

Rafael Araújo Ferreira
Orientador: Flávio Henrique Duarte de Araújo
Co-orientadora: Alcilene Dalília de Sousa

Detecção e Segmentação do Disco Óptico em Imagens da Retina

Picos - PI
Novembro de 2018

Rafael Araújo Ferreira
Orientador: Flávio Henrique Duarte de Araújo
Co-orientadora: Alcilene Dalília de Sousa

Detecção e Segmentação do Disco Óptico em Imagens da Retina

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
Novembro de 2018

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

F383d Ferreira, Rafael Araújo.

Detecção e segmentação do disco óptico em imagens da retina / Rafael Araújo Ferreira.– 2018.

CD-ROM : il.; 4 ¾ pol. (31 f.)

Trabalho de Conclusão de Curso (Curso Bacharelado em Sistemas de Informação) – Universidade Federal do Piauí, Picos, 2018.

Orientador(A): Orientador: Flávio Henrique Duarte de Araújo

Co-orientadora: Alcilene Dalília de Sousa

1. Aprendizado Profundo. 2. Glaucoma Imagens Médicas. 3. Processamento de Imagens. 4. Redes Totalmente Convolucionais. I. Título.

CDD 006.6

DETECÇÃO E SEGMENTAÇÃO DO DISCO ÓPTICO EM IMAGENS DA RETINA

RAFAEL ARAÚJO FERREIRA

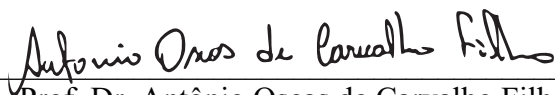
Monografia APROVADA como exigência parcial para obtenção do grau de
Bacharel em Sistemas de Informação.

Data de Aprovação

Picos – PI, 26 de Novembro de 2018



Prof. Dr. Flávio Henrique Duarte de Araújo
Orientador



Prof. Dr. Antônio Oseas de Carvalho Filho
Membro



Prof. Dr. Romuete Rodrigues Veloso e Silva
Membro

Agradecimentos

"Por que quem compreendeu a mente do Senhor? [...] Porque d'Ele e por Ele, e para Ele, são todas as coisas; glória, pois, a Ele eternamente. Amém. - Romanos 11:34 a, 36"

Aos meus pais, Moisés e Jucileide, também à minha irmã Rute, minhas colunas de sustentação durante toda a minha vida. À minha namorada Karen, minha inspiração pra ser o melhor que eu puder. A todos os meus amigos de curso que levarei comigo enquanto viver, em especial aos queridos Kayke, Dieguinho, Kirito, Fabrício, Guilherme, Jota P, Kécyo, Leo Preto, Leo de Jesus e Mateus PGx, companheiros da família que encontrei e que persiste até aqui. À minha prima Naara e sua família, suporte fundamental que encontrei nessa cidade, bem como seus amigos de turma Daví, Douglas, Estevão, Mateus (O Mateuzão), Milton e tantos outros, que fizeram parte de uma corrente de alunos que tomou pra si as dores dos demais, dando tudo o que podiam para ajudar uns aos outros. Aos irmãos da Segunda Igreja Batista de Picos, em especial ao Pastor Neyron, Sua esposa Juciane e meus amigos Nathan e José Victor, que me acolheram e cuidaram de mim em todo este período. À todos vocês, muitíssimo obrigado.

Um singelo e especial "obrigado!" vai para o professor Flávio Henrique, que não apenas me aceitou de última hora como seu orientando, como suportou e fez cobranças e incentivos sem os quais eu não teria chegado aqui.

Não foi uma jornada fácil. Os problemas surgem, pessoas criticam e julgam. Todos dizem ter melhor entendimento sobre o que fazer com sua vida. Tudo parece urgente e todos esperam de você uma maturidade e capacidade de decisão que talvez ainda não tenha. Sobre tudo isto, paciência e perseverança. Seguir em frente, independente de qualquer situação. Entender que a vida é mais do que isto aqui.

"Não é de chegar primeiro, mas é de continuar" Paulo Nazareth

Resumo

É notável o crescimento do número de casos tardiamente diagnosticados de glaucoma que ocasionam cegueira em uma boa parcela dos indivíduos afetados, bem como os esforços da comunidade de tecnologia da informação em desenvolver métodos de diagnóstico automático desta patologia. Estes métodos requerem, na maior parte dos casos, ferramentas que auxiliem na preparação das imagens que serão utilizadas como base em seu desenvolvimento. Por esta razão se propôs a criação de um sistema de segmentação automática do Disco Óptico em imagens médicas de fundo de olho, capaz de preparar imagens para serem utilizadas em sistemas de diagnóstico automático. No decorrer deste trabalho se demonstra a viabilidade de adoção das Redes Totalmente Convolucionais como ferramenta para segmentação automática de imagens médicas de fundo de olho. Tal demonstração é feita através do treinamento e teste de quatro modelos de arquitetura de rede em condições semelhantes. A base de dados utilizada é constituída de 400 imagens de fundo de olho em alta qualidade, disponibilizadas publicamente com suas respectivas máscaras e informações que constituem o *ground truth* (verdade-terrestre). Foi demonstrado que dentro das condições utilizadas durante o estudo, as Redes Totalmente Convolucionais são uma solução viável para a segmentação automática do Disco Óptico, tendo seus melhores resultados quando utilizando a arquitetura VGG. Essa arquitetura alcançou uma acurácia de pixel de 88%, um bom resultado quando comparado a trabalhos semelhantes encontrados na literatura.

Palavras-chaves: Aprendizado Profundo, Glaucoma, Imagens Médicas, Processamento de Imagens, Redes Totalmente Convolucionais.

Abstract

It is notable the growth in the number of late diagnosed cases of glaucoma that cause blindness in a good portion of the affected individuals, as well as the efforts of the information technology community in developing methods of automatic diagnosis of this pathology. These methods require, in most cases, tools that assist in the preparation of images that will be used as a basis for their development. For this reason it was proposed the creation of an automatic optical disk segmentation system in medical background images, capable of preparing images for use in automatic diagnostic systems. In the course of this work, it is demonstrated the feasibility of adopting the Fully Convolutional Networks as a tool for automatic segmentation of medical images of the eye fundus. Such demonstration is done by training and testing four network architecture models under similar conditions. The database used consists of 400 high quality eye fund images, publicly available with their respective masks and information that constitute the ground truth. It was demonstrated that under the conditions used during the study, the Fully Convolutional Networks are a viable solution for the automatic segmentation of the optical disk, having its best results when using the VGG architecture. This architecture achieved a pixel accuracy of 88 %, a good result when compared to similar works found in the literature.

