

Walef Alves da Silva

Orientador: Romuere Rodrigues Veloso e Silva

Segmentação do Disco Óptico em Imagens Retinoscópicas Utilizando a Rede Neural Convolutiva U-Net

Picos - PI

21 de novembro de 2019

Walef Alves da Silva
Orientador: Romuere Rodrigues Veloso e Silva

Segmentação do Disco Óptico em Imagens Retinoscópicas Utilizando a Rede Neural Convolutacional U-Net

Monografia submetida ao Curso de Bacharelado em Sistemas de Informação como requisito parcial para obtenção de grau de Bacharel em Sistemas de Informação.

Universidade Federal do Piauí
Campus Senador Helvídio Nunes de Barros
Bacharelado em Sistemas de Informação

Picos - PI
21 de novembro de 2019

FICHA CATALOGRÁFICA
Serviço de Processamento Técnico da Universidade Federal do Piauí
Biblioteca José Albano de Macêdo

S586s Silva, Walef Alves da.
Segmentação do disco óptico em imagens retinoscópicas utilizando a rede neural convolucional U-Net. / Walef Alves da Silva. -- Picos,PI, 2019.
34 f.
CD-ROM: 4 ¾ pol.

Trabalho de Conclusão de Curso (Bacharelado em Sistemas de Informação). – Universidade Federal do Piauí, Picos, 2020.

“Orientador(A): Prof. Dr. Romuere Rodrigues Veloso e Silva.”

1. Retinopatia Diabética. 2. Segmentação - Imagem Digital.
3. U-Net. I. Título.

CDD 005.1

Elaborada por Rafael Gomes de Sousa CRB 3/1163

SEGMENTAÇÃO DO DISCO ÓPTICO EM IMAGENS RETINOSCÓPICAS UTILIZANDO A
REDE NEURAL CONVOLUCIONAL U-NET

WALEF ALVES DA SILVA

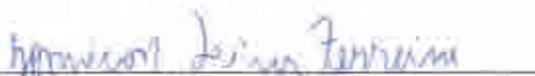
Monografia Aprovada como exigência parcial para obtenção do
grau de Bacharel em Sistemas de Informação.

Data de Aprovação

Picos - Pl. 03 de dezembro de 2019


Prof(a). Romere Rodrigues Veloso e Silva


Prof(a). Flavio Henrique Duarte de Araujo


Prof(a). Jonnison Lima Ferreira

Agradecimentos

Queria agradecer inicialmente aos meus pais Francilene Alves da Cruz e Adonias Oliveira da Silva, por todo esforço e dedicação para juntos conseguirmos a realização desse sonho não só meu mas também deles.

Nessa minha trajetória tem algumas pessoas importantes, como meu irmão Walison que me motivou a entrar no curso e foi fundamental na minha adaptação, não poderia deixar de citar também a importância da Larissa (minha namorada) que de todas as pessoas ela é a que mais sabe de tudo que eu passei no curso.

Não poderia deixar de agradecer a família que eu ganhei aqui em Picos pessoas com toda certeza quero levar para minha vida inteira, eles sempre estiveram presentes em meus momentos, a turma do corredor, o lixão, meus amigos de bar, a galera do PAAD e do DTI, um cara importante demais nessa trajetória foi Matheus um cara muito esforçado e com um grande coração.

Queria agradecer também ao meu orientador Prof. Dr. Romuere Rodrigues Veloso e Silva por todo o suporte e conhecimento acadêmico transferido ao longo do desenvolvimento desse trabalho.

Não poderia deixar de agradecer também a família DNA meus amigos irmãos que conhecem toda a minha história e minhas dificuldades.

*A tarefa não é tanto ver aquilo que ninguém viu, mas pensar o que ninguém ainda
pensou sobre aquilo que todo mundo vê.*

Arthur Schopenhauer

Resumo

A Retinopatia Diabética (RD) é uma complicação comum e potencialmente prejudicial a visão do diabético e é uma das principais doenças causadoras de cegueira. Um diagnóstico precoce da RD para localizar lesões em seu estado inicial pode fazer com que alguns tratamentos reduzam o risco da perda da visão. O diagnóstico da RD pode se dar por métodos computacionais automatizados através de imagens retinoscópicas como também pode ser feita de maneira manual através da análise de um especialista. Quando o processo de identificação de lesões (microaneurismas, exsudatos duros, moles e disco óptico (DO)) é executado de maneira automática, a análise destas torna-se mais concisa e precisa, auxiliando a elaboração do diagnóstico do estágio da doença. Este trabalho analisa o desempenho da Rede Neural Convolutiva (RNC) U-Net baseado no modelo original da arquitetura com o ajuste na quantidade de camadas para melhorar o seu desempenho no problema da segmentação do disco óptico em imagens retinoscópicas contendo ou não retinopatia diabética. Os resultados obtidos utilizando as bases *I-DRID*, *RIM-ONE* e *DRIONS-DB* foram de dice superior a 85% e acurácia superior a 87% mostrando ser uma metodologia promissora para a segmentação do disco óptico.

Palavras-chaves: Retinopatia Diabética, Segmentação, U-Net.

Abstract

Diabetic retinopathy (DR) is a common and potentially detrimental complication in diabetic vision and is a major disease causing blindness. An early diagnosis of DR to locate lesions in their early stages may cause some treatments to reduce the risk of vision loss. The diagnosis of DR can be done by automated computational methods through retinoscopic images as well as by manual analysis by a specialist. When the process of identifying lesions (micro-aneurysms, hard exudates, moles and optic disc (DO)) is performed automatically, their analysis becomes more concise and accurate, helping to diagnose the stage of the disease. This paper analyzes the performance of U-Net Convolutional RedeNeural Conventional (RNC) based on the original architecture model as adjustment in the number of layers to improve its performance in the problem of optical disc segmentation in retinoscopic images with or without diabetic retinopathy. The results obtained using the bases I-DRID, RIM-ONE and DRIONS-DB had an index superior to 85% and an accuracy superior to 87% showing a promising methodology for the optical disc segmentation.

Lista de ilustrações

Figura 1 – Representação da Arquitetura das Redes Convolucionais.	17
Figura 2 – Representação da Camada de Convolução.	18
Figura 3 – Representação da Camada de Subamostragem.	18
Figura 4 – Representação da Arquitetura U-Net.	20
Figura 5 – Etapas do método proposto..	23
Figura 6 – Representação do resultado da segmentação com a base <i>I-DRID</i>	24
Figura 7 – Representação do resultado da segmentação com a base <i>RIM-ONE</i>	24
Figura 8 – Representação do resultado da segmentação com a base <i>DRIONS-DB</i>	25
Figura 9 – U-Net arquitetura. Cada caixa azul corresponde a um mapa de recursos multicanal. O número de canais é indicado na parte superior da caixa. O tamanho X-Y é fornecido no canto inferior esquerdo da caixa, caixas brancas representam mapas de recursos copiados. As setas denotam as diferentes operações (RONNEBERGER PHILIPP FISCHER, 2015).	26
Figura 10 –Gráfico do resultado da validação	30

Lista de tabelas

Tabela 1 – Resumo dos trabalhos relacionados.	22
Tabela 2 – Resultados obtidos no método proposto com a U-Net treinada e testada apenas como uma única base.	28
Tabela 3 – Resultados obtidos no método proposto com a U-Net com variações na base de teste.	29
Tabela 4 – Resultados obtidos no método proposto com a U-Net com variações na base de treino.	29
Tabela 5 – Comparação do método proposto com alguns trabalhos relacionados. .	30

