENTOS TEÓRICOS

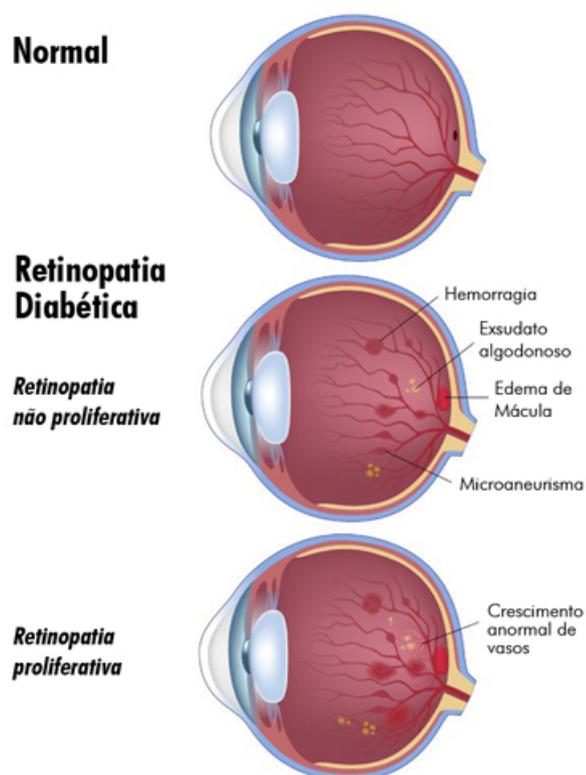
Este capítulo apresenta alguns fundamentos relevantes para a compreensão dos assuntos tratados ao longo deste trabalho. Assim, serão abordados aspectos em oftalmologia tais como a retinopatia diabética, microaneurismas, redes neurais artificiais, redes neurais convolucionais e arquiteturas de redes neurais convolucionais para segmentação e reconhecimento de imagens

2.1 Aspectos em oftalmologia

No que concerne a oftalmologia, é importante discutir alguns aspectos relevantes relacionados à identificação de microaneurismas em imagens de fundo de olho. Embora este texto não tenha a intenção de esgotar o assunto, suas informações têm como objetivo facilitar o entendimento inicial.

2.1.1 Retinopatia diabética

No Brasil, de acordo com o Ministério da Saúde (GOVDIABETES, 2023), 13 milhões de pessoas vivem com diabetes, de 24% a 39% dessas pessoas (PCDT, 2021) é afetada com uma condição médica chamada RD, que é uma doença ocular grave que pode levar à cegueira. Como é possível observar na Figura 2.1 existem dois tipos de RD: a Retinopatia Diabética Não Proliferativa (RDNP) e a Retinopatia Diabética Proliferativa (RDP). A forma mais branda é a RDNP, que é caracterizada por exsudatos e pode causar visão turva. A RDP, por sua vez, é a forma avançada da doença e é caracterizada pelo crescimento de novos vasos sanguíneos na retina, que podem vaziar ou sangrar facilmente, causando sangramentos na retina (MSDMANUALS, 2022).

Figura 2.1: Retinopatia diabética proliferativa e não proliferativa.

Fonte: <<https://www.centrooftalmologicomg.com.br/blog/retinopatia-diabetica/>>

Nos estágios iniciais da RD, os sinais mais comuns são os exsudatos e os microaneurismas. Os exsudatos são acúmulos de fluidos e proteínas na retina, resultantes do dano causado pelos altos níveis de glicose no sangue. Já os microaneurismas se apresentam como pequenos pontos vermelhos arredondados no fundo do olho. Conforme a doença progride, a retina pode sofrer danos mais graves, incluindo o bloqueio completo dos vasos sanguíneos e a formação de lesões nos vasos sanguíneos, como hemorragias (MSDMANUALS, 2022).

2.1.2 Avaliação da retinografia

A avaliação manual de imagens de fundo de olho para detecção de Retinopatia Diabética (RD) é realizada por oftalmologistas treinados e experientes, eles usam um aparelho chamado retinógrafo, Figura 2.2 para obter a imagem de fundo de olho e em seguida realizam uma análise minuciosa.

O retinógrafo é um aparelho que permite a captura de imagens digitais da retina, com alta resolução e nitidez. Durante o exame, o paciente é posicionado diante do aparelho, que emite flashes de luz para registrar imagens do fundo do olho. Essas imagens são então processadas e armazenadas em um computador, permitindo a avaliação por parte do oftalmolo-

Figura 2.2: Retinógrafo Modelo TRC-NW8 Non-Mydriatic Retinal Camera.

Fonte: <<https://topconhealthcare.com/products/trc-nw8f/>>

gista(SAUDEBEMESTAR, 2023)

O exame manual feito pelo especialista envolve a análise de vários aspectos da imagem da retina, como a presença de lesões, como microaneurismas, exsudatos e hemorragias, bem como a presença de anomalias nos vasos sanguíneos da retina. O oftalmologista também avalia o tamanho e a localização das lesões, e pode usar escalas de classificação para determinar o estágio da RD, a Figura 2.3 mostra essa classificação: Sem retinopatia aparente, RDNP leve, RDNP moderada, RDNP grave, RDNP muito grave e RDP.

Figura 2.3: Níveis de gravidade da retinopatia e alterações observáveis ao exame de fundo de olho.

Nível de gravidade da retinopatia	Alterações observáveis à fundoscopia dilatada
Sem retinopatia aparente	Sem alterações.
RDNP leve	Apenas microaneurismas.
RDNP moderada	Presença de microaneurismas, mas que ainda não caracterize RDNP grave. Hemorragias intra-retinianas nos 4 quadrantes ou
RDNP grave com um dos critérios ao lado RDNP muito grave com dois critérios	Alterações venosas em conta em 2 ou mais quadrantes ou IRMA moderada em 1 ou mais quadrantes.
RDP	Neovascularização de disco ou de retina ou Hemorragia vítrea ou pré-retiniana.

RDNP = Retinopatia Diabética não Proliferativa; IRMA = Alterações Microvasculares Intra-Retinianas; RDP = Retinopatia Diabética Proliferativa.

Fonte: <https://www.gov.br/saude/pt-br/assuntos/protocolos-clinicos-e-diretrizes-terapeuticas-pcdt/arquivos/2021/portal-portaria-conjunta_pcdt_retinopatia-diabetica_.pdf>

Essa avaliação pode ser um processo demorado e trabalhoso, pois o oftalmologista precisa

examinar cuidadosamente a imagem da retina para identificar todos os sinais de RD. No entanto, para alguns pacientes, especialmente nos estágios iniciais, é difícil detectar a doença de forma precisa.

Para auxiliar os profissionais de saúde na detecção dessas estruturas, diversas técnicas de processamento de imagem e análise computacional têm sido desenvolvidas. Essas técnicas empregam algoritmos sofisticados para automatizar o processo de identificação de microaneurismas, aumentando a eficiência e a precisão do diagnóstico.

2.1.3 Microaneurisma

Os microaneurismas são pequenas dilatações ou sacos nas paredes dos capilares sanguíneos localizados na retina, a camada sensível à luz no fundo do olho. Eles são uma característica comum e importante da retinopatia diabética, uma complicação ocular associada ao diabetes crônico.

Essas estruturas se formam devido a danos progressivos nos vasos sanguíneos da retina causados pelo alto nível de glicose no sangue ao longo do tempo. O elevado nível de glicose danifica as células endoteliais, que revestem os capilares sanguíneos, tornando-os frágeis e propensos a se dilatar. Com o tempo, essa fragilidade leva à formação dos microaneurismas.

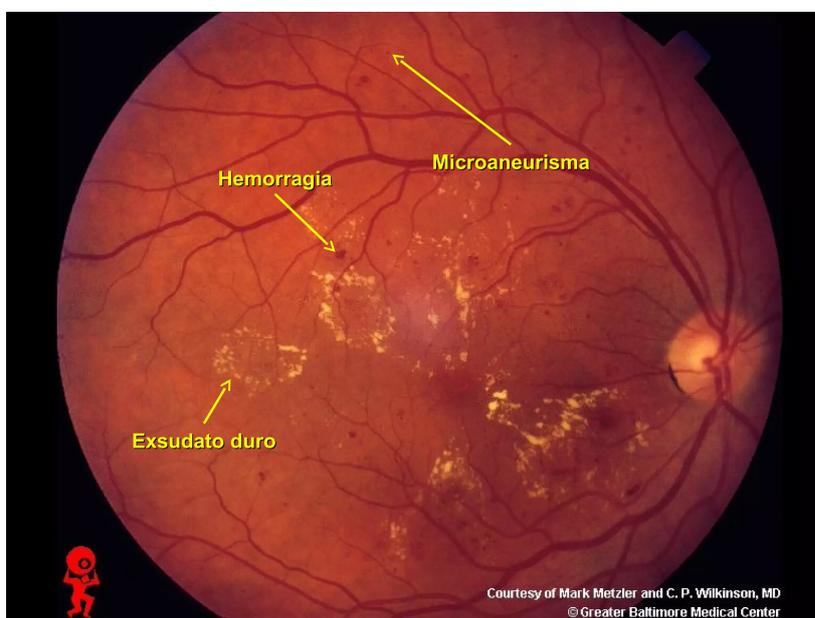
Os microaneurismas podem ser observados durante um exame oftalmológico como pequenos pontos vermelhos dispersos na retina, conforme a Figura 2.4, nela observamos hemorragia, exsudato duro e um microaneurisma, essas são todas características da retinopatia diabética. Geralmente, eles têm um tamanho muito pequeno, variando de 10 a 100 micrômetros de diâmetro. Essa dimensão diminuta significa que eles não afetam diretamente a visão, mas sua presença é um sinal de alerta para a presença de danos nos vasos sanguíneos da retina.