Lista de ilustrações

Figura 1 – Escavação em Olho Saudável (a) e Glaucomatoso (b).	14
Figura 2 – Ilustração de uma Rede Neural Artificial.	17
Figura 3 – Ilustração de uma Rede de Aprendizado Profundo.	18
Figura 4 – Ilustração do Funcionamento de uma Rede Neural Convolutacional. . . .	19
Figura 5 – Ilustração do Funcionamento de uma Rede Totalmente Convolutacional.	20
Figura 6 – Demonstrativo das Imagens Contidas na Base de Dados Utilizada. . . .	21
Figura 7 – Ilustração da Arquitetura ResNet.	25
Figura 8 – Ilustração da Arquitetura do modelo VGG.	25
Figura 9 – Ilustração da Arquitetura do modelo Segnet.	26
Figura 10 – Demonstração das Segmentações Feitas por Cada Modelo.	28
Figura 11 – Demonstração de Falhas na Segmentação.	28

Lista de tabelas

Tabela 1 – Descrição do Experimento.	26
Tabela 2 – Descrição do Ambiente de Desenvolvimento.	27
Tabela 3 – Resultados Obtidos.	27
Tabela 4 – Comparativo Entre os Resultados Obtidos.	29

Lista de abreviaturas e siglas

CNO	Cabeça do Nervo Óptico
DL	<i>Deep Learning</i> (Aprendizagem Profunda)
DO	Disco Óptico
PDI	Processamento Digital de Imagens
RNA	Redes Neurais Artificiais
RNC	Redes Neurais Convolucionais (<i>Convolutional Neural Networks</i>)
RTC	Redes Totalmente Convolucionais (<i>Fully Convolutional Networks</i>)
TCO	Tomografia de Coerência Óptica

Lista de símbolos

Σ Somatório

Sumário

1	Introdução	13
1.1	Contexto e Problema	13
1.2	Objetivos	14
1.2.1	Objetivo Geral	14
1.2.2	Objetivos Específicos	15
1.3	Contribuição	15
1.4	Metodologia de Trabalho	15
1.5	Disposição do Trabalho	15
2	Referencial Teórico	17
2.1	Redes Neurais Artificiais	17
2.2	Aprendizado Profundo	18
2.3	Redes Neurais Convolucionais	19
2.4	Redes Totalmente Convolucionais	19
2.5	Base de Imagens	20
2.5.1	RIM-ONE	20
3	Trabalhos Relacionados	22
4	Método Proposto	24
4.1	Modelos de Rede	24
4.2	Treinamento e Avaliação	25
4.3	Resultados e Discussões	27
5	Conclusão	30
5.1	Trabalhos Futuros	30
	Referências	31

1 Introdução

O termo Glaucoma é atualmente associado à uma série de neuropatias ópticas progressivas que tem em comum uma lenta e crescente degeneração das células retinais ganglionares e seus axônios, o que resulta em uma aparência peculiar do Disco Óptico (DO) e um simultâneo padrão de perda visual (WEINREB; KHAW, 2004). Dada a viabilidade de diagnóstico através da avaliação da aparência do DO (BUSSEL; WOLLSTEN; SCHUMAN, 2013), abre-se um leque de possibilidades a respeito das formas plausíveis de auxílio tecnológico para um diagnóstico comumente realizado apenas por profissionais humanos com o devido gabarito.

1.1 Contexto e Problema

Com origens que remontam ao período da Grécia antiga, o termo glaucoma (derivado do grego *glaukos*) era utilizado para identificar um grupo de distúrbios oculares que levavam à cegueira, normalmente indicados por características como coloração (esverdeada ou esbranquiçada, como se observa na cegueira) ou o próprio tamanho do globo ocular, normalmente alterado pelo aumento da pressão intraocular (sintoma frequentemente observado em diversos quadros de disfunção ótica). Não havia, porém, pertinência a nenhuma patologia em especial (MONTEIRO, 2014).

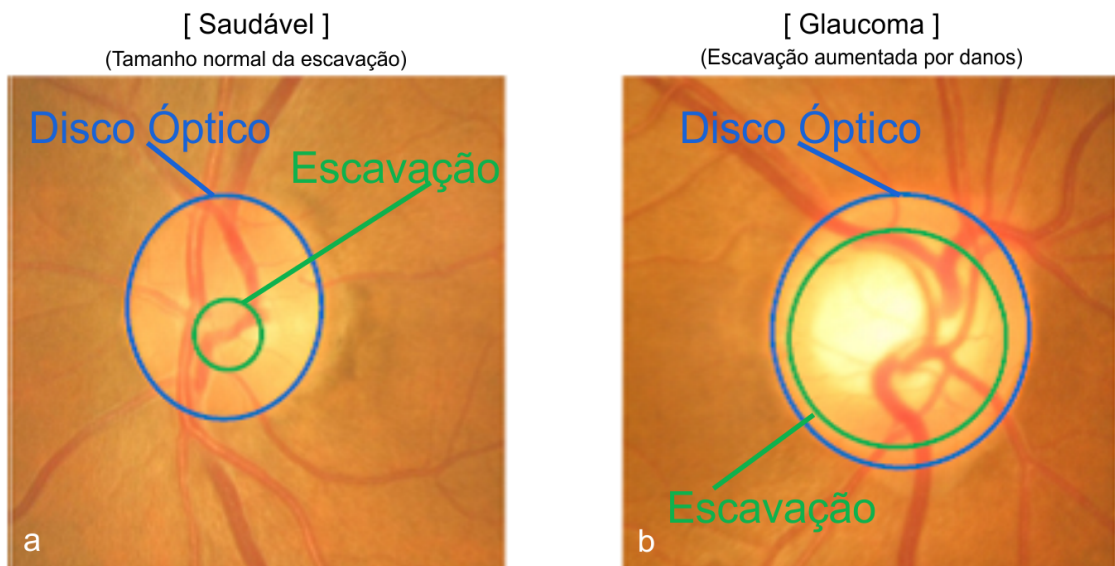
Atualmente o Glaucoma é associado a um conjunto de neuropatias ópticas lentas e progressivas que resultam em um padrão de perda visual, cuja progressão é frequentemente observada concomitantemente com uma peculiar alteração da aparência do Disco Óptico (WEINREB; KHAW, 2004). Por se tratar de uma patologia crônica degenerativa, o diagnóstico durante os estágios iniciais é de evidente importância e pode ser realizado por meio de exames clínicos como a tonometria, que determina a pressão intraocular (LAURETTI; FILHO, 1997), ou ainda através de métodos de imagem, como por exemplo a Tomografia de Coerência Óptica (TCO), um exame não invasivo que permite a obtenção de imagens da Cabeça do Nervo Óptico (CNO) e da retina, comumente utilizado para a avaliação de danos estruturais glaucomatosos (BUSSEL; WOLLSTEN; SCHUMAN, 2013).

DO, principal área do olho humano afetada pela patologia em estudo, é uma denominação clínica criada para descrever as estruturas que envolvem o tecido neural posicionado internamente à cabeça do nervo óptico (REIS ALEXANDRE SOARES CASTRO; BURGOS; CHAUHAN, 2012).