Lista de abreviaturas e siglas

A	Acurácia
RD	Retinopatia diabética
DM	Diabetes Mellitus
RNA	Rede Neural Artificial
FCN	Rede Completamente Convolutacional
ReLU	Unidade Linear Retificada
DO	Disco Óptico
RNC	Rede Neural Convolutacional
CAC	Campo Aleatório Condicional
RGB	Rede Green Blue
OMS	Organização Mundial de Saúde
MLP	Percepção de multicamadas

Lista de símbolos

Sumário

1	Introdução	13
1.1	Objetivos	13
1.1.1	Objetivo Geral	13
1.1.2	Objetivos Específicos	13
1.2	Organização do Trabalho	14
2	Referencial Teórico	15
2.1	Retinopatia diabética	15
2.2	Redes Neurais Convolucionais	16
2.2.1	Rede Completamente Convolucional	19
2.3	U-Net	19
3	Trabalhos Relacionados	21
4	Materiais e Métodos	23
4.1	Base de Imagens	23
4.1.1	<i>I-DRID</i>	24
4.1.2	<i>RIM-ONE</i>	24
4.1.3	<i>DRIONS-DB</i>	24
4.2	Segmentação	25
4.2.1	U-Net	25
4.3	Métricas de desempenho	26
4.3.1	Validação	27
5	Resultados e Discussões	28
6	Conclusão e Trabalhos Futuros	31
	Referências	32

1 Introdução

A retinopatia diabética (RD) proliferativa é uma complicação comum e potencialmente prejudicial à visão do diabético, indivíduos incapazes de controlar a glicemia tendem a sofrer com a retinopatia (B. BEAULIEU WESLEY T., 2019).

A RD é uma das principais causas de cegueira. Apesar dos tratamentos disponíveis, alguns pacientes só descobrem a doença tardiamente dificultando o tratamento. Se a retinopatia for detectada no início, alguns fatores de risco modificáveis (glicose, pressão arterial) podem retardar a progressão e reduzir o risco da cegueira (SCANLON, 2019).

A detecção da RD se dá pela presença de lesões na retina. A segmentação de microaneurismas, exsudatos duros e moles e do disco óptico é crucial no diagnóstico da RD. No entanto, tamanhos variados, características semelhantes (intensidade e cor), variação da localização e formas tornam este trabalho desafiador. A delimitação manual feita por oftalmologistas em imagens retinoscópica é demorada e exaustiva, o que leva a uma alta demanda para desenvolver um método de segmentação automática eficaz e preciso (XUE SHUO YAN, 2019).

Com base na necessidade e a importância, apontadas acima, da segmentação como processo de análise e diagnóstico da retinopatia diabética em imagens retinoscópicas. Este trabalho emprega a Rede Neural Convolutiva (RNC) U-Net no problema da segmentação do disco óptico em imagens que contém ou não a RD. O desempenho deste trabalho foi avaliado utilizando as métricas: Acurácia (PROVOST; DOMINGOS, 2000) e o dice (GINSBERG; YOUNG, 1992)

1.1 Objetivos

1.1.1 Objetivo Geral

A partir do contexto apresentado, o presente trabalho possui como objetivo principal aplicar uma RNC U-Net para a segmentação do disco óptico em imagens doentes ou saudáveis.

1.1.2 Objetivos Específicos

1. Avaliar o desempenho da segmentação utilizando a RNC U-NET.
2. Gerar uma segmentação coerente a máscara disponibilizada pelo especialista.
3. Construir uma arquitetura que forneça ao especialista uma nova proposta na segmentação do disco óptico.

4. Testar o desempenho do método em diferentes bases.

1.2 Organização do Trabalho

Além da introdução, este trabalho está dividido na seguinte ordem: no Capítulo 2, apresenta-se o referencial teórico; no Capítulo 3, são apresentados os trabalhos relacionados; no Capítulo 4, descreve-se a metodologia proposta; no Capítulo 5, são descritos os resultados e discussões da execução do projeto; e por fim, são apresentadas as conclusões e sugestões para trabalhos futuro no Capítulo 6.

2 Referencial Teórico

Este capítulo apresenta a fundamentação teórica utilizada no desenvolvimento deste trabalho contendo conceitos, conhecimentos e informações pertinentes ao trabalho e necessária para compreensão das técnicas utilizadas para alcançar os objetivos esperados pelo mesmo.

2.1 Retinopatia diabética

Segundo a Organização Mundial de Saúde (OMS), a RD é uma doença que pode se desenvolver em pacientes diabéticos que estão em uma faixa etária de idade denominada ativa, casos que ocorrem em pacientes com menos de 20 anos e com mais e 70 anos não são comuns. Esta doença é normalmente progressiva e possui 5 estágios, o primeiro é caracterizado pela ausência da doença em pacientes que possuem diabetes, no segundo estágio o paciente apresenta pequenas lesões acompanhadas de hemorragia, mas de forma discreta. o terceiro estágio é denominado moderado, e já possui diversos tipos de aneurismas e hemorragias. O quarto estágio já está em um nível avançado da doença, neste caso o tratamento já se faz necessário devido a diversas lesões como hemorragia, aneurismas e má formação de vasos vasculares, entre outros. O último estágio da doença já traz complicações maiores e pode levar a cegueira do paciente, desta maneira já existe a necessidade de cirurgias, injeções, etc.

A RD tem como a principal causa de problemas visuais o edema macular, o mesmo pode estar presente desde o início da retinopatia até em estágios mais avançados da doença, atacando 30% dos pacientes com mais de 20 anos de diabetes. O estágio mais avançado da doença é o proliferativo que é aquele que se relaciona mais frequente com a perda visual grave, devido a eventos oculares possivelmente causadores de cegueira irreversível incluindo o edema macular e o descolamento tracional de retina. Estima-se que em olhos com RD proliferativa não tratada a taxa de evolução para cegueira seja de 50%, em 5 anos.([OLIVEIRA RENAN MAGALHÃES MONTENEGRO JUNIOR, 2018](#))

"A RD é umas das principais complicações relacionadas ao diabetes mel-litus (DM) é a principal causa de cegueira em pessoas com idade entre 20 e 74 anos. Aproximadamente 12% dos novos casos de cegueira legal,isto é, a diminuição da acuidade visual a um nível que impeça o exercício de atividades laborais, são causados pela RD. Após 20 anos de doença,mais de 90% dos diabéticos tipo 1 e 60% daqueles com o tipo 2 apresentarão algum grau de retinopatia."([OLIVEIRA RENAN MAGALHÃES MONTENEGRO JUNIOR, 2018](#)).

Shu, Chui e FRANZCO (2016) citam em seu trabalho uma estimativa feita pela OMS onde a RD é responsável por cerca de 37 milhões de casos de cegueira registrados em todo

o mundo, essa quantidade é equivalente a 4,8% dos casos. Ainda nesse trabalho, foi feita uma análise conjunta de 22.896 pessoas com diabetes que residem nos EUA, Austrália, Europa e Ásia, onde se concluiu que a prevalência geral de qualquer DR é de 34,6%.

Graças a duas importantes séries de estudos é que foi possível se basear para elaborar definições como a de tratamento com fotocoagulação a laser e os critérios necessários para o diagnóstico e caracterização da RD. Os dois estudos em questão são: Diabetic Retinopathy Study (DRS – 1976 a 1979) e Early Treatment Diabetic Retinopathy (ETDRS – 1985 a 1997), o primeiro foi responsável pela definição de critérios diagnósticos para a RD proliferativa, enquanto o segundo definiu critérios utilizados referentes à RD não proliferativa e ao edema macular, assim como a utilização da fotocoagulação a laser nesses pacientes. Buscando menor fragmentação e maior padronização na classificação da RD ETDRS (FONG MD, 2016).