Essas estruturas são consideradas biomarcadores importantes da retinopatia diabética, pois sua ocorrência está associada a outras alterações microvasculares, como hemorragias, exsudatos (depósitos de fluido ou sangue) e edema macular (acúmulo de líquido na região central da retina). Os microaneurismas indicam a presença de danos nos vasos sanguíneos da retina, que são fundamentais para a saúde ocular.

É crucial realizar um diagnóstico precoce e monitorar regularmente os microaneurismas, pois sua progressão pode levar a complicações mais graves, incluindo o descolamento da retina e a perda permanente da visão. O controle adequado dos níveis de glicose no sangue, juntamente com um bom controle da pressão arterial e dos níveis de colesterol, desempenham

um papel fundamental na prevenção e no manejo da retinopatia diabética, reduzindo o risco de desenvolvimento de microaneurismas e outras lesões oculares relacionadas à doença. Fonte: (MSDMANUALS, 2022)

Figura 2.4: Microaneurisma, hemorragia e exsudato duro.



Fonte: <<https://pt.slideshare.net/PietrodeAzevedo/retinopatia-diabtica>> Slide 6.

2.2 Redes neurais artificiais

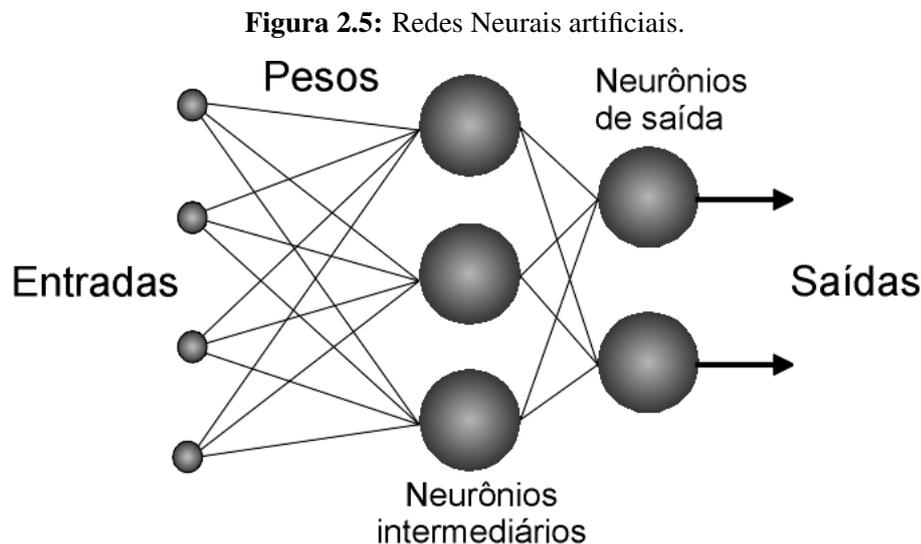
Redes neurais artificiais são modelos de computação inspirados no cérebro humano que ajudam a aprender e reconhecer padrões. Elas são compostas por um grupo de “neurônio” artificiais conectados, que processam informações e fazem cálculos para produzir um resultado. Conforme mencionado por (PEDRYCZ; GOMIDE, 2007) citados por (DAMBROS, 2020)), as redes neurais artificiais foram desenvolvidas com o propósito de efetuar a identificação e aprendizado de padrões.

A Figura 2.5 ilustra uma rede neural artificial simples, nela pode-se observar uma "camada de entrada", uma "camada oculta" e uma "camada de saída". O número de camadas varia dependendo do que a rede precisa fazer, podendo ter várias camadas em uma única rede neural. Cada "neurônio" artificial na rede neural tem entradas, pesos e uma "ativação", conforme demonstrado na Figura 2.6. As entradas são os dados que o "neurônio" recebe de outras camadas ou da entrada original, e os pesos são valores que determinam a importância de cada entrada para o resultado do "neurônio". A "ativação" do "neurônio" é calculada somando as entradas multiplicadas pelos pesos, passando por uma função que dá um resultado final.

Durante o treinamento da rede neural, os pesos são ajustados repetidamente para reduzir a diferença entre o resultado da rede e o resultado desejado para um conjunto de dados de treinamento. Isso é feito usando algoritmos especiais que calculam quão "errada" a rede está e ajustam os pesos na direção certa.

A escolha da estrutura e do número de camadas em uma rede neural é importante para o seu desempenho em uma determinada tarefa de aprendizado de máquina. Camadas adicionais podem ajudar a rede a identificar características mais complexas e a ser mais versátil. No entanto, muitas camadas podem levar a um problema chamado "overfitting", onde a rede memoriza os dados de treinamento em vez de aprender padrões úteis (DEMAR; ZUPAN, 2021).

Conforme (KRENKER et al., 2011), com um treinamento adequado, uma rede neural pode ser usada para realizar várias tarefas, como classificar imagens, reconhecer fala, traduzir automaticamente e prever padrões em séries de dados. As redes neurais são usadas em muitas aplicações do mundo real, como reconhecimento de voz em assistentes virtuais, classificação de imagens em aplicativos de fotos e detecção de fraudes em transações financeiras.

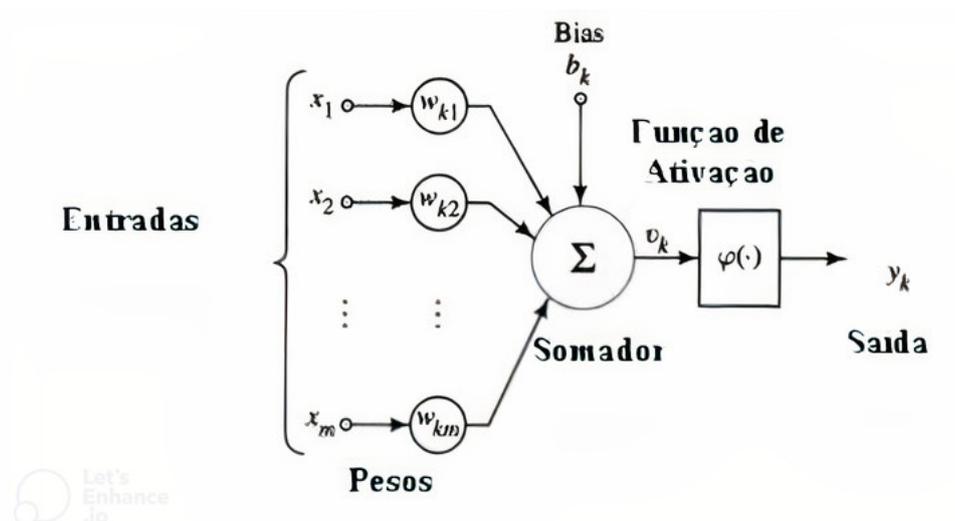


Fonte: <<http://www.cerebromente.org.br/n05/tecnologia/rna.htm>>

2.2.1 Redes neurais convolucionais

Com o avanço da tecnologia, surgiram as redes neurais profundas, que podem realizar tarefas complexas, como reconhecimento de imagens (RODRIGUES, 2018). Dentro dessas novas redes, deve-se destacar as Redes Neurais Convolucionais (CNNs), que também apresentam uma estrutura organizada em camadas, e incluem camadas de convolução nas quais são realizadas

Figura 2.6: Neurônio artificial.



Fonte: <<https://ifgjatai.webcindario.com/rna.html>>

operações de convolução de forma que essas camadas desempenham um papel fundamental na identificação de características das imagens. Pode-se pensar nelas como filtros que percorrem a imagem, cada um procurando por uma característica específica (DESHPANDE, 2016).

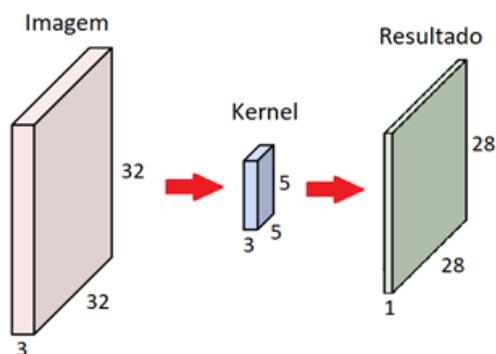
Como exemplo de tais características podemos citar bordas retas, curvaturas distintas, cores simples e outros elementos abstratos que são comuns a todas as imagens. Assim, as camadas de convolução ajudam na extração dessas características, fazendo com que a rede neural aprenda a reconhecer padrões visuais pertinentes a tarefa.

Parker (2010) (FERREIRA, 2017) define a convolução da seguinte maneira:

[...] uma operação matemática entre duas funções f e g , produzindo uma terceira função, que pode ser interpretada como uma função modificada de f . No processamento de imagens, onde a imagem é definida como uma função bidimensional, a convolução é útil para detecção de bordas, suavização de imagem, extração de atributos, entre outras aplicações. (FERREIRA, 2017, p 27).

A convolução desempenha um papel essencial nas redes neurais convolucionais, sendo empregada para extrair características de uma imagem de entrada. A Figura 2.7 demonstra que esse processo envolve a aplicação de um filtro(kernel) que percorre a imagem, realizando a multiplicação de cada trecho da imagem com o filtro. Esse processo resulta em uma nova imagem alterada que ressalta as características mais relevantes.

As redes neurais convolucionais são especialmente projetadas para processar dados

Figura 2.7: Convolução.

Fonte: <<https://viceri.com.br/insights/entendendo-de-vez-a-convolucao-base-para-processamento-de-imagens/>>

estruturados, como imagens, por meio da aplicação de operações convolucionais. Essas operações permitem que as CNNs capturem características locais dos dados e preservem a relação espacial entre os pixels. Por outro lado, as redes neurais não convolucionais são mais adequadas para dados não estruturados, como sequências de texto ou séries temporais. Elas geralmente consistem em camadas totalmente conectadas, onde cada neurônio está conectado a todos os neurônios da camada anterior, permitindo a aprendizagem de padrões e relações complexas em dados sequenciais. Portanto, as redes neurais convolucionais são especialmente eficazes em tarefas de visão computacional, enquanto as redes neurais não convolucionais são mais adequadas para processar dados sequenciais ou não estruturados.