A escavação patológica do Disco Óptico está frequentemente associada ao glaucoma e outras condições neuro-oftálmicas. - (ZHANG YI-XIN; WEI, 2014)

Com isso é possível assegurar a existência de características físicas observáveis que podem indicar fortemente a presença do Glaucoma ou condições semelhantes. A Figura 1 ilustra a marcação das áreas de interesse em imagem de olho saudável (a) e de olho glaucomatoso (b) sendo que o círculo azul sinaliza o DO enquanto o verde a Escavação.

Figura 1: Escavação em Olho Saudável (a) e Glaucomatoso (b).



Fonte: Adaptação de ETH Zurich Department of Computer Science. ¹

Uma vez observada a viabilidade de diagnóstico do Glaucoma através da avaliação da aparência do Disco Óptico, decidimos iniciar a criação de um sistema que segmentasse imagens médicas de fundo de olho. Isto deve resultar num aumento de praticidade na preparação de imagens para sistemas de diagnóstico automático.

1.2 Objetivos

Nesta seção são abordados mais profundamente os objetivos geral e específicos deste trabalho.

1.2.1 Objetivo Geral

Desenvolver um método de segmentação do Disco Óptico em imagens de fundo de olho, que seja robusto e performe tão bem quanto, ou melhor que os disponíveis na literatura.

¹ Disponível em <<http://www.ise.inf.ethz.ch/education/studentprojects/completedprojects/automatic-optic-cup-and-optic-disc-segmentation-for-glaucoma-dia0.html>> - Acesso em 10/11/18 11:34h

1.2.2 Objetivos Específicos

1. Aplicar Redes Totalmente Convolucionais (RTC's) para segmentação do DO como área de interesse.
2. Determinar a melhor arquitetura de RTC para cumprir este objetivo.

1.3 Contribuição

A principal contribuição deste trabalho é a avaliação das Redes Neurais Totalmente Convolucionais na segmentação semântica (segundo o contorno real do DO) de imagens médicas de fundo de olho.

1.4 Metodologia de Trabalho

A investigação da literatura permitiu identificar as técnicas baseadas em Aprendizado Profundo (AP) mais utilizadas, tanto para fins gerais quanto especificamente para tratamento de imagens médicas. Consequentemente, após um levantamento das características de cada uma das técnicas utilizadas, foi possível determinar que as Redes Totalmente Convolucionais seriam primordialmente a técnica mais adequada para este trabalho, tendo sido utilizada na detecção (ABDULLA, 2017) e segmentação semântica de objetos (SHELHAMER; LONG; DARRELL, 2016), mas não para segmentação com foco na preparação de imagens para classificadores de imagens médicas e ferramentas de diagnóstico automatizado. Com isto, seguiram-se as etapas de desenvolvimento, testes laboratoriais, correções de código, e consequente treinamento dessa rede para adequação de pesos e posterior obtenção de resultados, que foram comparados às segmentações realizadas por profissionais da área médica, gerando os índices apresentados na Seção 4.3 deste trabalho.

1.5 Disposição do Trabalho

Nos capítulos que se seguem o trabalho obedece à seguinte disposição. No Capítulo 2 são apresentados conceitos fundamentais ao desenvolvimento e compreensão do trabalho, bem como a base de dados e técnicas utilizadas.

Em sequência, o Capítulo 3 contém brevemente descritos os principais trabalhos relacionados encontrados na literatura, bem como suas principais semelhanças e diferenças em relação a este trabalho.

Logo em seguida, no Capítulo 4, são abordadas as principais características deste trabalho, tais como ambiente de desenvolvimento, método desenvolvido, resultados alcançados e outras informações úteis sobre a solução aqui implementada.

Por fim, no Capítulo 5 são apresentadas as principais conclusões deste trabalho, bem como possíveis trabalhos futuros.

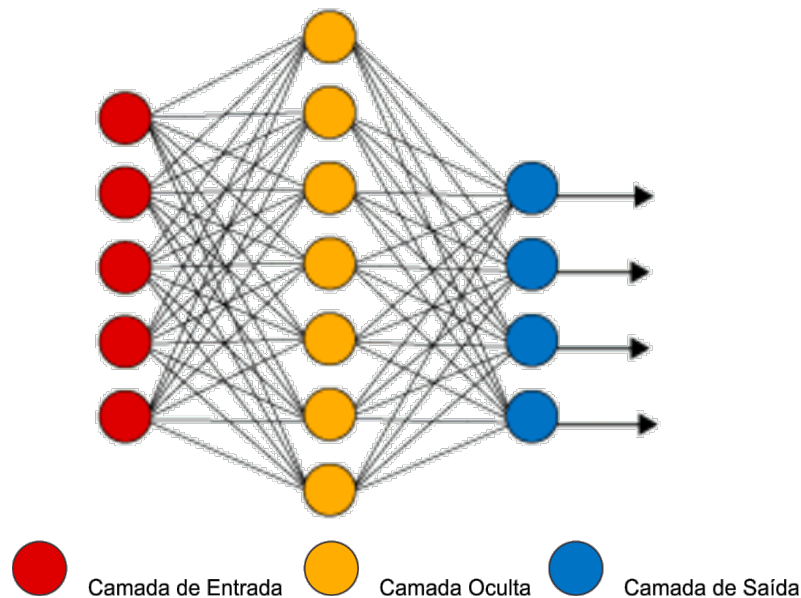
2 Referencial Teórico

Neste capítulo são apresentados os principais conceitos que fundamentaram a construção do método desenvolvido.

2.1 Redes Neurais Artificiais

Redes Neurais Artificiais (RNA) são uma classe de algoritmos de aprendizagem de máquina que são capazes de aprender à partir de coleções de dados e se especializar no reconhecimento de padrões, inspirados na estrutura e funcionamento do cérebro, onde células especializadas em tratar estímulos (neurônios) recebidos como impulsos elétricos se organizam em rede com o propósito de interpretar os estímulos recebidos e determinar quais atitudes devem ser tomadas. Semelhantemente, as redes neurais artificiais são constituídas de componentes lógicos (Neurônios Artificiais) que recebem entradas adaptadas (metadados) e “descobrem” quais as saídas adequadas para cada uma destas entradas (ROSEBROCK, 2017). A Figura 2 traz a ilustração de uma RNA.

Figura 2: Ilustração de uma Rede Neural Artificial.



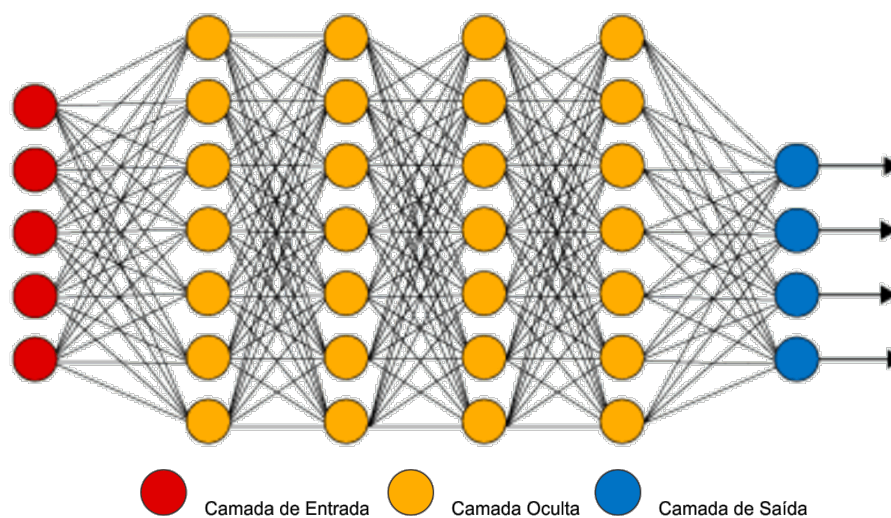
Fonte: Adaptação de *Deep Learning Book*.¹

¹ Disponível em <<http://deeplearningbook.com.br/o-que-sao-redes-neurais-artificiais-profundas/>> - acesso em 03/12/18.