2.2 Redes Neurais Convolucionais

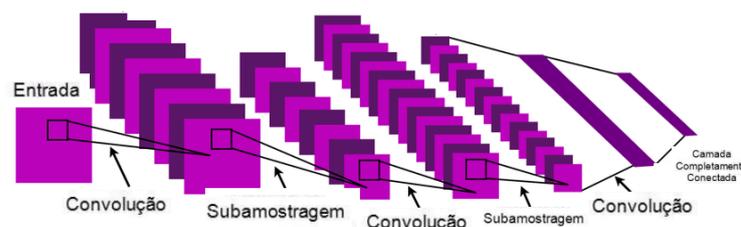
As redes neurais convolucionais (CNN - do inglês *Convolutional Neural Network*) são modelos biologicamente inspirados que podem aprender características de forma hierárquica (LECUN; BENGIO; HINTON, 2015). Este tipo de rede neural vem sendo amplamente utilizada em aplicações de segmentação, classificação, detecção e reconhecimento em imagens e vídeos. As mesmas são especialmente projetadas para lidar com a variabilidade em dados bidimensionais (2D), como as imagens em formato matricial. Contudo, as RNCs tem como característica sua versatilidade, e possuem a capacidade de processar em diversos formatos de matrizes: 1D para sinais e sequências (linguagem); 2D para imagens ou espectrogramas de áudio; e 3D para imagens em vídeos ou volumétricas. As RNCs possuem quatro princípios que a aproximam de sinais naturais: conexões locais, pesos compartilhados, agrupamento e o uso de muitas camadas (ALMOUSLI, 2014; LECUN; BENGIO; HINTON, 2015; PERONA; FINK, 2013).

De forma geral uma RNC tem sua arquitetura composta de três tipos de camadas: convolução, subamostragem e completamente conectada (KANG; WANG, 2014). As camadas convolucionais são responsáveis pela extração de características e a camada de subamostragem é responsável pela redução de dimensionalidade dos dados a serem classificados. As características são extraídas de uma camada para outra da rede, onde um neurônio na camada atual conecta-se a um campo receptivo local na camada anterior. Dessa forma, as características são combinadas sequencialmente para, assim, serem extraídas características de mais alto nível (FERNANDES, 2013). Essas camadas são seguidas por camadas completamente conectadas, análogas à percepção de multicamada (MLP), responsável por fazer a classificação com base nas características extraídas nas camadas anteriores.

A Figura 1 ilustra a arquitetura de uma RNC. Onde a entrada da rede é normalmente uma matriz tridimensional com altura e largura (de acordo com as dimensões da imagem)

e profundidade, determinada pela quantidade de canais de cores. geralmente as imagens utilizam três canais, RGB, com os valores de cada pixel. As convoluções funcionam como filtros que enxergam pequenos quadrados e vão percorrendo por toda a imagem aprendendo os traços mais marcantes como representado na Figura 2. As funções de ativação servem para trazer a não-linearidades ao sistema, para que a rede consiga aprender qualquer tipo de funcionalidade. Há muitas funções, como sigmoid, tanh e softmax, mas a mais indicada para redes convolucionais é a Relu por ser mais eficiente computacionalmente sem grandes diferenças de acurácia quando comparada a outras funções. A camada de pooling serve para simplificar a informação da camada anterior. Assim como na convolução. A unidade é responsável por resumir a informação daquela área em um único valor. Ao final da rede é colocada uma camada completamente conectada, onde sua entrada é a saída da camada anterior e sua saída são N neurônios, com N sendo a quantidade de classes do seu modelo para finalizar a classificação.

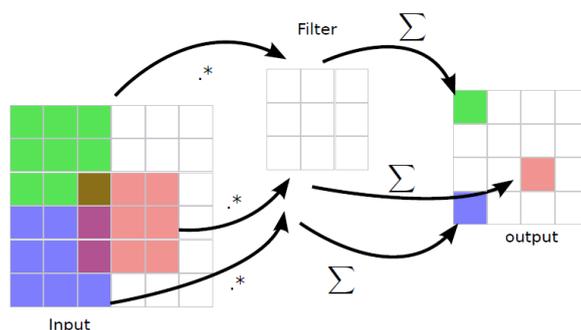
Figura 1: Representação da Arquitetura das Redes Convolucionais.



Fonte: Adaptado de (LECUN et al., 2010)

As camadas de convolução são compostas por filtros treináveis que são aplicados em toda a imagem de entrada, gerando assim, vários mapas de características (LECUN et al., 2010). Os filtros determinam uma pequena área (3×3 , 5×5 , 7×7 pixels) e cada neurônio é ligado somente aos neurônios nas proximidades da camada anterior. Os neurônios compartilham os pesos entre si, levando os filtros a aprender os padrões frequentes que ocorrem em qualquer parte da imagem. Posteriormente ao treinamento da rede, os filtros tornam-se responsáveis pela detecção de características particulares que ocorrem na imagem (HAFEMANN, 2014a). Uma vez que a característica é detectada, sua localização exata na imagem de entrada torna-se irrelevante (LECUN et al., 1998). A Figura 2 ilustra como ocorre a convolução em uma imagem.

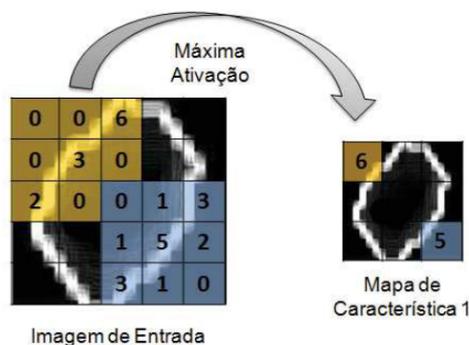
Figura 2: Representação da Camada de Convolução.



Fonte: (HAFEMANN, 2014b)

De acordo com Hafemann (2014a), as camadas de subamostragem implementam uma função de redução da solução não-linear, a fim de reduzir a dimensionalidade selecionando características invariantes à deslocamentos e distorções. A Figura 3 apresenta o processo da camada de subamostragem do tipo máxima ativação, em que apenas o pixel de maior intensidade do campo receptivo anterior é mantido. Existem outras formas de ativação na camada de subamostragem, como por exemplo: média, mediana, mínima e entre outras.

Figura 3: Representação da Camada de Subamostragem.



Fonte: (SILVA et al., 2017)

Após as operações realizadas nas camadas de convolução e subamostragem que são encarregadas de extrair atributos da imagem, os *pixels* de todos os mapas de características da camada anterior são dados como entrada às camadas completamente conectadas, que por sua vez, são responsáveis pela classificação dos padrões de entrada, semelhante à MLP.

Para o problema de classificação de imagens, as arquiteturas baseadas no modelo mostrado na Figura 1 determinam o estado da arte. Long, Shelhamer e Darrell (2015) propôs uma arquitetura sem camadas completamente conectadas: a rede completamente convolucional (FCN do inglês *Fully convolutional networks*). Esta rede gera uma saída nas mesmas dimensões da imagem de entrada, realizando uma segmentação *pixel-a-pixel* da imagem de entrada.