2.2.2 Arquiteturas de redes neurais convolucionais para segmentação e reconhecimento de imagens

A arquitetura de uma rede neural refere-se à organização das camadas de uma rede neural e à maneira como os neurônios estão conectados entre si. Existem diferentes tipos de arquiteturas de redes neurais, mas uma das mais comuns é a chamada "rede neural feedforward" ou "rede neural de alimentação direta".

A arquitetura de uma rede neural pode variar em termos de número de camadas ocultas, número de neurônios em cada camada, tipo de função de ativação usada nos neurônios e outras características específicas. Essas variações permitem que as redes neurais sejam adaptadas para diferentes tipos de problemas, como classificação de imagens, reconhecimento de voz, previsão de séries temporais, entre outros.

Métodos baseados em aprendizado de máquina têm demonstrado grande capacidade e interesse nos últimos anos, e um advento que reforça esse fato é o surgimento das redes neurais

convolucionais (Convolutional Neural network - CNN). Muitas arquiteturas baseadas nas CNNs foram propostas, e atualmente são o estado da arte em tarefas como segmentação e classificação de imagens, além disso, os datasets como Mnist, Imagenet, GoogleNet, CIFAR10, AlexNet entre outros impulsionaram de modo considerável as pesquisas científicas em aprendizagem de máquinas (PAVLOVIČOVÁ; AL., 2018). No contexto de imagens biomédicas, uma arquitetura muito eficiente foi proposta por (RONNEBERGER et al., 2015) designada por U-net, esta arquitetura é baseada em contração e expansão no processo de convolução, o que lhe permite dar um melhor tratamento a esse tipo de imagem, que costuma apresentar grandes dimensões, baixo contraste e ruído.

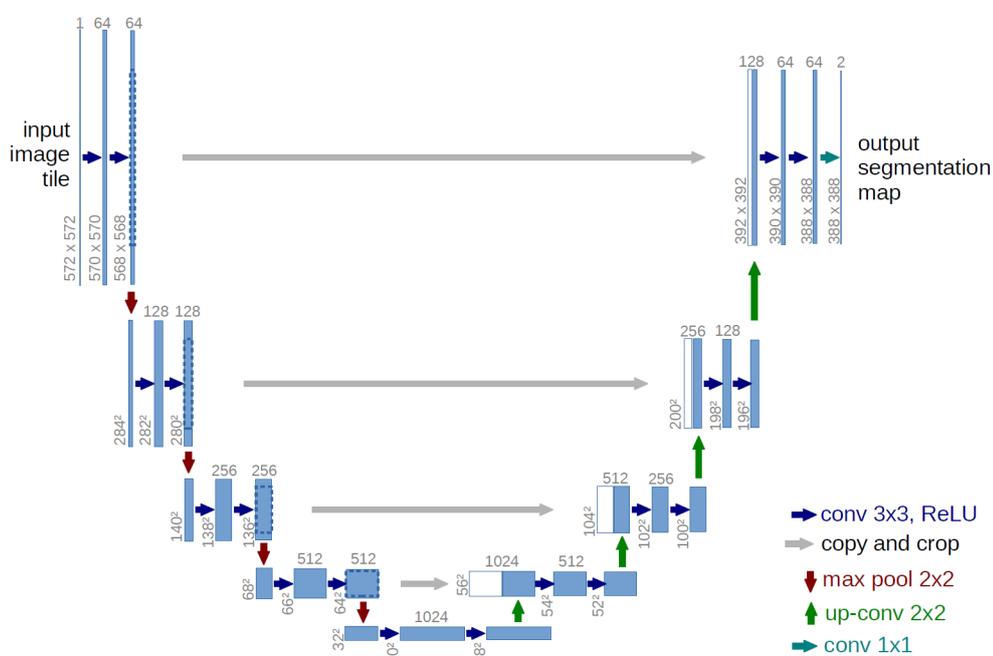
A U-Net é uma arquitetura de rede neural convolucional totalmente convolucional, também conhecida como FCN (Fully Convolutional Network) em inglês. Foi criada e implementada por Olaf Ronneberger et al. em 2015, na Universidade de Freiburg, na Alemanha (RONNEBERGER et al., 2015). Diferentemente das CNNs tradicionais, a U-Net substitui as camadas totalmente conectadas por camadas convolucionais que operam localmente, incluindo camadas de convolução, agrupamento e upsampling (SHELHAMER et al., 2017). Essa arquitetura foi desenvolvida especificamente para segmentação de imagens biomédicas e se tornou um dos métodos mais populares para qualquer tarefa de segmentação semântica atualmente.

A U-Net é uma rede neural totalmente convolucional projetada para aprender com menos amostras de treinamento. Ela consiste em um caminho de contração e um caminho de expansão. O caminho de contração segue a arquitetura típica de uma rede convolucional, com a aplicação repetida de duas convoluções 3x3 (sem preenchimento), seguidas de uma unidade linear retificada (ReLU) e uma operação de max pooling 2x2 com passo 2 para redução da amostragem. Em cada passo de redução de amostragem, o número de canais de características (recursos computacionais que representam as informações extraídas da imagem na codificação e decodificação.) é dobrado. Cada passo no caminho de expansão consiste em um aumento da amostra do mapa de características (subconjuntos de recursos de uma determinada camada da rede que são selecionados e usados para melhorar a precisão da segmentação.), seguido de uma convolução 2x2 ("convolução ascendente") que reduz pela metade o número de canais de características, uma concatenação com o mapa de características correspondente cortado do caminho de contração e duas convoluções 3x3, cada uma seguida por uma unidade ReLU. O corte é necessário devido à perda de pixels de borda em cada convolução. Na camada final, uma convolução 1x1 é usada para mapear cada vetor de características de 64 componentes para o

número desejado de classes. No total, a rede tem 23 camadas convolucionais.

A U-Net é simples, eficiente, fácil de entender e construir, e pode ser treinada com conjuntos de dados pequenos. Ela se baseia no forte uso de aumento de dados para usar amostras anotadas (Imagens de treinamento) disponíveis de forma mais eficiente. Como pode se observar na Figura 2.8, a rede é composta por quatro blocos de codificador e quatro blocos de decodificador que são conectados por uma ponte. O codificador reduz o tamanho dos mapas de características pela metade enquanto dobra o número de canais de características após cada bloco convolucional. O decodificador faz o oposto, dobrando o tamanho dos mapas de características enquanto reduz pela metade o número de canais de características após cada bloco convolucional.

Figura 2.8: Arquitetura da U-Net.



Fonte: (RONNEBERGER et al., 2015)

CAPÍTULO 3

MATERIAIS E MÉTODOS

Este capítulo detalha os materiais e métodos utilizados no desenvolvimento do trabalho, inicialmente são apresentadas as métricas utilizadas, em seguida as características no projeto da arquitetura, as ferramentas e a arquitetura proposta.

3.1 Métricas

A avaliação de métodos de segmentação é altamente subjetiva pois o resultado desse processo é uma imagem em que os objetos são separados por linhas ou cores, e a validação do método é feita comparando essa imagem resultante com uma imagem segmentada por um especialista, chamada de *ground truth* (verdade absoluta), que é obtida por meio de segmentação manual buscando contornar problemas como sombras, obstruções e outros (CARDOSO, 2019).

Ao comparar uma imagem segmentada por um especialista com uma imagem segmentada por um método, várias métricas podem ser definidas. Um "verdadeiro positivo" significa que o objeto alvo foi classificado pelo método de acordo com o que foi estabelecido pelo especialista, sendo que o objeto alvo pode ser qualquer elemento na imagem. No caso da análise do fundo de olho, os possíveis alvos são o disco óptico, os vasos sanguíneos, exsudatos, microaneurismas, e assim por diante. Conforme CARDOSO (2019), Purusharathi e Kurhade (2014) e Mittapalli e Kande (2016) definem os possíveis resultados dessa comparação:

Tabela 3.1: Classificação de um pixel na segmentação, em comparação com a segmentação feita por um especialista.

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

Considerando dois conjuntos distintos resultantes da segmentação manual: **A** - *pixels* pertencentes ao objeto alvo, e **Q** - *pixels* que não pertencem ao objeto alvo, e a FOV (*Field Of View*), que é a região de interesse, são estabelecidas as seguintes relações:

- $A \cup Q$: conjunto de todos os *pixels* no FOV;
- $TP \cup TN$: conjuntos dos *pixels* corretamente classificados;
- $FP = Q - TN$: o conjunto dos falso positivos é o complemento de **TN** em relação a **Q**;
- $FN = A - TP$: o conjunto dos falso negativos é o complemento de **TP** em relação a **A**.

Considerando a cardinalidade de um conjunto representada pela notação $\|\cdot\|$, são definidas as seguintes medidas de avaliação da qualidade dos métodos, (PURUSHARATHI; KURHADE, 2014; WANG et al., 2015; MITTAPALLI; KANDE, 2016):

- TPR (*True Positives Ratio*) - taxa de verdadeiros positivos, ou seja, fração de pixels corretamente classificados como pertencentes ao alvo:

$$TPR = \frac{\|TP\|}{\|A\|} \quad (3.1)$$

- FPR (*False Positives Ratio*) - taxa de falsos positivos, ou seja, a fração de pixels erroneamente classificados como pertencentes ao alvo:

$$FPR = \frac{\|FP\|}{\|Q\|} \quad (3.2)$$

- *Acc (Accuracy)* - precisão é a taxa de acertos em relação ao total de *pixels* na área sendo investigada - FOV. As medidas de verdadeiro positivos e verdadeiro negativos (TP e TN) refletem o total de acertos na segmentação:

$$Acc = \frac{\|TP + TN\|}{\|FOV\|} \quad (3.3)$$

A precisão (*Acc*) é amplamente utilizada na literatura como uma das principais métricas. Ela combina informações sobre a capacidade do método em distinguir, gerando um único valor que permite a comparação rápida, direta e eficiente entre uma variedade de algoritmos.