2.2 Aprendizado Profundo

Deep Learning, em português, Aprendizado Profundo é um subcampo da aprendizagem de máquina que, por sua vez, é um subcampo da inteligência artificial, e assim como seu campo geral tem como objetivo central desenvolver técnicas e algoritmos que possam ser usados para resolver problemas com os quais os humanos estão habituados a lidar de forma intuitiva e praticamente automática (ROSEBROCK, 2017). A Figura 3 demonstra uma Rede de Aprendizado Profundo.

Figura 3: Ilustração de uma Rede de Aprendizado Profundo.



Fonte: Adaptação de *Deep Learning Book*.²

”Métodos de Aprendizado Profundo são métodos de aprendizagem representativa com múltiplos níveis de representação, obtidos ao se compor módulos simples, mas não lineares capazes de transformar a representação em um nível baixo (iniciando com a entrada bruta) em uma representação em nível mais alto e sutilmente mais abstrato” - (LECUN; BENGIO; HINTON, 2017, p. 436)

Existem inúmeras abordagens de Aprendizado Profundo derivadas de modificações das Redes Neurais Artificiais, criadas e aplicadas para os mais diversos fins. Porém, no que tange ao PDI, algumas abordagens se destacam; entre elas as Redes Neurais Convolucionais (RNC) e as Redes Totalmente Convolucionais (RTC). Nas seções à seguir serão apresentados os fundamentos necessários para a compreensão destas técnicas.

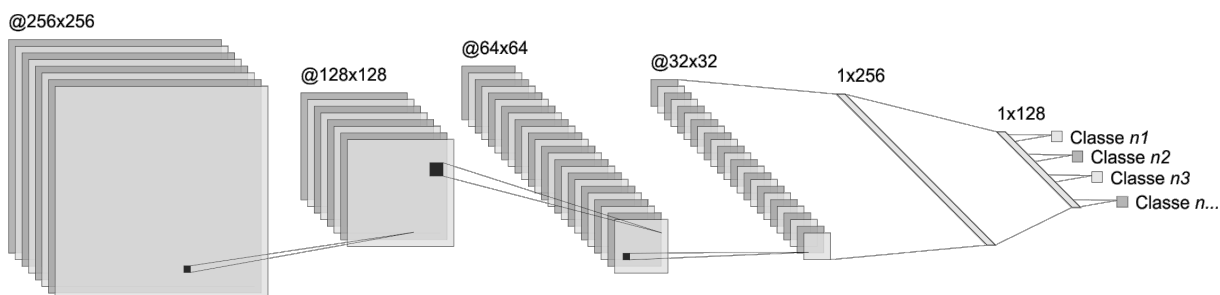
² Disponível em <<http://deeplearningbook.com.br/o-que-sao-redes-neurais-artificiais-profundas/>> - acesso em 03/12/18.

2.3 Redes Neurais Convolucionais

Análogas às Redes Neurais Artificiais tradicionais (RNA), Redes Neurais Convolucionais são compostas de neurônios artificiais que se auto otimizam através da aprendizagem. Cada neurônio recebe uma entrada e faz uso de uma função de ativação, comportamento fundamental de dezenas de Redes Neurais Artificiais. A única diferença notável entre RNC e RNA é o fato de as RNC serem primordialmente utilizadas no campo de reconhecimento de padrões em sinais, o que permite a codificação de características específicas das imagens na arquitetura, tornando-as mais recomendadas e utilizáveis para tarefas com foco em PDI (THEANO, 2013).

A grande vantagem da aplicação de RNC's para o Processamento Digital de Imagens é sua grande capacidade de extrair características, por trabalhar com convoluções e *pooling*, ou seja, criando “sub-quadros de bits” de forma que seja possível realçar contornos e áreas de interesse de cada imagem (ROSEBROCK, 2017), como ilustrado na Figura 4.

Figura 4: Ilustração do Funcionamento de uma Rede Neural Convolutional.



Fonte: Elaborado pelo autor.

O processo de subdivisão da imagem em subgrupos de *pixels* cada vez menores é denominado *pooling*, enquanto Convolução caracteriza pela extração de características destes subgrupos. Vale lembrar que o processo realizado em camadas de *pooling* está frequentemente associado a camadas de *padding* que auxiliam na diminuição do volume de dados, camadas de *dropout* que ignoram unidades aleatórias para evitar falso aprendizado, bem como camadas *upsampling*, que auxiliam na deconvolução.

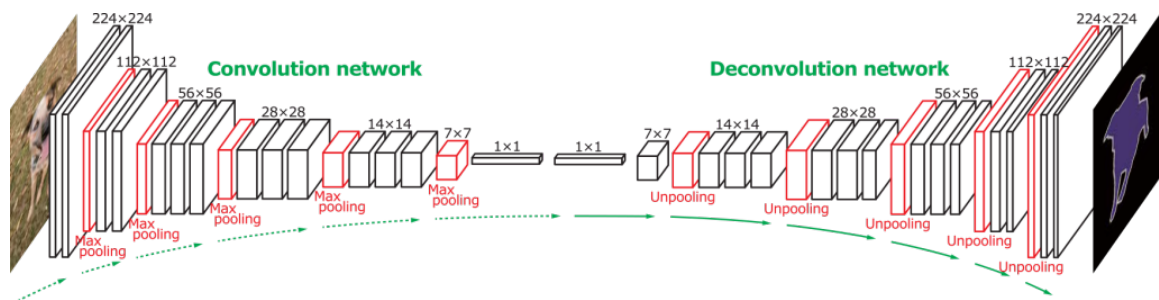
Como visto na Figura 4, em Redes Neurais Convolucionais as imagens de entrada são divididas em “janelas” que serão processadas e cujas características serão somadas e propagadas para as próximas camadas da rede neural. Contudo, considerando a necessidade de gerar imagens propriamente ditas como resultado do processamento, surgem as Redes Totalmente Convolucionais, apresentadas na seção 2.4.