2.2.1 Rede Completamente Convolutacional

Uma Rede completamente convolutacional(RNC) usa uma rede neural convolutacional para realizar a transformação dos pixels de imagem em categorias de pixel. Diferentemente das redes neurais convolutacionais citada na Seção 2.2, uma FCN converte a altura e a largura do mapa de recursos da camada intermediária de volta ao tamanho da imagem de entrada através da camada de uma convolução transposta, de modo que as previsões tenham uma correspondência individual com a imagem de entrada em dimensão espacial (altura e largura). Dada uma posição na dimensão espacial, a saída da dimensão do canal será uma previsão de categoria do pixel correspondente ao local. (Long; Shelhamer; Darrell, 2015)

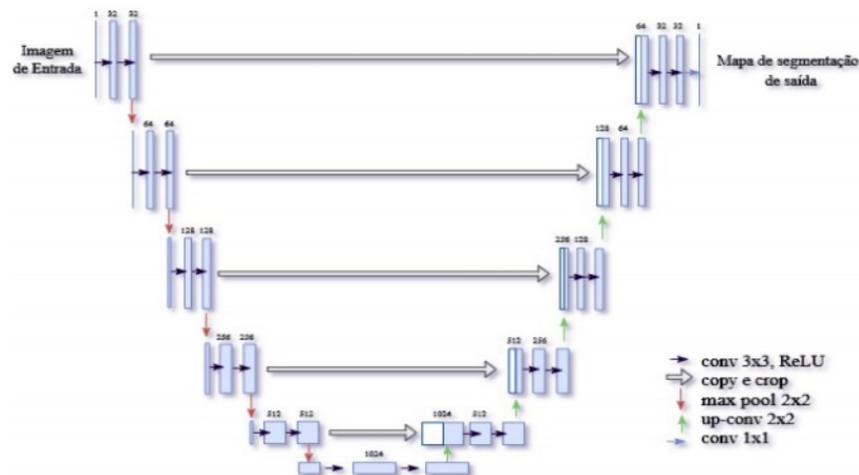
Uma FCN indica que a rede neural é composta de camadas convolutacionais sem nenhuma camada totalmente conectada ou MLP normalmente encontrada no final da rede, elas possuem em sua estrutura apenas convoluções e suas funções de ativação (sigmoid, relu e etc). As redes convolutacionais, são treinadas de ponta a ponta, pixels a pixels, que tenta regenerar tanto a entrada quanto a saída, verifica-se que camadas convolutacionais em diferentes níveis caracterizam o contexto sob diferentes perspectivas. Uma camada superior realiza a propagação de informações, enquanto uma camada inferior carrega as informações e pode separar melhor os pixels com aparência semelhante. Ambas as camadas são usadas em conjunto para realizar a segmentação. as FCNs são comumente utilizadas como auto-codificadores e também para tarefas supervisionadas, como classificação e segmentação. Dentre as arquiteturas FCNs famosas podemos ressaltar as utilizadas para segmentação como a UNet e a ResNet. (WANG et al., 2015)

2.3 U-Net

Ronneberger, Fischer e Brox (2015) apresentaram a U-Net, que é uma RNA (Rede Neural Artificial) completamente convolutacional. A mesma teve em sua concepção, o objetivo de realizar a segmentação de imagens biomédicas. A U-Net é uma rede totalmente convolutacional, cuja a estrutura consiste em um caminho de contração para registrar o contexto, e um caminho de expansão simétrico que possibilita uma segmentação com exatidão.

A U-Net distinguiu-se das demais arquiteturas FCN (Rede Completamente Convolutacional) pela junção entre os mapas de características do estágio de contração e seus equivalentes simétricos no estágio de expansão, permitindo a propagação de informações de contexto para os mapas de atributos de alta resolução. Em diversos desafios biomédicos de segmentação de imagens o presente modelo vem se destacando entre os mais eficientes métodos existentes (RONNEBERGER; FISCHER; BROX, 2015). A Figura 4 ilustra a arquitetura e as camadas que constituem a rede.

Figura 4: Representação da Arquitetura U-Net.



Fonte: (RONNEBERGER; FISCHER; BROX, 2015)

A arquitetura consiste em um caminho de contração (lado esquerdo) para capturar o contexto e um caminho de expansão (lado direito) proporcional que permite uma localização precisa. O caminho de contração é uma arquitetura comum de uma CNN, utilizando de inúmeras aplicações de duas convoluções 3x3, cada uma acompanhada por uma função de ativação Unidade Linear Retificada (ReLU) e uma operação de subamostragem com ativação máxima de 2x2 reduzindo a dimensão da imagem, com passo 2 para subamostragem. Em cada uma dessas etapas, o número de canais de características é duplicado.

O caminho expansivo baseia-se em um levantamento do mapa de características seguido de um sobre-amostragem 2x2 que faz a metade do número de canais de características, que é uma união com o mapa de características correspondentemente cortado do caminho de contratação e duas convoluções 3x3, cada uma seguida por uma função de ativação ReLU. Essa etapa do caminho expansivo é importante devido à perda de *pixels* de borda em cada convolução no caminho de contratação. Na última camada é utilizada uma convolução 1x1 para mapear cada vetor de características para o número desejado de classes.

Por último, essa rede é capaz de ser treinada do início ao fim a partir de poucas imagens, onde a U-Net simplesmente correlaciona os mapas de características do codificador para mapear mapas de características do decodificador em todas as etapas para gerar uma estrutura como escada. Essa arquitetura por suas conexões de concatenação permite que o decodificador em cada etapa aprenda as características pertinentes que são perdidas quando agrupadas no codificador. No presente trabalho, a U-Net foi empregada para gerar a segmentação da região do Disco Óptico em imagens de RD.

3 Trabalhos Relacionados

Neste capítulo expomos trabalhos que abordam problemas semelhantes ao nosso, segmentação do disco óptico em imagens retinoscópicas que contem ou não RD.

Xue *et al.* (2019) em seu trabalho propõe um novo método automático de segmentação de microaneurismas, exsudatos duros e DO para a avaliação diagnóstica da RD. Este método de segmentação multitarefa é baseado num novo sistema de membrana denominada de sistema de membrana dinâmica com estruturas híbridas, os resultados obtidos para a segmentação do DO foram: 96,7% para o dice e 98,1% para sensibilidade.

Wen-Bo *et al.* (2018) propôs um método de detecção de DO baseado em Convolutional Neural Network (CNN) e o Campo Aleatório Condicional (CAC). As imagens retinoscópicas foram pré-classificadas pela CNN e uma função de potência de primeira ordem foi construída do CAC. Em seguida, a combinação linear das funções do kernel gaussiano foram usadas para construir a função de segunda ordem CAC. Ao fim, aplicou-se o método de restrições regionais para à análise da consistência dos rótulos e corrigir o rótulo de cada pixel calculando a média de probabilidade posterior de super *pixels*, onde foi obtido 98,90% de acurácia.

Kowsalsya *et al.* (2018) em seu trabalho, implementou uma técnica de exame para extrair o DO para avaliar sua condição. Para minerar a região do DO a partir de imagens RGB da retina foi aplicado uma abordagem baseada na combinação da entropia de Kapur e agrupamento K-means. Após a extração do DO, análises foram realizadas comparando com as verificações de fundo do especialista, os resultados obtidos foram de 93,93% de dice e 96,73% de sensibilidade.

Xue *et al.* (2017) em seu estudo, desenvolveu um novo método de detecção do disco óptico usando modelo de saliência baseado em agrupamento, para simular o mecanismo de filtragem humana de sistema visual para detecção do DO de imagens do fundo do olho. Primeiro, candidatos a DO são extraídos com K-means. Segundo, duas saliências da imagem são selecionadas como a região DO. Terceiro, o contorno do DO original é extraído por elipse. Último, o contorno do DO pode ser segmentado com precisão por contorno ativo, o resultado obtido foi de 94% de acurácia.

A abordagem aqui proposta utiliza de uma rede específica para segmentação de imagens biomédicas que pode ser treinada de ponta a ponta com um conjunto pequeno, por isso o tempo de execução é diminuído onde todos os testes realizados foram rodados no Collab. É importante ressaltar que alguns dos trabalhos relacionados realizam a segmentação do disco óptico e também de microaneurismas.