3.2 Características no projeto de arquitetura de rede

Projetar uma arquitetura de rede neural envolve a definição cuidadosa de suas características, como o número e tipo de camadas, funções de ativação, estratégias de inicialização de pesos e outras decisões específicas. Esse processo requer experimentação e análise para encontrar a configuração mais adequada para o problema em questão.

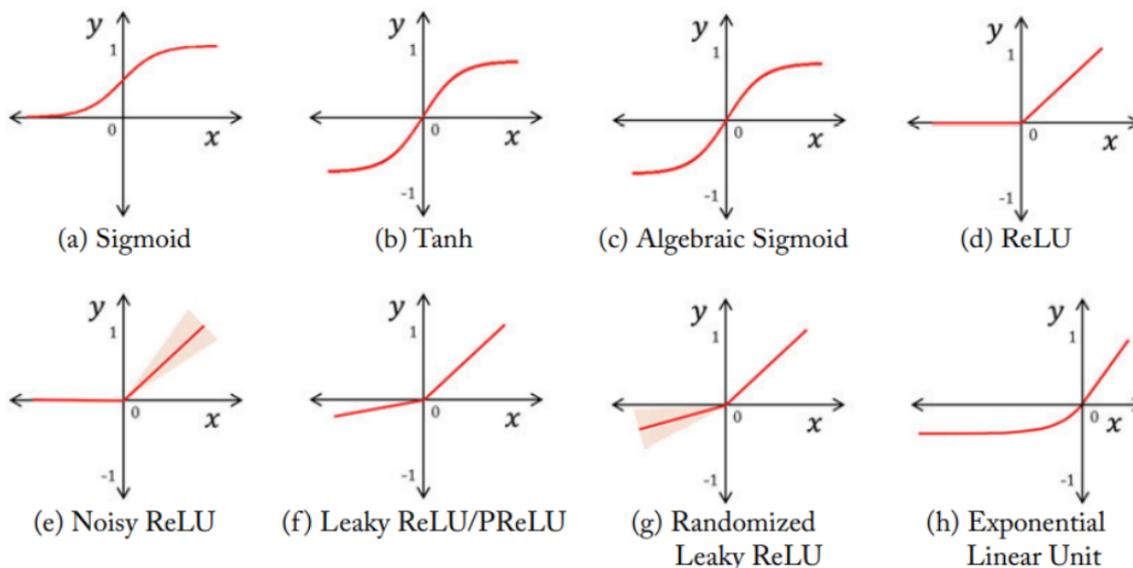
Conforme foi dito anteriormente, é possível compreender a convolução como a utilização de filtros (*kernels*) nos dados de entrada da rede, visando extrair características como linhas, arestas e cores de imagens. Essas informações são fundamentais para o treinamento da rede, bem como para o subsequente reconhecimento e classificação de padrões (RODRIGUES, 2018).

Quanto as funções de ativação, a Figura 3.1, exibe algumas delas. A ReLU é uma das funções de ativação comum em redes neurais profundas, que é usada para introduzir não-linearidade no modelo. A função retorna a entrada se for positiva e 0 se for negativa, matematicamente representada como $f(x) = \max(0, x)$. Essa função é importante para ajudar o modelo a aprender características mais complexas.

Uma outra operação importante na arquitetura de uma CNN é o *pooling*, que é uma operação utilizada em redes neurais convolucionais para reduzir a dimensionalidade do mapa de características, resumindo a informação em regiões e retendo as informações importantes. Segundo (ABREU, 2021) é possível utilizar diferentes abordagens para a aplicação do *pooling*, como o *max pooling* e o *pooling* de média, Figura 3.2. O *max pooling*, Figura 3.3, consiste em dividir a matriz em regiões e selecionar os valores mais altos dessas regiões como representantes. Por sua vez, o *pooling* de média realiza a média dos valores presentes nessas regiões.

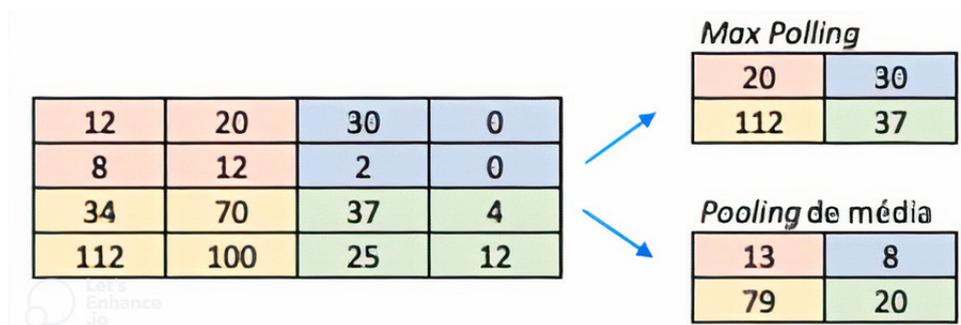
MaxPooling2D, Figura 3.4, é uma variante da operação de *pooling* usada para reduzir

Figura 3.1: Funções não lineares que são frequentemente utilizadas em arquiteturas de aprendizado profundo.



Fonte: (KHAN; AL., 2020)

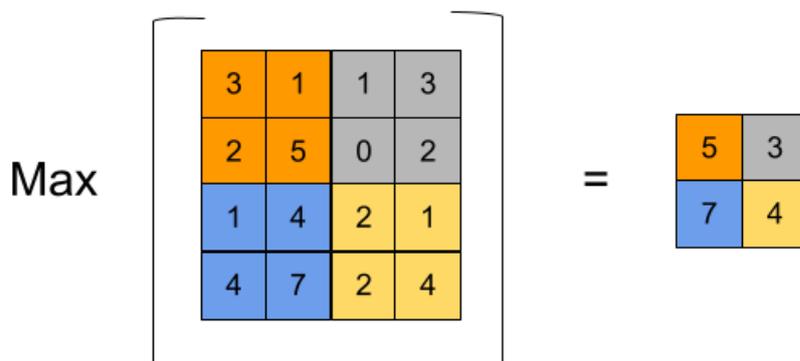
Figura 3.2: Pooling.



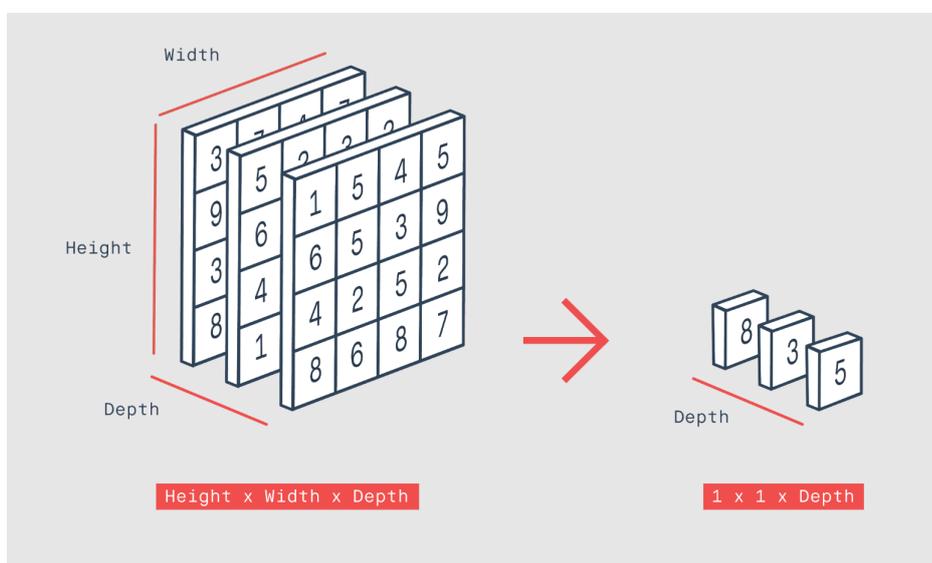
Fonte: (ABREU, 2021)

as dimensões espaciais de um mapa de características 2D. Ele funciona movendo uma janela sobre o mapa de características e tomando o valor máximo dentro dessa janela como saída para a localização correspondente. Essa técnica é útil para reduzir a complexidade computacional do modelo e prevenir o *overfitting*, além de fornecer invariância à translação (Fonte: *Keras*(MAXPOOLING2D..., 2023)). Em resumo, essas técnicas são fundamentais para a construção de redes neurais convolucionais eficazes em tarefas de processamento de imagem e visão computacional.

Dropout, exemplificado na Figura 3.5, é uma técnica criada por (SRIVASTAVA; AL., 2014) comumente usada em redes neurais para prevenir *overfitting*, que ocorre quando o modelo se ajusta muito bem aos dados de treinamento, mas não generaliza bem para novos dados. Essa

Figura 3.3: Max Pooling.