2.4 Redes Totalmente Convolucionais

Derivadas de uma modificação das RNC's convencionais, as Redes Totalmente Convolucionais (RTC's) são de grande valia para sistemas de segmentação automática e de

tecção de objetos, tendo em vista que sua arquitetura substitui as últimas camadas de uma RNC comum por camadas deconvolucionais, criando assim um processo reverso e tornando janelas de bits maiores a cada camada, o que permite que a saída da rede seja uma imagem semelhante à entrada, porém, segmentada. Após o devido treinamento as RTC's são capazes de detectar objetos no contexto das imagens, como ilustrado na Figura 5. Por observar esta capacidade e a inexistência de trabalhos semelhantes, as RTC's foram escolhidas como objeto de estudo desse trabalho.

Figura 5: Ilustração do Funcionamento de uma Rede Totalmente Convolutiva.



Fonte: Azoft Digital Products.³

A próxima seção apresenta a base de imagens utilizada para o treinamento e avaliação dos algoritmos no presente trabalho.

2.5 Base de Imagens

Para o desenvolvimento de um sistema de Processamento Digital de Imagens é preciso constituir uma base de imagens confiável para que seja possível realizar o processo de treinamento, bem como avaliar corretamente os resultados obtidos pelo mesmo. Com isso em mente, apresenta-se durante esta seção e suas subseções, a base de imagens utilizada para treinamento e avaliação do método desenvolvido.

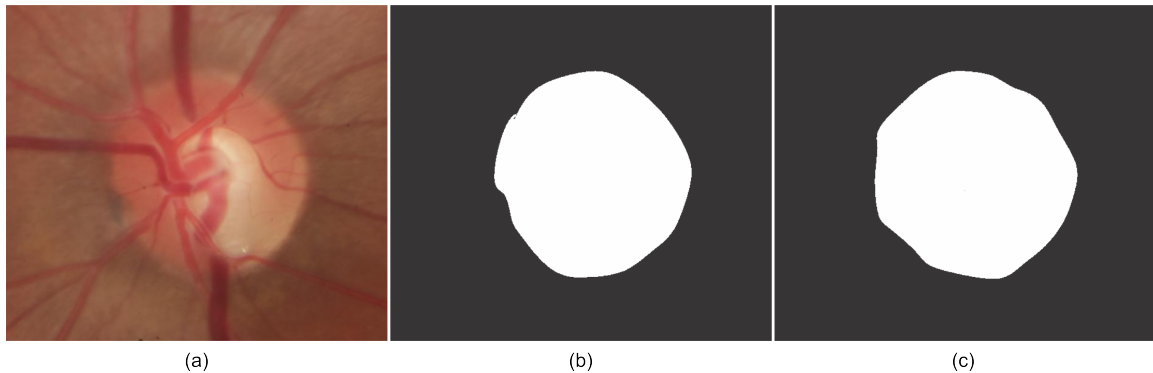
2.5.1 RIM-ONE

A biblioteca RIM-ONE é uma coleção pública de imagens de fundo de olho com elevados padrões de qualidade no que diz respeito à definição das fotografias e sua documentação. A mesma foi elaborada com o objetivo de prover segmentações acuradas de imagens da cabeça do Nervo Óptico que servissem de base para a concepção de novos sistemas de detecção automática do Glaucoma (FUMERO et al., 2011). O alto nível de qualidade da base é certificado pela presença em todas as suas versões de um arquivo de texto referente a cada uma das imagens. Este arquivo contém informações importante,

³ Disponível em <<https://www.azoft.com/blog/fully-convolutional-neural-networks/>> - acesso em 10/11/18.

como o centroide médio de todos os contornos feitos por profissionais, bem como eixos em 8 direções partindo do centroide até o contorno do Disco Óptico. A Figura 6 demonstra imagens desta base.

Figura 6: Demonstrativo das Imagens Contidas na Base de Dados Utilizada.



Fonte: RIM-ONE R1.

Observa-se na Figura 6 em (a) uma imagem de fundo de olho contida na primeira de três versões disponibilizadas da RIM-ONE. Em (b) e (c) observam-se segmentações manuais realizadas por profissionais que auxiliaram na construção da base. É importante ressaltar que cada uma das versões da biblioteca tem características únicas. Observe:

- RIM-ONE R1: Primeira versão da biblioteca, conta com um total de 158 imagens de 478 *pixels* de altura e 436 de largura, divididas e organizadas como: Glaucoma Profundo (14), Glaucoma Mediano (14), Glaucoma Leve (12), e Normais (118). Contém também cinco segmentações de profissionais para cada uma das imagens e um arquivo de metadados referentes a cada imagem.
- RIM-ONE R2: Segunda versão da biblioteca, constituída de um total de 455 imagens de 400 *pixels* de altura e 469 de largura, divididas e organizadas como Glaucoma e Suspeitos (200) e Normais (255). Contém também um arquivo de metadados referentes a cada imagem. Não contém, entretanto, imagens referentes à segmentação de profissionais, sendo o arquivo de metadados a única referência para *ground truth*, por esse motivo essas imagens não foram utilizadas neste trabalho.
- RIM-ONE R3: Terceira e mais recente versão da biblioteca, constituída de 159 imagens estéreo (olhos direito e esquerdo) de 2144 *pixels* de altura e 1424 de largura, divididas e organizadas como Glaucoma e Suspeitos (74) e Normais (85). Contém também duas imagens referentes à segmentações de profissionais para cada uma das imagens, bem como arquivos de metadados referentes a cada imagem.

As três versões da biblioteca são disponibilizadas gratuitamente no site do MIAG⁴, somando 1072 imagens de fundo de olho em alta qualidade.

⁴ Disponível em <<http://medimrg.webs.ull.es/>>. Acesso em 11/11/2018.

3 Trabalhos Relacionados

No que concerne aos trabalhos desta linha de pesquisa, pode-se observar uma tendência de crescimento no uso de técnicas de Aprendizado Profundo, tanto para segmentação de objetos em imagens comuns, quanto para segmentação de imagens médicas. A seguir são descritos os trabalhos relacionados com objetivos semelhantes a este.

1. [STAAL et al. \(2010\)](#) adotam a decomposição de alterações na aparência dos vasos sanguíneos baseado nas características impostas por cada condição clínica para detectar patologias óticas. Diferencia-se deste trabalho tanto por sua abordagem quanto pelas características observadas, não baseando o diagnóstico no aumento patológico da escavação do DO, mas na aparência dos vasos. Foi igualmente importante para este trabalho por trazer informações úteis a respeito da fisionomia do olho humano e características relevantes.
2. [RAJAPUT; RESHMI; SIDRAMAPPA \(2011\)](#) fazem uso de uma abordagem baseada em transformação de Hough para segmentação do DO e detecção da região da fóvea e seu centro. Assemelha-se a este trabalho tanto por seu objetivo, quanto pela base de imagens utilizada (RIM-ONE R1). Diferencia-se por utilizar uma abordagem diferente.
3. [ALMAZROA et al. \(2015\)](#) propõem um estudo a respeito das metodologias de segmentação do DO e escavação para detecção do glaucoma em imagens. Diferencia-se deste trabalho por não focar em uma técnica em específico, nem tampouco tratar da técnica aqui estudada. Foi de fundamental importância para este trabalho, principalmente por direcionar às linhas de pesquisa e abordagens já utilizadas para este fim e presentes na literatura.
4. [SRIVASTAVA et al. \(2015\)](#) faz o uso de Aprendizado Profundo na implementação de um sistema de segmentação automática do Disco Óptico em imagens da retina, resistente à Atrofia Parapapilar. Semelhantemente a este trabalho, utiliza uma abordagem de Aprendizado Profundo para segmentação automática e tem como saída uma imagem segmentada em “estado de arte”. Diferencia-se porém por sua abordagem ser baseada em camadas estacadas de codificadores automáticos.
5. [TAN et al. \(2017\)](#) faz o uso de uma RNC para segmentar as regiões do DO, Fóvea e Vasculatura Retinal em imagens médicas de fundo de olho. Se assemelha a este trabalho por ter entre seus objetivos a segmentação do DO em imagens médicas de fundo de olho. Diferencia-se abordagem utilizada para atingir este fim.