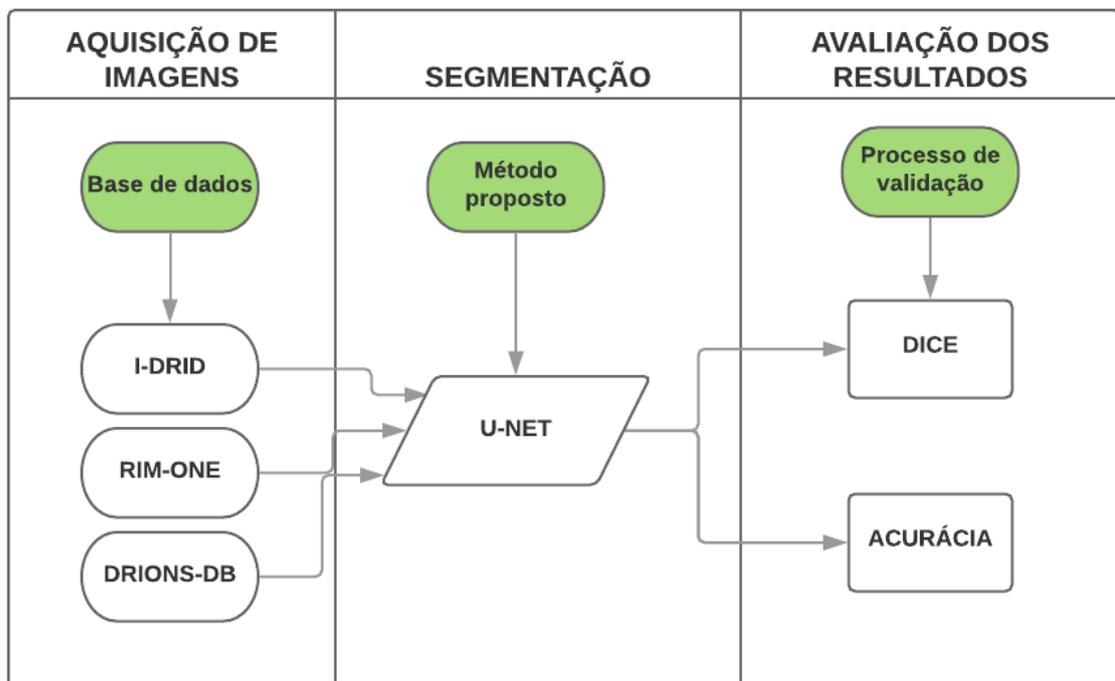
Tabela 1: Resumo dos trabalhos relacionados.

Trabalho	Método	<i>Dataset</i>	Amostras	Acurácia
(XUE SHUO YAN, 2019)	Segmentação Automática	IDRID	81	-
(WEN-BO DUNWEI WEN, 2018)	CNN	DRIONS	100	98,90%
(KOWSALYA A. KALYANI, 2018)	Kapur e K-means	RIM-ONE	169	-
(XUE JIA-WEN LIN, 2017)	Saliência baseada em agrupamento	MESSIDOR	1200	94,00%
Método proposto	U-Net	IDRID	54	99,52%

4 Materiais e Métodos

Este capítulo apresenta os métodos propostos para segmentação do disco óptico, a partir de imagens médicas do fundo do olho. Esse método é composto por três etapas, como pode ser observado na [Figura 5](#). Em síntese, a primeira etapa descreve os materiais utilizados, como as bases de imagens de RD. Na segunda etapa é realizado a segmentação. Na terceira e última etapa, o resultado da segmentação é avaliado.

Figura 5: Etapas do método proposto..



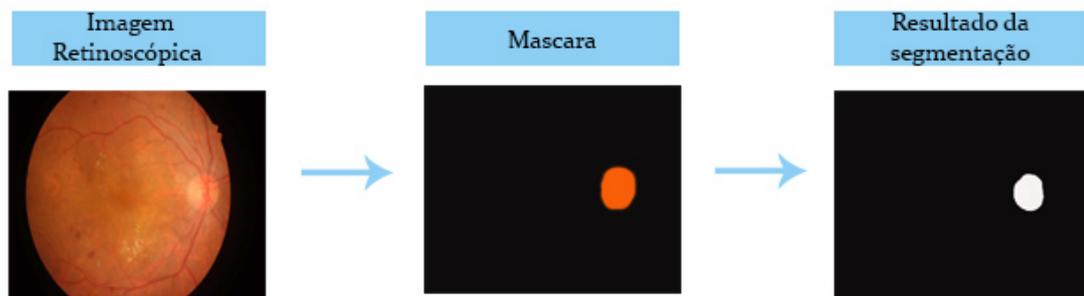
4.1 Base de Imagens

Foram 3 os conjuntos de dados utilizados neste trabalho, onde as imagens foram obtidas através da retinografia digital. No exame de retinografia ocular as imagens obtidas estão as cores se encontram em alta resolução, permitindo desta forma ver com detalhe todo o fundo ocular (retina), as imagens são capturadas pelo retinógrafo, por mais que sejam capturadas pelo mesmo exame possuem dimensões diferentes para cada base.

4.1.1 I-DRID

O dataset IDRID disponível como parte do desafio “IDRID - Diabetic Retinopathy”, organizado em conjunto com o IEEE International Symposium on Biomedical Imaging (IEEE, 2018). O dataset além das imagens fornece a máscara binária para avaliação e desempenho técnico de lesões e do disco óptico como pode ser observado na Figura 5. Este conjunto de imagens consiste em 81 imagens de fundo de cor com sinais de RD e 164 sem sinais de RD (PORWAL P.; MERIAUDEAU, 2018), porem apenas 54 imagens possuem marcação do disco óptico feita por especialistas e elas possuem dimensões de 4288x2848 *pixels*.

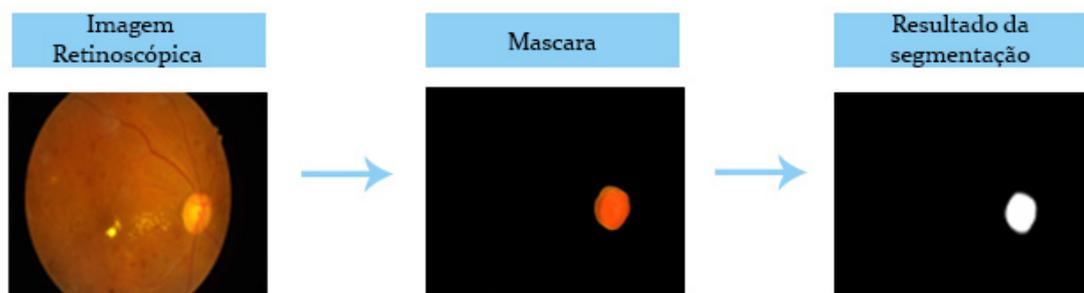
Figura 6: Representação do resultado da segmentação com a base *I-DRID*.



4.1.2 RIM-ONE

A *RIM-ONE* é uma base publica que possui um conjunto de 50 imagens biomédicas de fundo do olho com suas referentes mascaras binarias (FUMERO et al., 2011). As imagens possuem dimensões de 600x400 *pixels*.

Figura 7: Representação do resultado da segmentação com a base *RIM-ONE*.

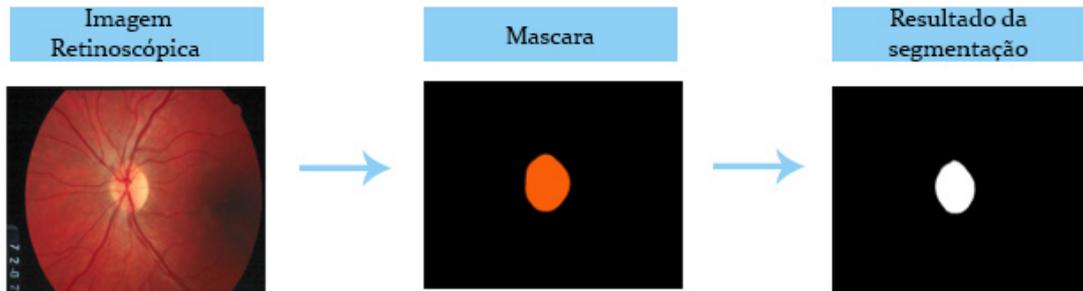


4.1.3 DRIONS-DB

DRIONS-DB possui em seus dados um conjunto de 37 imagens e suas respectivas mascaras (FEIJOO(A) JOSÉ MARÍA MARTÍNEZ DE LA CASA(A), 2008). As imagens

possuem dimensões de 530x462 pixels

Figura 8: Representação do resultado da segmentação com a base *DRIONS-DB*.