Fonte: <<https://androidkt.com/explain-pooling-layers-max-pooling-average-pooling-global-average-pooling-and-global-max-pooling/>>

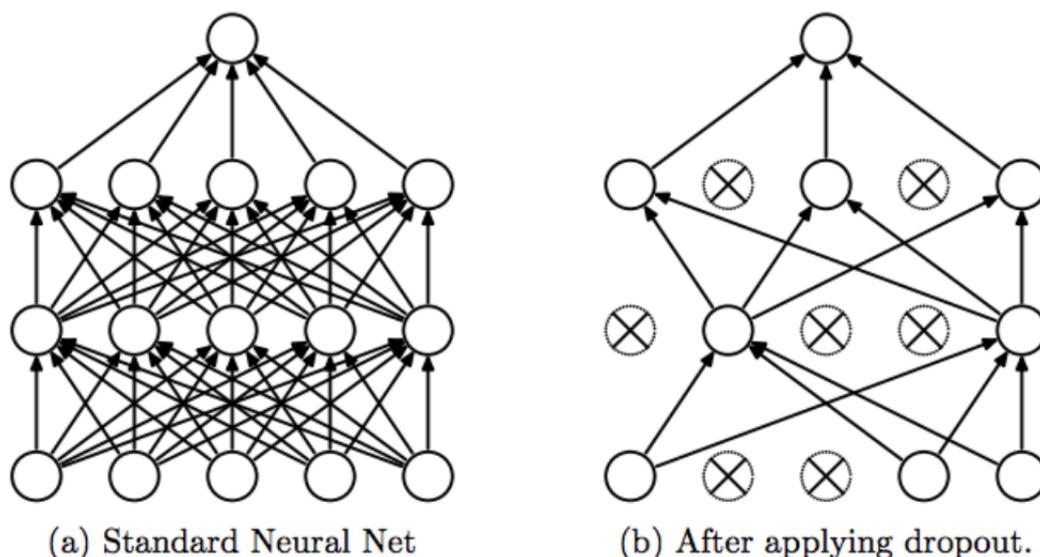
Figura 3.4: Max Pooling.

Fonte: <https://www.ahmedbesbes.com/static/c281d168ae48705ee23c8434efc945/124f1/global_average_pooling.png>

técnica consiste em aleatoriamente desligar um conjunto de neurônios durante o treinamento, fazendo com que os neurônios remanescentes sejam forçados a adquirir características mais robustas do problema em questão (VERGARA, 2018).

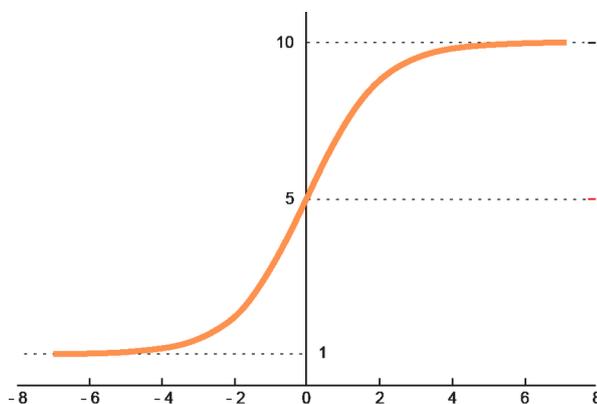
UpSampling2D é uma operação usada em redes neurais convolucionais para aumentar a resolução espacial de um mapa de características. Ele funciona replicando os valores da entrada e interpolando entre eles para produzir uma saída com dimensões maiores. Essa técnica é útil em tarefas como segmentação de imagens, em que é necessário produzir uma máscara com a mesma resolução espacial da imagem de entrada.

Softmax, Figura 3.6, é uma função de ativação comumente usada em redes neurais para transformar um conjunto de valores em uma distribuição de probabilidade. É amplamente

Figura 3.5: *Dropout*.

Fonte: (SRIVASTAVA; AL., 2014)

utilizada em tarefas de classificação, em que é necessário prever a probabilidade de cada classe. A função *Softmax* normaliza as saídas para que a soma de todas as saídas seja igual a 1, o que torna mais fácil interpretar as previsões do modelo.

Figura 3.6: *Softmax*.

Fonte: <[https:](https://www.researchgate.net/figure/Graphic-representation-of-the-softmax-activation-function_fig5_348703101)

[//www.researchgate.net/figure/Graphic-representation-of-the-softmax-activation-function_fig5_348703101](https://www.researchgate.net/figure/Graphic-representation-of-the-softmax-activation-function_fig5_348703101)>

O otimizador *sgd* (gradiente descendente estocástico) é um dos otimizadores mais simples e amplamente utilizados em redes neurais. Ele ajusta os pesos da rede para minimizar a função de perda (loss) usando o gradiente da função calculada em pequenos lotes (*batches*) de dados de treinamento. O *sgd* tem uma série de hiperparâmetros, como taxa de aprendizagem (learning rate), momento (*momentum*) e decaimento do *learning rate* (*learning rate decay*), que podem ser ajustados para melhorar o desempenho do modelo.

A loss *categorical_crossentropy* é uma função de perda comumente usada em tarefas de

classificação multiclasse. Ela calcula a diferença entre as probabilidades preditas pelo modelo e as probabilidades reais das classes, e retorna um valor que indica a qualidade da predição. A *loss categorical_crossentropy* é usada em conjunto com a função *Softmax* para calcular a probabilidade de cada classe.

A métrica *accuracy* é uma medida comum de desempenho usada em tarefas de classificação. Ela calcula a proporção de amostras de teste que foram classificadas corretamente pelo modelo. A métrica *accuracy* é simples de calcular e interpretar, mas pode ser enganosa em casos em que as classes estão desbalanceadas ou em tarefas em que a precisão não é o único objetivo.

Evitar o *overfitting* (ou sobreajuste) é um dos principais desafios enfrentados durante o treinamento de uma rede neural. O *overfitting* acontece quando o modelo treinado não consegue generalizar efetivamente o conhecimento adquirido. Em outras palavras, quando é submetido a dados externos não utilizados no treinamento, não consegue apresentar o desempenho esperado, sugerindo que se adaptou excessivamente apenas aos dados utilizados durante o treinamento (KHAN; AL., 2020)

3.3 Ferramentas

O algoritmo foi desenvolvido em um notebook Acer Nitro 5 AN515, com 16GB de memória RAM, NVidia GTX 1050 e processador intel i5-8300H.

Algumas ferramentas foram utilizadas no desenvolvimento deste trabalho, destacando-se o *TensorFlow*, o *Keras*, o *Spyder* (IDE - Integrated Development Environment) e a linguagem de programação Python. Essas ferramentas foram escolhidas levando em consideração diversos fatores que contribuem para o desenvolvimento eficiente e eficaz do trabalho.

O *TensorFlow* é uma das bibliotecas mais populares para a implementação de redes neurais e aprendizado de máquina. Ele oferece uma ampla gama de recursos e funcionalidades avançadas que facilitam o projeto e treinamento de arquiteturas de redes neurais. Sua flexibilidade, escalabilidade e suporte para processamento em *GPU* são vantagens importantes para lidar com grandes conjuntos de dados e modelos complexos. Além disso, o *TensorFlow* possui uma comunidade ativa e recursos de documentação abundantes, o que facilita o aprendizado e a resolução de problemas.

O *Keras* é uma biblioteca de alto nível construída sobre o *TensorFlow* que oferece uma interface simples e intuitiva para projetar e treinar redes neurais. Ele simplifica tarefas complexas, como a definição das camadas da rede neural, a configuração dos hiperparâmetros e a compilação

do modelo. O *Keras* também fornece uma variedade de funções de ativação, otimizadores e métricas de avaliação, tornando o processo de desenvolvimento mais ágil e produtivo.

A escolha do *Spyder* como ambiente de desenvolvimento integrado (IDE) está relacionada à sua facilidade de uso e ao suporte completo para a linguagem de programação *Python*. O *Spyder* oferece uma interface amigável e recursos específicos para desenvolvimento científico, como edição interativa de código, depuração, visualização de dados e integração com bibliotecas populares. Ele proporciona um ambiente eficiente para escrever, testar e depurar o código de implementação da rede neural.

A linguagem de programação *Python* foi escolhida por sua popularidade e versatilidade no campo do aprendizado de máquina e análise de dados. *Python* possui uma sintaxe clara e legível, vasta disponibilidade de bibliotecas especializadas em ciência de dados e aprendizado de máquina, além de uma comunidade ativa que promove o compartilhamento de conhecimento e a colaboração.

3.3.1 Tensor Flow

TensorFlow é uma biblioteca de código aberto desenvolvida pelo *Google* para realização de cálculos numéricos e treinamento de modelos de aprendizado de máquina (campo da inteligência artificial que utiliza algoritmos para analisar dados e aprender padrões, permitindo que um modelo melhore sua performance em tarefas específicas com base em exemplos passados.), incluindo redes neurais. A sua arquitetura flexível permite construir e treinar modelos de aprendizado de máquina com rapidez e facilidade, além de fornecer recursos de diferenciação automática para simplificar o processo de definição de novas redes.

Uma das principais vantagens do *TensorFlow* é a sua capacidade de rodar em múltiplas CPUs (Central Processing Unit) e GPUs (Graphics Processing Unit), o que o torna adequado para tarefas de aprendizado de máquina em larga escala. O *TensorFlow* também oferece suporte a diversas linguagens de programação, incluindo *Python* e C++, bem como a diversos *back-ends*, tanto em CPU quanto em GPU, em *desktops*, servidores ou plataformas móveis.

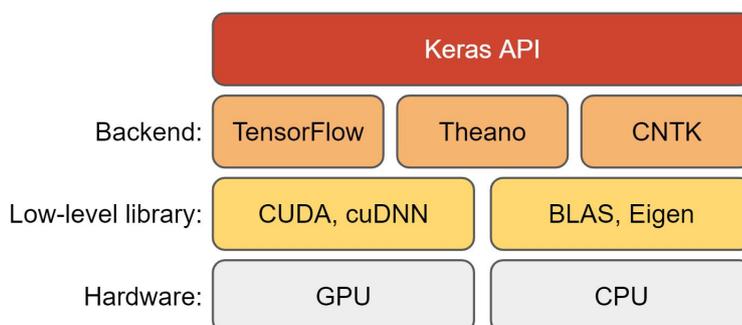
Com o *TensorFlow*, é possível criar uma ampla variedade de modelos de aprendizado de máquina, desde redes neurais simples até redes mais complexas, como as redes convolucionais e recorrentes. O *TensorFlow* também é usado para uma ampla gama de tarefas de aprendizado de máquina, como classificação, detecção de objetos, segmentação de imagem, tradução automática, processamento de linguagem natural, reconhecimento de voz, entre outros.