Dando sequência, são apresentados também os trabalhos relacionados que fazem uso de técnicas semelhantes para objetivos diferentes deste trabalho.

1. [PINHEIRO; COLLORBERT \(2014\)](#) utiliza Redes Neurais Convolucionais Recorrentes, uma técnica deveras semelhante à aplicada neste trabalho, para criar legendas de objetos em cenas (imagens). Assemelha-se com este trabalho por ter o objetivo de preparar imagens para uso posterior. Difere, porém, na técnica utilizada, bem como no espectro do objetivo, tratando de objetos em geral e não de áreas de interesse em imagens médicas.
2. [SHELHAMER; LONG; DARRELL \(2016\)](#) faz uso da abordagem RTC's para segmentação semântica de objetos sem utilização de maquinário posterior, ou seja, aplica a técnica para retirar diretamente dela uma saída em “estado de arte” (uma imagem propriamente dita). Diferencia-se deste trabalho pelo objetivo geral, bem como pela arquitetura da rede montada, sendo neste caso uma RTC pixel-a-pixel, e nesse uma arquitetura estacada de *pooling/unpooling* progressivo. Foi de extrema importância por ser o primeiro trabalho que considerava a saída em “estado de arte”, ou seja, a possibilidade de ter uma imagem propriamente dita como saída da rede.
3. [PAKHOMOV et al. \(2017\)](#) aplica as RTC's para segmentação e adição de máscaras com o intuito final de detecção e identificação de objetos. Difere deste trabalho no que diz respeito ao objetivo final, ao segmentar objetos e não regiões de interesse em imagens médicas. Entretanto, explica de forma elementar os modelos mais utilizáveis de arquitetura de rede, contribuindo amplamente para o desenvolvimento do presente trabalho.

Após analisar os trabalhos relacionados optou-se pela utilização de RTC's por perceber que elas obtiveram bons resultados na detecção e segmentação de imagens. Além disso, não foram encontrados trabalhos que lhes utilizassem para o propósito deste trabalho. O desenvolvimento do método proposto é descrito no Capítulo 4.

4 Método Proposto

No decorrer deste capítulo são abordados ambiente e metodologia utilizada durante o desenvolvimento do presente trabalho.

4.1 Modelos de Rede

Após avaliar a bibliografia encontrada a respeito do tema, foram encontrados diversos modelos de arquitetura compatíveis com a proposta desse trabalho, os quais foram comparadas e avaliadas para determinar qual a melhor performance dentro do objetivo do trabalho. Com isso, foi possível selecionar quatro principais modelos para estudo mais profundo e consequente avaliação da melhor performance. Estes modelos foram escolhidos por apresentar arquiteturas diferentes, seja pela quantidade de camadas ou pelas diferentes formas de abordar o processo de convolução/deconvolução característicos de uma RTC. Tais modelos são descritos à seguir.

- **Modelo 1 - RTC ResNet**

A RTC ResNet é uma arquitetura deveras custosa. Com mais de 100 camadas se dividem entre funções de *pooling* (ver seção 2.3), convolução, *batch normalization* (normalização de lotes), *merge* (fusão), ativação e *upsampling* (sobre-amostragem) (HUNDT, 2017). Uma ilustração da arquitetura da RTC ResNet é mostrada na Figura 7

- **Modelo 2 - RTC VGG**

Construída com apenas 30 camadas e consideravelmente mais leve que o modelo ResNet, a RTC VGG é uma arquitetura menos custosa na qual a maior parte das camadas são de *pooling* e convolução (JU, 2016). Uma ilustração de sua arquitetura pode ser vista na Figura 8.

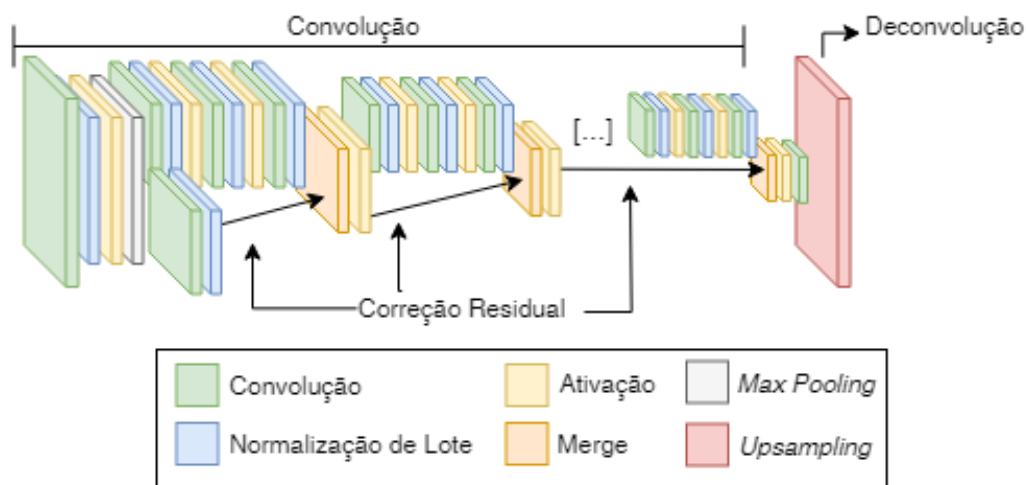
- **Modelo 3 - RTC Segnet**

Utilizando 38 camadas, a arquitetura RTC Segnet é principalmente composta de camadas convolucionais, aliadas a camadas de *pooling* e *padding* (ver seção 2.3). É possível observar uma ilustração desta arquitetura na Figura 9.

- **Modelo 4 - RTC Unet**

Constituída de mínimas 24 camadas, a RTC Unet é com certeza a mais simples das arquiteturas utilizadas, se mantendo fiel ao conceito mais básico de RTC, já ilustrado na Figura 5, por utilizar camadas convolucionais de tamanhos decrescentes "sobrepostas" com camadas deconvolucionais.