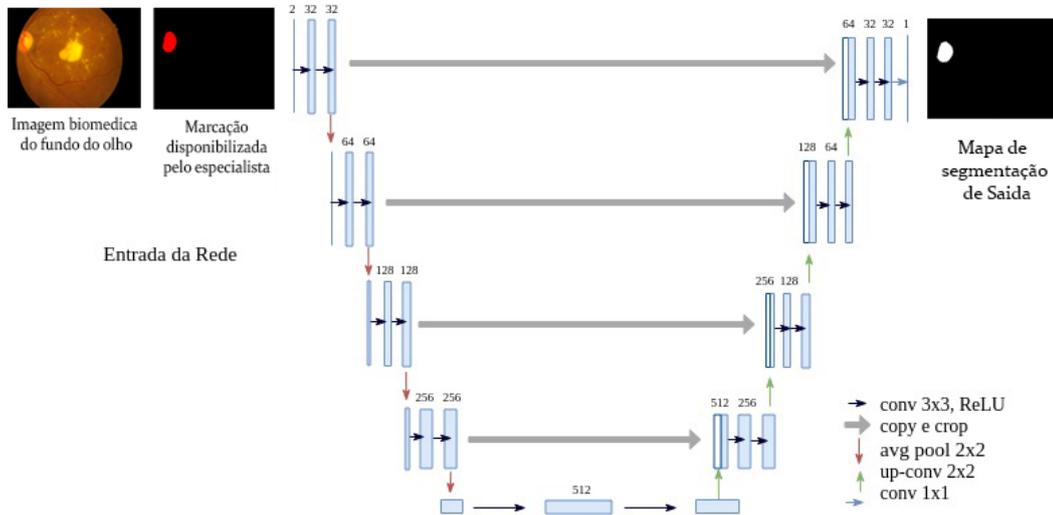
4.2 Segmentação

A segmentação do Disco Óptico é necessária pois em alguns estágios a doença pode apresentar uma lesão que se assemelha com o mesmo dificultando assim o diagnóstico. Esse trabalho apresenta um método para segmentação do DO, utilizando uma rede convolucional baseada na arquitetura U-Net. O método possui a entrada composta pela imagem de RD e pelas marcações dos especialistas dos exames de treino.

4.2.1 U-Net

A estratégia utilizada para a segmentação é um modelo de aprendizagem profunda baseado na U-Net proposta por [Ronneberger, Fischer e Brox \(2015\)](#) descrita na Subseção 2.3. A segmentação do disco óptico é realizada em cada pixel da imagem, sendo assim uma segmentação 2D. A entrada da rede consiste em uma imagem do fundo do olho acompanhada pela marcação específica do especialista.

Figura 9: U-Net arquitetura. Cada caixa azul corresponde a um mapa de recursos multicanal. O número de canais é indicado na parte superior da caixa. O tamanho X-Y é fornecido no canto inferior esquerdo da caixa, caixas brancas representam mapas de recursos copiados. As setas denotam as diferentes operações (RONNEBERGER PHILIPP FISCHER, 2015).



A rede utilizada para segmentação difere da U-Net original na quantidade de camadas. No método proposto a rede possui apenas quatro camadas de subamostragem com função de ativação do tipo média, onde a média de intensidade dos *pixels* da janela é mantida. Para o caminho de expansão é utilizada a convolução transposta. Cada convolução é seguida por uma camada de normalização de *batch* e uma camada de *Dropout*, com uma taxa de 20% para desativação de neurônios. Por fim, a segmentação é feita por meio de uma convolução (1 x 1), com função de ativação sigmóide. A rede é treinada com a função de perda *Dice Loss*, representada pela [Equação 4.1](#).

$$S = 1 - \frac{2VP}{2VP + FP + FN}. \quad (4.1)$$

4.3 Métricas de desempenho

As métricas utilizadas para validação dos resultados foram acurácia (PROVOST; DOMINGOS, 2000) e o dice (GINSBERG; YOUNG, 1992)

O índice Dice é bastante aplicado na literatura como medida para desempenho da segmentação e indica a proporção de *pixels* compartilhadas entre duas amostras em relação ao total de *pixels*.

Já acurácia, calcula o total de acertos baseado em todas as instâncias segmentadas, calculado a partir de todos os casos (segmentados corretamente e incorretamente) (PROVOST; DOMINGOS, 2000).

A equação Dice (DSC) é dada pela dobro de verdadeiros positivos (VP) comuns em duas amostras dividido pela soma do número total de amostra somado aos falsos positivos (FP) e falsos negativos (FN), Equação [Equação 4.2](#), onde a faixa de valores compreendidos

por esta medida vão de 0% (nenhuma correspondência entre os volumes) a 100% (total correspondência entre os volumes).

$$DSC = \frac{2VP}{2VP + FP + FN}. \quad (4.2)$$

A equação da Acurácia (A) é expressa através da razão do acerto total (dado por VP + VN) pelo total de casos da base. A equação para calculo da acurácia pode ser observado na [Equação 4.3](#).

$$A = \frac{VP + VN}{VP + VN + FP + FN}. \quad (4.3)$$

4.3.1 Validação

Para que seja realizada a análise dos resultados obtidos pelo presente trabalho, foram definidas algumas métricas para a validação da eficiência do mesmo. Este trabalho utiliza métricas comumente empregadas na literatura para análise de desempenho de sistemas baseados em processamento de imagens e reconhecimento de padrões para segmentação de imagens. Estas métricas são Acurácia (A), Dice ([Seção 4.3](#)). Tais métricas têm como objetivo medir o desempenho do modelo proposto como aceitável ou não, além de auxiliar na identificação de pontos positivos e negativos para melhoria futura deste trabalho nas fases de treinamento, validação e teste.

5 Resultados e Discussões

Este capítulo apresenta e discute os resultados obtidos pela metodologia proposta para segmentação do disco óptico em imagens retinoscópicas. Todos os testes foram implementados na linguagem de programação Python com o auxílio da biblioteca de aprendizagem profunda Keras (CHOLLET et al., 2018),

Para a realização destes experimentos, foi utilizada a plataforma Google Colaboratory (CARNEIRO et al., 2018). A mesma trata-se de um serviço baseado em *cloud* que utiliza Jupyter Notebooks para disseminar a pesquisa em aprendizado de máquina. No Colab, como também é conhecido, é fornecido um ambiente totalmente configurado e pronto para ser utilizado, com acesso gratuito a uma GPU robusta.

Para os testes da metodologia foram utilizadas todas as 54 imagens da base *I-DRID*, 50 imagens da base *RIM-ONE* e 37 imagens da base *DRIONS-DB*. As bases foram divididas aleatoriamente em 3 partes: treinamento, validação e teste. Da base *I-DRID* foram utilizadas 37 imagens para treino, 8 para validação e 9 para teste, a base *RIM-ONE* foram utilizadas 35 imagens para treino, 7 para validação e 8 para teste e a base *DRIONS-DB* foram utilizadas 24 imagens para treino, 6 para validação e 7 para teste.

Inicialmente foi realizada a quantização não-uniforme dos exames da base. Onde os valores de intensidade dos *voxels* são normalizados entre 0 e 255, tornando a região do disco óptico mais uniforme e realizando uma representação da imagem com uma quantidade menor de bits.