Além disso, o *TensorFlow* oferece uma ampla gama de ferramentas e recursos para auxiliar no treinamento e validação de modelos de aprendizado de máquina, como o TensorBoard para visualização de dados e o *TensorFlow Hub* para compartilhamento de modelos treinados.

Em resumo, o *TensorFlow* é uma biblioteca de aprendizado de máquina poderosa e versátil, adequada para uma ampla variedade de tarefas de aprendizado de máquina e capaz de ser executada em uma ampla variedade de plataformas e dispositivos.

3.3.2 Keras

Figura 3.7: Funcionamento do *Keras*.



Fonte: <<https://www.devopsschool.com/blog/what-is-keras-and-how-it-works-an-overview-and-its-use-cases-2/>>

O *Keras* é uma Application Programming Interface (API) em código aberto para redes neurais de alto nível, escrita em *Python* e capaz de ser executada no *TensorFlow* e *Theano* (uma biblioteca *Python* usada para definir, otimizar e avaliar expressões matemáticas envolvendo matrizes multidimensionais.).

A Figura 3.7 demonstra o funcionamento do *Keras*, tendo como base o hardware, GPU ou CPU, após isso uma biblioteca de baixo nível, um backend e por fim a API em si.

O *Keras* foi criado com o objetivo de tornar a construção e treinamento de redes neurais mais acessível e fácil para desenvolvedores de todos os níveis de experiência. Ele oferece uma série de recursos que permitem criar rapidamente modelos de redes neurais profundas (modelos de aprendizado de máquina compostos por várias camadas de neurônios que permitem a extração de representações de alta complexidade a partir de dados brutos.), como camadas convolucionais, camadas densas, funções de ativação e otimizadores.

Um dos principais benefícios do *Keras* é sua capacidade de abstrair muitos dos detalhes de baixo nível envolvidos na construção e treinamento de redes neurais, permitindo que os usuários se concentrem em suas ideias e experimentos em vez de se preocupar com a programação de baixo nível. Isso torna o processo de construção de redes neurais mais rápido e intuitivo.

Além disso, o *Keras* é altamente modular e permite que os usuários criem modelos de redes neurais profundas personalizados e adaptados às suas necessidades específicas. Ele também possui uma comunidade ativa e crescente de desenvolvedores, o que significa que há muitos recursos e exemplos disponíveis para ajudar os usuários a começar a trabalhar rapidamente.

Apesar de ser uma excelente ferramenta para construir e treinar redes neurais profundas, o *Keras* tem algumas limitações. Uma delas é que ele não fornece a maioria dos modelos pré-treinados de última geração, o que significa que os usuários precisam construir seus próprios modelos ou adaptar modelos existentes para seus projetos.

Além disso, o *Keras* não é a melhor opção para tarefas que exigem desempenho máximo, como reconhecimento de imagens em tempo real ou processamento de grandes conjuntos de dados.

Em resumo, o *Keras* é uma API em código aberto para redes neurais de alto nível que oferece uma interface amigável e modular para construir e treinar redes neurais profundas. Ele é altamente personalizável e permite que os usuários criem modelos adaptados às suas necessidades específicas. Embora o *Keras* tenha algumas limitações, ele é uma ferramenta valiosa para desenvolvedores e cientistas de dados que desejam explorar o campo das redes neurais profundas.

3.3.3 Spyder

O *Spyder* é uma interface gráfica que fornece um ambiente interativo para utilização do Python, oferecendo recursos semelhantes ao *Matlab*. Ele simplifica a edição de *scripts*, teste, depuração e visualização gráfica de maneira conveniente.

3.4 Base de dados

A DIARETDB1 (*DIAbetic RETinopathy DataBase*) é uma base de dados reconhecida na área de detecção de retinopatia diabética. Ela consiste em 89 imagens, das quais 84 apresentam retinopatia diabética não proliferativa (RDNP) intermediária, caracterizada pela presença de microaneurismas, enquanto as outras 5 imagens não possuem qualquer sinal de retinopatia

diabética. As imagens foram capturadas utilizando uma câmera de retinografia com um ângulo de 50° e possuem dimensões de 1500 × 1152 *pixels*, com uma profundidade de cor de 8 bits por canal RGB.

Essa base de dados fornece informações precisas, pois conta com a anotação de especialistas em retinopatia diabética. Esses especialistas marcaram os microaneurismas, hemorragias, exsudatos duros e manchas algodonosas presentes nas imagens. Além disso, cada marcação recebeu um grau de confiança atribuído pelo especialista (KAUPPI et al., 2007).

A DIARETDB1 é amplamente utilizada como referência em pesquisas relacionadas à detecção de retinopatia diabética, proporcionando uma base sólida para o desenvolvimento e avaliação de algoritmos e técnicas de detecção nesse campo. A disponibilidade de imagens com características específicas e a presença do *ground truth* fornecido pelos especialistas contribuem para uma análise precisa e confiável dos métodos de detecção propostos. Ela foi escolhida justamente por essas vantagens.

3.5 Arquitetura da rede proposta

A segmentação de imagens é uma tarefa importante na análise de imagens médicas e outras áreas de aplicação. Neste trabalho, propomos uma modificação da arquitetura U-Net original para segmentação de imagens. A rede proposta, Figura 3.8, consiste em seis camadas, mantendo a ativação "relu", o *padding* "same" e o *data_format* "channels_first" em todas as camadas.

A primeira e a segunda camada contêm três camadas convolucionais 2D com filtro de tamanho 3x3, uma camada convolucional 2D com filtro de tamanho 1x1, um *dropout* de 0.15 e uma camada de *max pooling 2D* com janela de tamanho 2x2. A primeira camada possui 32 neurônios e a segunda possui 64.

Na terceira camada, há quatro camadas convolucionais 2D com filtro de tamanho 3x3, uma camada convolucional 2D com filtro de tamanho 1x1, um *dropout* de 0.15 e um total de 128 neurônios.

Na quarta camada, temos uma camada de *UpSampling2D* com janela de tamanho 2x2, seguida de quatro camadas convolucionais 2D com filtro de tamanho 3x3, um *dropout* de 0.15 e 64 neurônios.

Na quinta camada, temos uma camada de *UpSampling2D* com janela de tamanho 2x2, seguida de três camadas convolucionais 2D com filtro de tamanho 3x3, uma camada convolucional

2D com filtro de tamanho 1x1, um *dropout* de 0.15 e 32 neurônios.

Na sexta camada, temos uma camada convolucional 2D com filtro de tamanho 1x1, dois neurônios, reshape (2, height * width) e permute (2, 1). A ativação usada foi a *Softmax*. Na compilação, o otimizador foi o *sgd*, o loss foi o *categorical_crossentropy* e as métricas foram *accuracy*.

A rede foi treinada com 200 epochs e um batch size de 1.

Figura 3.8: Arquitetura de rede proposta

```

123 def get_vini_unet(channels, height, width):
124     inputs = Input(shape=(channels, height, width))
125
126     # convolution_2d_1 = Conv2D(32, (3, 3), activation='relu', padding='same', data_format='channels_first')(inputs)
127     convolution_2d_1 = Conv2D(32, (3, 3), activation='relu', padding='same', data_format='channels_first')(inputs)
128
129     # convolution_2d_1 = Dropout(0.25)(convolution_2d_1)
130     convolution_2d_1 = Conv2D(32, (3, 3), activation='relu', padding='same', data_format='channels_first')(convolution_2d_1)
131     convolution_2d_1 = Conv2D(32, (3, 3), activation='relu', padding='same', data_format='channels_first')(convolution_2d_1)
132     max_pooling_1 = MaxPooling2D((2, 2))(convolution_2d_1)
133
134     convolution_2d_2 = Conv2D(64, (3, 3), activation='relu', padding='same', data_format='channels_first')(max_pooling_1)
135     convolution_2d_2 = Dropout(0.25)(convolution_2d_2)
136     convolution_2d_2 = Conv2D(64, (3, 3), activation='relu', padding='same', data_format='channels_first')(convolution_2d_2)
137     convolution_2d_2 = Conv2D(64, (3, 3), activation='relu', padding='same', data_format='channels_first')(convolution_2d_2)
138     max_pooling_2 = MaxPooling2D((2, 2))(convolution_2d_2)
139
140     convolution_2d_3 = Conv2D(128, (3, 3), activation='relu', padding='same', data_format='channels_first')(max_pooling_2)
141     convolution_2d_3 = Conv2D(128, (3, 3), activation='relu', padding='same', data_format='channels_first')(convolution_2d_3)
142     convolution_2d_3 = Dropout(0.25)(convolution_2d_3)
143     convolution_2d_3 = Conv2D(128, (3, 3), activation='relu', padding='same', data_format='channels_first')(convolution_2d_3)
144     convolution_2d_3 = Conv2D(128, (3, 3), activation='relu', padding='same', data_format='channels_first')(convolution_2d_3)
145
146     up_sampling_1 = UpSampling2D(size=(2, 2))(convolution_2d_3)
147     up_sampling_1 = concatenate([convolution_2d_2, up_sampling_1], axis=1)
148     convolution_2d_4 = Conv2D(64, (3, 3), activation='relu', padding='same', data_format='channels_first')(up_sampling_1)
149     convolution_2d_4 = Conv2D(64, (3, 3), activation='relu', padding='same', data_format='channels_first')(up_sampling_1)
150     convolution_2d_4 = Dropout(0.25)(convolution_2d_4)
151     convolution_2d_4 = Conv2D(64, (3, 3), activation='relu', padding='same', data_format='channels_first')(convolution_2d_4)
152     convolution_2d_4 = Conv2D(64, (3, 3), activation='relu', padding='same', data_format='channels_first')(convolution_2d_4)
153
154     up_sampling_2 = UpSampling2D(size=(2, 2))(convolution_2d_4)
155     up_sampling_2 = concatenate([convolution_2d_1, up_sampling_2], axis=1)
156     convolution_2d_5 = Conv2D(32, (3, 3), activation='relu', padding='same', data_format='channels_first')(up_sampling_2)
157     convolution_2d_5 = Dropout(0.25)(convolution_2d_5)
158     convolution_2d_5 = Conv2D(32, (3, 3), activation='relu', padding='same', data_format='channels_first')(convolution_2d_5)
159
160     convolution_2d_6 = Conv2D(2, (1, 1), activation='relu', padding='same', data_format='channels_first')(convolution_2d_5)
161     convolution_2d_6 = Reshape((2, height * width))(convolution_2d_6)
162     convolution_2d_6 = Permute((2, 1))(convolution_2d_6)
163
164     activation_1 = Activation('softmax')(convolution_2d_6)
165
166     model = Model(inputs=inputs, outputs=activation_1)
167
168     model.compile(optimizer='sgd', loss='categorical_crossentropy', metrics=['accuracy'])
169
170     return model

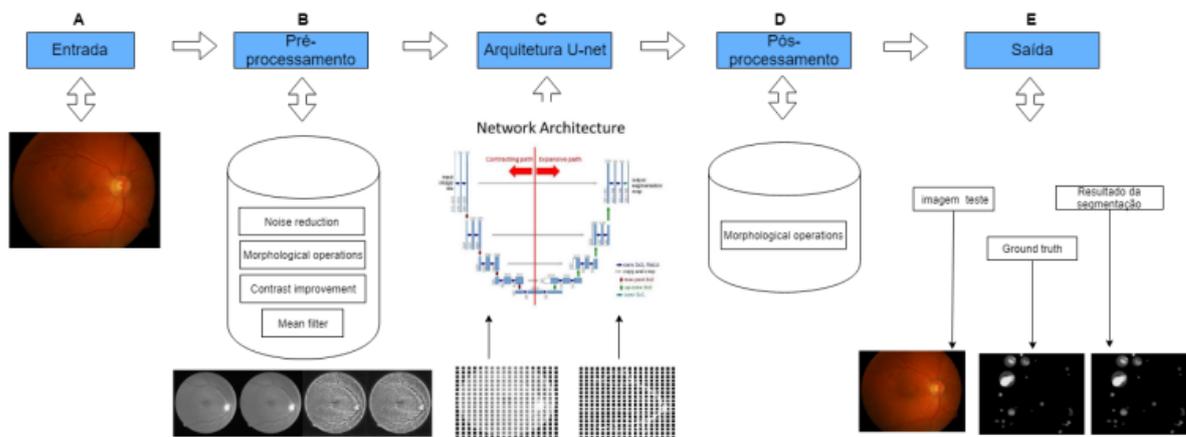
```