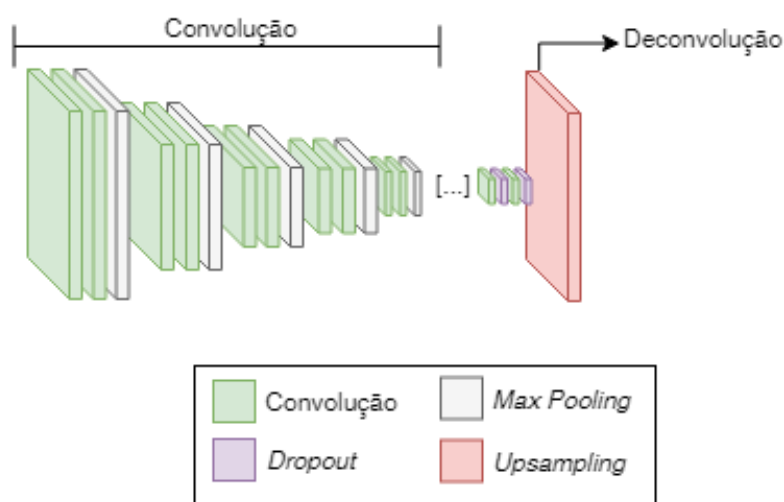
Figura 7: Ilustração da Arquitetura ResNet.



Fonte:

Elaborado pelo autor.

Figura 8: Ilustração da Arquitetura do modelo VGG.



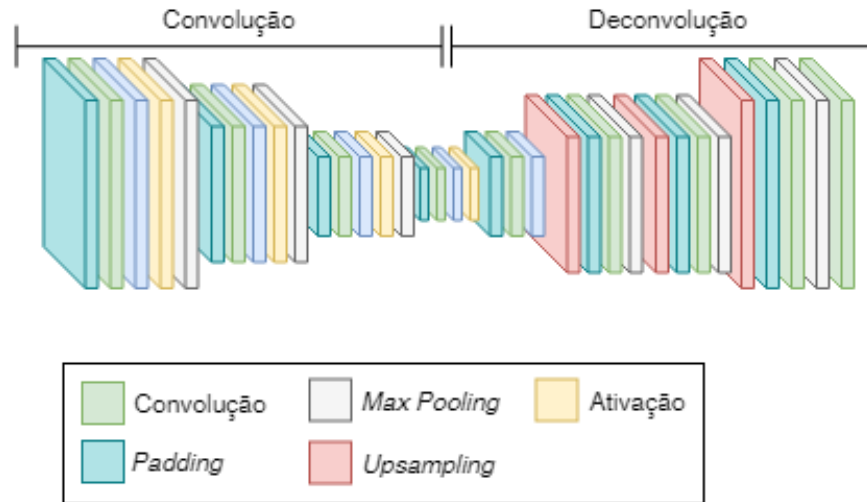
Fonte: Elaborado pelo autor.

4.2 Treinamento e Avaliação

No processo de treinamento dos modelos avaliados foram utilizadas 400 imagens integrantes da primeira e terceira versão da biblioteca RIM-ONE. Isso porque a segunda versão não possui imagens da segmentação feita por profissionais (*ground truth*). Em contrapartida, a terceira versão da biblioteca é constituída de imagens estéreo, o que nos garante duas unidades para cada imagem, bem como suas referentes segmentações. A quantidade de imagens utilizadas em cada etapa são descritas na Tabela 1.

A avaliação de performance dos modelos foi realizada com base em quatro métricas

Figura 9: Ilustração da Arquitetura do modelo Segnet.



Fonte: Elaborado pelo autor.

Tabela 1: Descrição do Experimento.

Etapa	Total de Imagens	Nº Glaucomatosas	Nº Normais	Épocas
Treinamento e Validação	300	138	162	100
Teste	100	50	50	-

Fonte: Elaborado pelo autor.

consideravelmente difundidas. Tais técnicas são: Acurácia de Pixels, Acurácia Média, média da Intercessão sobre União (SHELHAMER; LONG; DARRELL,2016) e Coeficiente Dice de Similaridade (DICE, 1945). Considerando que n_{ij} seja o número de *pixels* da classe i preditos como pertencendo à classe j , onde existem n_{cl} classes diferentes, que $t_i = \sum_j n_{ij}$ seja o número total de *pixels* da classe i , e que A e B são conjuntos de *pixels* pertencentes às áreas de interesse em duas imagens, calculamos:

- **Acurácia de Pixels:** $Ac_Pixel = \sum_i n_{ii} / \sum_i t_i$.
- **Acurácia Média:** $Ac_Média = (1/n_{cl}) \sum_i n_{ii} / t_i$.
- **Média da Intercessão sobre União:** $MIU = (1/n_{cl}) \sum_i n_{ii} / (t_i \sum_j n_{ji} - n_{ii})$.
- **Coeficiente Dice de Similaridade:** $Dice = 2(A \cap B) / (A + B)$.

Quando em se tratando do Coeficiente *Dice* de similaridade, temos comparações matemáticas entre a segmentação realizada pelos modelos sob avaliação e a segmentação disponibilizada pelos produtores da base de dados (*ground truth*), considerando boas seg-

mentações os resultados que apresentem pontuação final superior a 0,70 (ZOU et al., 2004).

Os resultados obtidos ao fim das etapas de treinamento e testes avaliativos serão apresentados e discutidos na Seção 4.3.

4.3 Resultados e Discussões

Para a realização dos experimentos aqui descritos, utilizou-se um computador com acesso à internet, cujas especificações são descritas na Tabela 2.

Tabela 2: Descrição do Ambiente de Desenvolvimento.

Característica	Descrição
Modelo	XL555LF
RAM	12GB
Memória de Vídeo	2GB
Placa de Vídeo	Nvidia GeForce 930M
Processador	Intel I5 5200 @ 2.20GHz
Sistema Operacional	Linux Ubuntu 14.0

Fonte: Elaborado pelo autor.

Todos os modelos foram treinados nas condições descritas na Tabela 1 e tiveram seu desempenho observado e avaliado de acordo com as métricas descritas na Seção 4.2. O resultado dos testes executados são demonstrados na Tabela 3

Tabela 3: Resultados Obtidos.

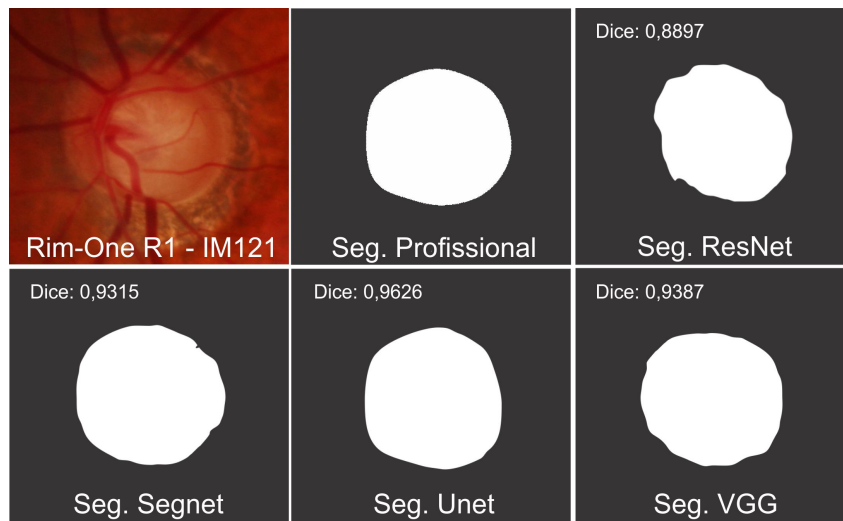
Modelo	Ac. Pixel	Ac. Média	MIU	Dice
Resnet	0,7204	0,4795	0,3486	0,7062
Segnet	0,8327	0,6143	0,4829	0,8138
Unet	0,8725	0,6309	0,4987	0,8516
VGG	0,8891	0,6537	0,5330	0,8729

Fonte: Elaborado pelo autor.