Logo após, se inicia a etapa de segmentação. Essa etapa foi executada separadamente para cada base e em seguida, executada juntando as bases de duas em duas em uma única base e fazendo variações de treino e teste.

Tabela 2: Resultados obtidos no método proposto com a U-Net treinada e testada apenas como uma única base.

Base de dados	Acurácia	Dice
I-DRID	99,52%	96,38%
RIM-ONE	98,90%	91,29%
DRIONS-DB	97,10%	81,53%

Diante dos resultados exibidos na Tabela 2 podemos observar resultados satisfatórios, na base *I-DRID* e na base *RIM-ONE* apresentam resultados semelhantes podemos atribuir isso porque as mesmas possuem quantidades semelhantes de imagens, já a base da *DRIONS-DB* apresenta pouca discrepância na acurácia porem o Dice apresenta resultados não tão apropriados, podemos atribuir isso ao menor numero de imagens existentes na base.

Tabela 3: Resultados obtidos no método proposto com a U-Net com variações na base de teste.

Base de treinamento	Base de teste	Acurácia	Dice
I-DRID	DRIONS-DB+ RIM-ONE	87,50%	83,01%
RIM-ONE	DRIONS-DB+ I-DRID	88,15%	85,70%
DRIONS-DB	I-DDRID+ RIM-ONE	85,40%	86,00%

Diante dos resultados exibidos na Tabela 3 podemos observar a variação na base de teste, onde as demais bases foram agrupadas de duas em duas. Com isso o conjunto de teste ficou bem maior que o de treino onde causou interferência nos resultados mostrando uma queda nos mesmos.

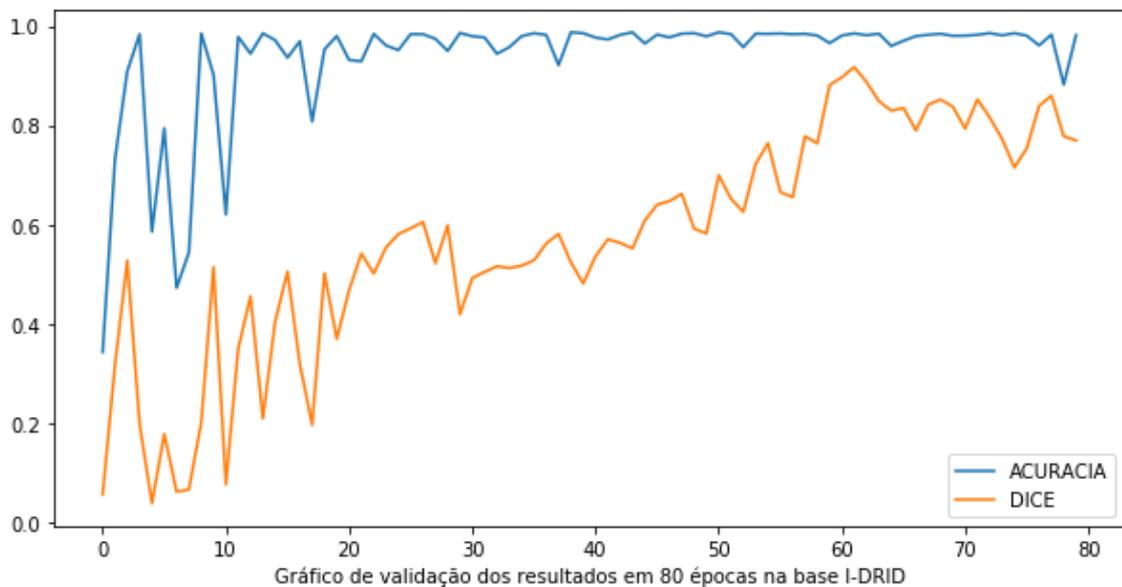
Tabela 4: Resultados obtidos no método proposto com a U-Net com variações na base de treino.

Base de treinamento	Base de teste	Acurácia	Dice
DRIONS-DB+ I-DRID	RIM-ONE	98,54%	97,23%
DRIONS-DB+ RIM-ONE	I-DRID	99,54%	95,39%
I-DRID+ RIM-ONE	DRIONS-DB	98,85%	97,43%

Com a variação exibida na Tabela 4 podemos observar que com um conjunto de treinamento maior interfere significativamente nos resultados da rede, mostrando uma melhora nos mesmos.

A Imagem 10 mostra a validação dos resultados na base *I-DRID* ao longo de 80 épocas executada na plataforma Google Colaboratory.

Figura 10: Gráfico do resultado da validação



A comparação com outros trabalhos dentro da literatura é uma tarefa um pouco complicada, já que há algumas diferenças nas maneiras de abordagem do problema e também nas ferramentas e técnicas utilizadas para solucioná-los. A Tabela 5 exibe um estudo comparativo entre o trabalho proposto e outros quatro da literatura exibindo o método utilizado, dataset e quantidade de amostras, onde podemos observar que os resultados obtidos pela metodologia proposta são semelhantes e por isso podemos considerá-los satisfatórios.

Tabela 5: Comparação do método proposto com alguns trabalhos relacionados.

Trabalho	Método	Dataset	Amostras	Sensibilidade	Acurácia	Dice
(XUE SHUO YAN, 2019)	Segmentação Automática	IDRID	81	98,1%	-	96,7%
(WEN-BO DUNWEI WEN, 2018)	CNN	DRIONS	100	-	98,90%	-
(KOWSALYA A. KALYANI, 2018)	Kapur e K-means	RIM-ONE	169	96,73%	-	93,93%
(XUE JIA-WEN LIN, 2017)	Saliência baseada em agrupamento	MESSIDOR	1200	-	94,00%	-
Método proposto	U-Net	IDRID	54	-	99,52%	96,38%

6 Conclusão e Trabalhos Futuros

A retinopatia diabética segundo a OMS (Organização Mundial de Saúde) está entre as cinco doenças que mais causam cegueira. A detecção precoce da retinopatia é uma estratégia para encontrar microaneurismas, assim, possibilitar melhor chance de tratamento. Para uma detecção precisa, é necessário imagens retinoscópicas do fundo do olho. Isso evidencia a importância do desenvolvimento de pesquisas com o objetivo de segmentar automaticamente o disco óptico, pois a doença em estágio avançado pode apresentar lesão semelhante ao disco óptico dificultando assim a segmentação do mesmo.

Desta forma, este trabalho propõe um método para segmentação do disco óptico em exames de RD, baseados em técnicas de aprendizado profundo. O método utiliza uma rede completamente convolucional com arquitetura baseada na U-Net.

Os resultados obtidos através de experimentos utilizando as bases de imagens públicas de exames de RD do olho, *I-DRID*, *RIM-ONE* e *DRION-DB*, mostram quão promissor é o método proposto nesse trabalho. O método produz resultados com o Dice acima de 85% utilizando as três bases, comparáveis aos trabalhos existentes na literatura.

O objetivo da metodologia não é substituir a função do especialista, e sim oferecer a ele uma segunda opinião, corroborando para o aumento da produtividade. Dessa forma, tornando mais ágil e menos exaustiva a análise de exames pelos especialistas. Os resultados apresentados mostraram-se bastantes promissores por apresentarem valores satisfatórios para as métricas de avaliação.

Como trabalhos futuros, destacamos:

- Realizar novos testes com uma quantidade maior de imagens / outros *datasets*;
- Realizar o ajuste fino para obter melhores resultados;
- Realizar testes com a segmentação de microaneurismas.