No capítulo seguinte serão discutidos os resultados obtidos a partir da arquitetura proposta.

4.1 Etapas do processo de detecção de microaneurismas

A metodologia adotada visa um balanceamento entre aspectos qualitativos e quantitativos e estão distribuídas nos estágios de A a E da Figura 4.1. Inicialmente obtém-se a base de dados, em seguida, realiza-se pré-processamento dos dados de treinamento para se adequarem ao padrão da U-Net, em seguida, na fase de treinamento (estágio C da Figura 4.1), o conjunto de treinamento é separado do conjunto de testes e submetido a arquitetura escolhida com informações significativas sobre o conteúdo sendo buscado na imagem, denomina-se a segmentação manual realizada por um especialista de *ground truth*, esta informação é disponibilizada juntamente com a base de dados DIARETDB1 por (KAUPPI et al., 2007) e é importante pois norteia o aprendizado da rede. Em algumas situações, é realizado o pós-processamento das imagens de saída, sendo que no presente contexto isso não se fez necessário. Por fim, tem-se a saída, na qual gera-se os objetos detectados no conjunto de imagens teste separadamente, no entanto, optou-se por uso de métricas para avaliação no resultado, uma vez que essa abordagem facilita a comparação com outros trabalhos da literatura, portanto, os acertos da rede são automaticamente quantificados sob a métrica *accuracy*, que é amplamente aceita neste tipo de problema.

Acc (*Accuracy*) - como foi já dito, é a taxa de acertos em relação ao total de pixels na área sendo investigada – FOV, sendo que as medidas de verdadeiro positivos e verdadeiro negativos são utilizadas para a obtenção desta métrica. Um verdadeiro positivo (TP) é um pixel assinalado pelo método e pelo especialista como pertencente ao objeto de interesse (neste caso microaneurisma), um verdadeiro negativo é um pixel assinalado pelo método e pelo especialista como não pertencente ao objeto de interesse.

Figura 4.1: Etapas.

4.2 Resultado

Como pode ser observado, a rede U-Net foi adaptada com sucesso, obtendo resultados satisfatórios.

Figura 4.2: Resultados obtidos pela rede.

```
Test score: 0.10105981065167322
Test accuracy: 0.9620127254062228
```

Tabela 4.1: Accuracy entre os modelos

Rede	Accuracy (%)
(DAI et al., 2018)	96.1
Arquitetura proposta	96.2
U-Net padrão	94

Dai et al. (2018) apresentam uma abordagem original para detecção de microaneurismas, que combina recursos de monitoramento clínico com um modelo de CNN híbrido. Essa metodologia permite identificar microaneurismas e realizar a classificação de imagens da retina com alta precisão. Essa eficácia é resultado da identificação de potenciais regiões de microaneurismas no espaço de recursos, seguida pela aplicação de uma rede neural convolucional de triagem múltipla. O desempenho do modelo é comprovado pela alta taxa de precisão e um F1-score próximo a 1.

Diante dos resultados obtidos, pode-se concluir que a abordagem proposta neste trabalho, que consistiu na adaptação da rede neural U-Net para a segmentação de imagens de fundo de

olho, se mostrou eficaz e obteve resultados satisfatórios em relação à acurácia.

Com a utilização de técnicas de pré-processamento, como a normalização e o aumento de dados, e a implementação de melhorias na rede, como a inclusão de camadas adicionais e a utilização de funções de ativação específicas, será possível aprimorar o desempenho da rede em relação ao estado da arte na literatura. Com isso, os resultados obtidos neste trabalho sugerem que a abordagem proposta pode ser útil para auxiliar no diagnóstico em imagens de fundo de olho, permitindo uma segmentação mais precisa das imagens e, conseqüentemente, uma avaliação mais precisa do problema.

Portanto, pode-se concluir que a adaptação da rede U-Net para a segmentação de imagens de fundo de olho é uma abordagem promissora e que pode contribuir significativamente para o avanço da área médica e da saúde como um todo.

Em conclusão, neste trabalho, foi proposta uma adaptação do modelo U-Net para a detecção de microaneurismas em imagens de fundo de olho. Embora não tenha sido possível exibir os microaneurismas detectados em uma imagem reconstruída, os resultados obtidos até o momento são altamente promissores. O modelo U-Net modificado demonstrou uma notável taxa de precisão de 96,2%, o que destaca sua eficácia na identificação dessas estruturas vasculares relevantes.

A modificação realizada na U-Net permitiu explorar características específicas dos microaneurismas, levando a uma melhor capacidade de detecção em comparação com o modelo original, os resultados preliminares são encorajadores e fornecem uma base sólida para futuras pesquisas nessa área. Esses resultados ressaltam o potencial promissor da U-Net como uma ferramenta valiosa na prática clínica, auxiliando médicos e profissionais de saúde no diagnóstico precoce e tratamento de doenças oculares relacionadas aos microaneurismas.

4.3 Trabalhos Futuros

Diante dos resultados obtidos, como trabalhos futuros pretende-se:

Implementar a funcionalidade de segmentar as imagens e separá-las por quadrantes de forma que seja possível contar quantos microaneurismas ocorrem em cada quadrante.

Detectar hemorragias e classificar o nível de retinopatia diabética.

Investigar a utilização de outras redes neurais para a segmentação de imagens de fundo de olho, comparando seu desempenho com a abordagem proposta neste trabalho.

Realizar experimentos em um conjunto de dados maior e mais diverso, a fim de avaliar a robustez e generalização da abordagem proposta.

Investigar a utilização de outras técnicas de pré-processamento e pós-processamento, como a remoção de ruído e a correção de iluminação, para aprimorar ainda mais a precisão da segmentação.

Investigar a utilização de informações adicionais, como dados de OCT (tomografia de coerência óptica), para melhorar a segmentação de estruturas específicas na imagem.

Aplicar a abordagem proposta em imagens de fundo de olho de pacientes com diferentes patologias, a fim de avaliar sua utilidade para o diagnóstico e acompanhamento de doenças oculares.

Investigar a possibilidade de integrar a abordagem proposta em sistemas de telemedicina, permitindo a avaliação remota de imagens de fundo de olho e o auxílio no diagnóstico de pacientes em áreas remotas ou com acesso limitado a especialistas em oftalmologia.

REFERÊNCIAS BIBLIOGRÁFICAS

ABREU, L. R. *Redes neurais convolucionais aplicadas à detecção de objetos no domínio do futebol de robôs humanoides*. Dissertação (Dissertação de Mestrado) — Centro Universitário FEI, São Bernardo do Campo, 2021. Programa de pós-graduação Stricto Sensu em Engenharia Elétrica.

AGARAP, A. F. A neural network architecture combining gated recurrent unit (gru) and support vector machine (svm) for intrusion detection in network traffic data. *arXiv preprint arXiv:1709.03082*, 2017.

ALALSHEKMUBARAK, A.; SMITH, L. S. A novel approach combining recurrent neural network and support vector machines for time series classification. In: *2013 9th International Conference on Innovations in Information Technology (IIT)*. Al Ain, United Arab Emirates: [s.n.], 2013. p. 42–47.

CARDOSO, C.; CARRIJO, G. A.; FERREIRA, J. C.; AL. et. Automatic optic disc segmentation using a 3d morphologic fuzzy histon roughness index. *Research in Biomedical Engineering*, v. 35, p. 89–102, 2019. Disponível em: <<https://doi.org/10.1007/s42600-019-00011-4>>.

CARDOSO, C. D. F. D. S. *SEGMENTAÇÃO AUTOMÁTICA DO DISCO ÓPTICO E DE VASOS SANGUÍNEOS EM IMAGENS DE FUNDO DE OLHO*. Tese (Doutorado) — Universidade Federal de Uberlândia, 2019.

COVER, T.; HART, P. Nearest neighbor pattern classification. *IEEE Transactions on Information Theory*, v. 13, n. 1, p. 21–27, 1967.

DAI, L.; FANG, R.; LI, H.; HOU, X.; SHENG, B.; WU, Q.; JIA, W. Clinical report guided retinal microaneurysm detection with multi-sieving deep learning. *IEEE Transactions on Medical Imaging*, v. 37, n. 5, p. 1149–1161, May 2018.

DAMBROS, W. H. *Modelagem Computacional para a Identificação Não Intrusiva de Cargas Residenciais Utilizando Redes Neurais Artificiais*. Dissertação (Dissertação de Mestrado) — Universidade Regional do Noroeste do Estado do Rio Grande do Sul – UNIJUÍ, Ijuí, 2020. Programa de Pós-graduação em Modelagem Matemática.

DEMAR, J.; ZUPAN, B. Hands-on training about overfitting. *PLoS Computational Biology*, v. 17, 2021.

DESHPANDE, A. *A Beginner's Guide to Understanding Convolutional Neural Networks*. 2016. <<https://adeshpande3.github.io/adeshpande3.github.io/A-Beginner's-Guide-ToUnderstanding-Convolutional-Neural-Networks/>>. Acesso em: 20 jun. 2019.

DORION, T. *Manual de exame do fundo de olho*. [S.l.]: Editora Manole Ltda, 2002.

FERREIRA, A. S. *Redes Neurais Convolucionais Profundas na Detecção de Plantas Daninhas em Lavoura de Soja*. Tese (Doutorado) — Universidade Federal de Mato Grosso do Sul, Campo Grande, 2017. Dissertação de mestrado, Pós graduação em Ciências da Computação.

FRAZ, M.; REMAGNINO, P.; HOPPE, A.; UYYANONVARA, B.; RUDNICKA, A.; OWEN, C.; BARMAN, S. Blood vessel segmentation methodologies in retinal images – a survey. *Computer Methods and Programs in Biomedicine*, v. 108, n. 1, p. 407–433, 2012. ISSN 0169-2607. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0169260712000843>>.

GOVDIABETES. 2023. <<https://www.gov.br/saude/pt-br/assuntos/saude-de-a-a-z/d/diabetes#:~:text=De%20acordo%20com%20a%20Sociedade,%2C9%25%20da%20popula%C3%A7%C3%A3o%20nacional>>. Acesso em 17 de julho de 2023.

HENDRICK, A. M.; GIBSON, M. V.; KULSHRESHTHA, A. Diabetic retinopathy. *Primary Care: Clinics in Office Practice*, v. 42, n. 3, p. 451–464, 2015. ISSN 0095-4543. Primary Care Ophthalmology. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S009545431500038X>>.

IAPB. 2013. <<https://www.iapb.org/learn/resources/diabetic-retinopathy/>>. Acesso em 17 de julho de 2023.

JIMÉNEZ, S.; ALEMANY, P.; BENJUMEA, F. N.; SERRANO, C.; ACHA, B.; FONDÓN, I.; CARRAL, F.; SÁNCHEZ, C. I. [automatic detection of microaneurysms in colour fundus images]. *Archivos de la Sociedad Espanola de Oftalmologia*, v. 86 9, p. 277–81, 2011.

KAUPPI, T.; KALESNYKIENE, V.; KAMARAINEN, J.; LENSU, L.; SORRI, I.; RANINEN, A.; VOUTILAINEN, R.; UUSITALO, H.; AINEN, H.; PIETILA, J. Diaretdb1 diabetic retinopathy database and evaluation protocol. In: UNIVERSITY OF WARWICK. *Proceedings of the British Machine Vision Conference 2007*. Warwick, UK, 2007. p. 61–65. Disponível em: <<https://doi.org/10.5244/C.21.15>>.

KHAN, A.; AL. et. A survey of the recent architectures of deep convolutional neural networks. *Artificial Intelligence Review*, v. 53, p. 5455–5516, 2020.

KRENKER, A.; BESTER, J.; KOS, A. Introduction to the artificial neural networks. In: . [S.l.: s.n.], 2011.

MAXPOOLING2D layer. 2023. <https://keras.io/api/layers/pooling_layers/max_pooling2d/>. Acesso em 17 de julho de 2023.

MCLACHLAN, G. J. *Discriminant Analysis and Statistical Pattern Recognition*. [S.l.]: John Wiley & Sons, 2004. v. 544.

MITTAPALLI, P. S.; KANDE, G. B. Segmentation of optic disk and optic cup from digital fundus images for the assessment of glaucoma. *Biomedical Signal Processing and Control*, v. 24, p. 34–46, 2016. ISSN 1746-8094.

- MOOKIAH, M. R. K.; ACHARYA, U. R.; CHUA, C. K.; MIN, L. C.; NG, E.; MUSHRIF, M. M.; LAUDE, A. Automated detection of optic disk in retinal fundus images using intuitionistic fuzzy histon segmentation. *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine*, SAGE Publications Sage UK: London, England, v. 227, n. 1, p. 37–49, 2013.
- MOOKIAH, M. R. K.; ACHARYA, U. R.; CHUA, C. K.; LIM, C. M.; NG, E.; LAUDE, A. Computer-aided diagnosis of diabetic retinopathy: A review. *Computers in biology and medicine*, Elsevier, v. 43, n. 12, p. 2136–2155, 2013.
- MSDMANUALS. 2022. <<https://www.msdmanuals.com/pt-pt/casa/dist%C3%BArbios-oftalmol%C3%B3gicos/doen%C3%A7as-da-retina/retinopatia-diab%C3%A9tica>>. Acesso em 29 de maio de 2023.
- NIEMEIJER, M.; STAAL, J.; GINNEKEN, B. V.; LOOG, M.; ABRAMOFF, M. D. Comparative study of retinal vessel segmentation methods on a new publicly available database. In: SPIE. *Medical imaging 2004: image processing*. [S.l.], 2004. v. 5370, p. 648–656.
- OMS. 2021. <<https://apps.who.int/iris/bitstream/handle/10665/328717/9789241516570-por.pdf>>. Acesso em 17 de julho de 2023.
- PAVLOVIČOVÁ, J.; AL. et. Bright lesions detection on retinal images by convolutional neural network. In: IEEE. *PROCEEDINGS ELMAR - INTERNATIONAL SYMPOSIUM ELECTRONICS IN MARINE 2018*. 2018. Disponível em: <<https://ieeexplore.ieee.org/document/8534658/>>.
- PCDT. 2021. <https://www.gov.br/saude/pt-br/assuntos/protocolos-clinicos-e-diretrizes-terapeuticas-pcdt/arquivos/2021/portal-portaria-conjunta_pcdt_retinopatia-diabetica_.pdf>. Acesso em 17 de julho de 2023.
- PEDRYCZ, W.; GOMIDE, F. *Fuzzy Systems Engineering: Toward Human-Centric Computing*. New York: John Wiley & Sons, 2007.
- PURUSHARATHI, S.; KURHADE, B. Identification of diabetic retinopathy in retinal images using support vector machine. *International Journal of Computing and Technology*, v. 1, n. 2, p. 158–167, 2014. Disponível em: <<http://www.ijcat.org/IJCAT-2014/1-2/Identification-of-Diabetic-Retinopathy-in-Retinal-Images-using-Support-Vector-Machine.pdf>>.
- RODRIGUES, D. A. *Deep Learning e Redes Neurais Convolucionais: Reconhecimento automático de caracteres em placas de licenciamento automotivo*. Dissertação (Dissertação de Mestrado) — Universidade Federal da Paraíba – Centro de Informática, João Pessoa, 2018.
- RONNEBERGER, O.; FISCHER, P.; BROX, T. U-net: Convolutional networks for biomedical image segmentation. In: NAVAB, N.; AL. et (Ed.). *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*. Cham: Springer International Publishing, 2015. p. 234–241.
- SANTOS, J. C. M. dos; CARRIJO, G. A.; CARDOSO, C. de Fátima dos S.; AL. et. Fundus image quality enhancement for blood vessel detection via a neural network using clahe and wiener filter. *Research in Biomedical Engineering*, v. 36, p. 107–119, 2020. Disponível em: <<https://doi.org/10.1007/s42600-020-00046-y>>.

SAUDEBEMESTAR. 2023. <<https://www.saudebemestar.pt/pt/clinica/oftalmologia/retinografia/>>. Acesso em 17 de julho de 2023.

SHELHAMER, E.; LONG, J.; DARRELL, T. Fully convolutional networks for semantic segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 39, n. 4, p. 640–651, 2017.

SRIVASTAVA, N.; AL. et. Dropout: A simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research*, v. 15, p. 1929–1958, 06 2014.

TANG, Y. Deep learning using linear support vector machines. *arXiv preprint arXiv:1306.0239*, 2013.

VERGARA, R. F. *Detecção de alterações cerebrais anatômicas associadas à esquizofrenia com base em redes convolucionais aplicadas a imagens de ressonância magnética*. Dissertação (Monografia (Pós-Graduação)) — Faculdade Gama - FGA, Universidade de Brasília, Brasília, DF, 2018. 89 p.

WANG, S.; YIN, Y.; CAO, G.; WEI, B.; ZHENG, Y.; YANG, G. Hierarchical retinal blood vessel segmentation based on feature and ensemble learning. *Neurocomputing*, v. 149, n. Part B, p. 708–717, 2015. ISSN 0925-2312.