Referências

- ALMOUSLI, H. *Recognition of Facial Expressions with Autoencoders and Convolutional-Nets*. Tese (Doutorado) — Université de Montréal, 2014. Disponível em: <<http://hdl.handle.net/1866/10688>>. Citado na página 16.
- B. BEAULIEU WESLEY T., G. A. R. G. J. G. M. M. C. E. P. M. R. J. L. M. B. S. Photocoagulation versus ranibizumab for proliferative diabetic retinopathy should baseline characteristics affect choice of treatment? In: *for the Diabetic Retinopathy Clinical Research Network*. [S.l.]: Publishing Press, 2019. p. 1646–1654. Citado na página 13.
- CARNEIRO, T. et al. Performance analysis of google colab as a tool for accelerating deep learning applications. *IEEE Access*, IEEE, v. 6, p. 61677–61685, 2018. Citado na página 28.
- CHOLLET, F. et al. Keras: The python deep learning library. *Astrophysics Source Code Library*, 2018. Citado na página 28.
- FEIJOO(A) JOSÉ MARÍA MARTÍNEZ DE LA CASA(A), H. M. S. M. R. Z. M. B. M. E. J. C. S. J. G. Drions-db: Digital retinal images for optic nerve segmentation database. In: . [S.l.: s.n.], 2008. Citado na página 24.
- FERNANDES, B. J. T. Redes neurais com extração implícita de características para reconhecimento de padrões visuais. Universidade Federal de Pernambuco, 2013. Citado na página 16.
- FONG MD, M. L. A. M. P. T. W. G. M. G. L. K. M. G. B. D. S. Retinopathy in diabetes. 2016. Citado na página 16.
- FUMERO, F. et al. Rim-one: An open retinal image database for optic nerve evaluation. In: . [S.l.: s.n.], 2011. p. 1 – 6. Citado na página 24.
- GINSBERG, J. R.; YOUNG, T. P. Measuring association between individuals or groups in behavioural studies. *Animal Behaviour*, Elsevier Science, v. 44, n. 1, p. 377–379, 1992. Citado 2 vezes nas páginas 13 e 26.
- HAFEMANN, L. G. An analysis of deep neural networks for texture classification. 2014. Citado 2 vezes nas páginas 17 e 18.
- HAFEMANN, L. G. An analysis of deep neural networks for texture classification. 2014. Citado na página 18.
- IEEE. Ieee international symposium on biomedical imaging. *IEEE Pulse*, 2018. Citado na página 24.
- KANG, K.; WANG, X. Fully convolutional neural networks for crowd segmentation. *arXiv preprint arXiv:1411.4464*, 2014. Citado na página 16.

- KOWSALYA A. KALYANI, C. J. C. R. S. M. J. V. R. N. An approach to extract optic-disc from retinal image using k-means clustering. In: *4th International Conference on Biosignals, Images and Instrumentation (ICBSII)*. [S.l.]: Publishing Press, 2018. Citado 3 vezes nas páginas 21, 22 e 30.
- LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. *Nature*, v. 521, n. 7553, p. 436–444, 2015. ISSN 0028-0836. Citado na página 16.
- LECUN, Y. et al. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, IEEE, v. 86, n. 11, p. 2278–2324, 1998. Citado na página 17.
- LECUN, Y. et al. Convolutional networks and applications in vision. In: *ISCAS*. [S.l.: s.n.], 2010. p. 253–256. Citado na página 17.
- LONG, J.; SHELHAMER, E.; DARRELL, T. Fully convolutional networks for semantic segmentation. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. [S.l.: s.n.], 2015. p. 3431–3440. Citado na página 18.
- Long, J.; Shelhamer, E.; Darrell, T. Fully convolutional networks for semantic segmentation. In: *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. [S.l.: s.n.], 2015. p. 3431–3440. ISSN 1063-6919. Citado na página 19.
- OLIVEIRA RENAN MAGALHÃES MONTENEGRO JUNIOR, S. V. José Egídio Paulo de. Retinopatia diabética. *Sociedade Brasileira de Diabetes: SBD*, 2018. Citado na página 15.
- PERONA, P.; FINK, M. *The Full Images for Natural Knowledge Caltech Office DB*. California, 2013. Citado na página 16.
- PHD, G. C. M. C. F. D. S. W. T. M.; FRANZCO, T. Y. W. P. Diabetic retinopathy: global prevalence, major risk factors, screening practices and public health challenges: a review. *Clinical and Experimental Ophthalmology*, 2016. Citado na página 15.
- PORWAL P., P. S. K. R. K. M. D. G. S. V.; MERIAUDEAU, F. Indian diabetic retinopathy image dataset (idrid): a database for diabetic retinopathy screening research. In: *Journal Data*. [S.l.]: Publishing Press, 2018. Citado na página 24.
- PROVOST, F.; DOMINGOS, P. Well-trained pets: Improving probability estimation trees. Citeseer, 2000. Citado 2 vezes nas páginas 13 e 26.
- RONNEBERGER, O.; FISCHER, P.; BROX, T. U-net: Convolutional networks for biomedical image segmentation. In: SPRINGER. *International Conference on Medical image computing and computer-assisted intervention*. [S.l.], 2015. p. 234–241. Citado 3 vezes nas páginas 19, 20 e 25.
- RONNEBERGER PHILIPP FISCHER, T. B. O. U-net: Convolutional networks for biomedical image segmentation. In: *Spring International Publishing Switzerland*. [S.l.]: Publishing Press, 2015. Citado 2 vezes nas páginas 8 e 26.
- SCANLON, P. H. Diabetic retinopathy. In: *Elsevier Ltd*. [S.l.: s.n.], 2019. Citado na página 13.

- SILVA, G. L. F. d. et al. Diagnóstico de nódulos pulmonares em imagens de tomografia computadorizada usando redes neurais convolucionais evolutivas. Universidade Federal do Maranhão, 2017. Citado na página 18.
- WANG, L. et al. Visual tracking with fully convolutional networks. In: *The IEEE International Conference on Computer Vision (ICCV)*. [S.l.: s.n.], 2015. Citado na página 19.
- WEN-BO DUNWEI WEN, A. A. D. Y. Y. W. K. H. Automatic detection of optic disc in retina image using cnn and crf. In: *IEE SmartWorld*. [S.l.]: Publishing Press, 2018. Citado 3 vezes nas páginas 21, 22 e 30.
- XUE JIA-WEN LIN, X.-R. C. S.-H. Z. L. Y. L.-Y. Optic disk detection and segmentation for retinal images using saliency model based on clustering. In: *Journal of Computers Vol. 29 N° 5*. [S.l.]: Publishing Press, 2017. p. 66–79. Citado 3 vezes nas páginas 21, 22 e 30.
- XUE SHUO YAN, J. Q. F. Q. C. Q.-M. C. T. L. D. L. X. L. J. Deep membrane systems for multitask segmentation in diabetic retinopathy. In: *Elsevier B.V.* [S.l.]: Publishing Press, 2019. Citado 4 vezes nas páginas 13, 21, 22 e 30.



**TERMO DE AUTORIZAÇÃO PARA PUBLICAÇÃO DIGITAL NA BIBLIOTECA
“JOSÉ ALBANO DE MACEDO”**

Identificação do Tipo de Documento

- () Tese
() Dissertação
(x) Monografia
() Artigo

Eu, Waldy Alves da Silva,
autorizo com base na Lei Federal nº 9.610 de 19 de Fevereiro de 1998 e na Lei nº 10.973 de
02 de dezembro de 2004, a biblioteca da Universidade Federal do Piauí a divulgar,
gratuitamente, sem ressarcimento de direitos autorais, o texto integral da publicação

Segmentação do disco óptico em imagens retinográficas
utilizando a rede neural U-Net

de minha autoria, em formato PDF, para fins de leitura e/ou impressão, pela internet a título
de divulgação da produção científica gerada pela Universidade.

Picos-PI _____ de _____ de 20 _____

Waldy Alves da Silva
Assinatura

Assinatura