Ao observar os resultados obtidos pela RTC VGG, bem como a notável performance dos demais modelos em estudo, não apenas se prova a viabilidade de utilização de RTC's na segmentação semântica de imagens médicas, mas também se indica uma abordagem de Aprendizado Profundo ainda pouco explorada, mas de grande valia para uma linha de pesquisa inteira. Observe na Figura 10 exemplos de segmentações realizadas pelos modelos em estudo.

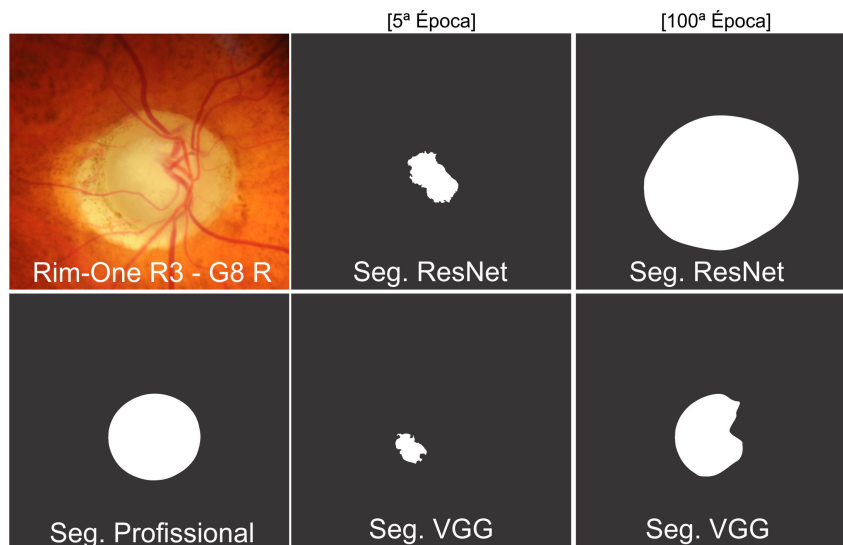
Existiram também falhas de segmentação, tanto durante o treinamento como também ao fim, durante os testes. Observe na Figura 11 a demonstração da segmentação de uma imagem utilizando pesos da 5ª e 100ª épocas, ambas consideradas incorretas.

Figura 10: Demonstração das Segmentações Feitas por Cada Modelo.



Fonte: Adaptação de RIM-ONE R1.

Figura 11: Demonstração de Falhas na Segmentação.



Fonte: Adaptação de RIM-ONE R3.

Realizando um rápido comparativo entre os trabalhos apresentados no Capítulo 3 que se utilizam de abordagens e métricas avaliativas semelhantes, apresentado na Tabela 4, percebe-se que o sistema proposto teve uma performance aceitável, podendo ser defendido como uma solução viável para a segmentação automática de imagens médicas. É possível notar que o trabalho que obteve índices de acurácia superior aos obtidos por este trabalho utilizou uma quantidade muito maior de amostras na sua base de treinamento, sendo estas amostras de bancos de imagens privados.

É importante também ressaltar que a utilização de maiores bases de treinamento implica na utilização de maior poder computacional, o que se pode observar quando

Tabela 4: Comparativo Entre os Resultados Obtidos.

Trabalho	Ac_Pixel (%)	Nº de Amostras
SRIVASTAVA et al.	90,3	200.000
RAJAPUT; RESHMI; SIDRAMAPPA	88,7	169
RTC VGG	88,9	400

Fonte: Elaborado pelo autor.

(TAN et al., 2017, p. 12) mencionam ao descrever seu ambiente, uma estação de trabalho equipada com dois processadores Intel Xeon de 2.20Ghz (E5-2650) e 512GB de RAM para processar suas mais de 75000 imagens. Embora TAN et al. possua objetivos semelhantes, não foi mencionado neste comparativo por fazer o uso de métricas avaliativas diferentes.

5 Conclusão

Neste trabalho avaliamos o desempenho de quatro arquiteturas de Redes Totalmente Convolucionais quando aplicadas para segmentação automática do Disco Óptico em imagens médicas de fundo de olho. Os resultados obtidos demonstram que nas condições do nosso experimento, a aplicação das RTC's em arquitetura VGG performou com maior acurácia e menores índices de perda.

De forma geral, considera-se cumprido o objetivo primário do trabalho. Os resultados obtidos serão de considerável valia para a continuação da pesquisa, bem como de futuros trabalhos de terceiros que poderão ter como base os dados aqui levantados.

São identificadas como principais contribuições deste trabalho o desenvolvimento de um sistema de segmentação automática de imagens médicas, a utilização de uma abordagem de Aprendizado Profundo ainda não explorada para este objetivo, como também a comparação de arquiteturas com diferentes número de camadas.

As prováveis contribuições futuras derivadas deste trabalho são descritas na Seção 5.1.

5.1 Trabalhos Futuros

São muitas as possibilidades de continuação desta linha de pesquisa. Entre elas são destacadas a identificação de características e parâmetros que possam melhorar o desempenho do modelo implementado, implementação da ferramenta de detecção automática de danos glaucomatosos em imagens de fundo de olho, com a possibilidade de diagnose em tempo real com base em *stream* de vídeo, aproveitando do sistema criado para segmentação em tempo real; exploração do potencial das RTC's na segmentação de imagens; levantamento de dados e elaboração de artigos que sejam relevantes neste campo de pesquisa.

Referências

- ABDULLA, W. *Mask R-CNN for object detection and instance segmentation on Keras and TensorFlow*. [S.l.]: Github, 2017. https://github.com/matterport/Mask_RCNN. Citado na página 15.
- ALMAZROA, A. et al. Optic disc and optic cup segmentation methodologies for glaucoma image detection: A survey. *Journal of Ophthalmology*, 2015. Citado na página 22.
- BUSSEL, I.; WOLLSTEN, G.; SCHUMAN, J. Oct for glaucoma diagnosis, screening and detection of glaucoma progression. *The British Journal of Ophthalmology*, 2013. Citado na página 13.
- DICE, L. Measures of the amount of ecologic association between species. *Ecology*, p. 297–302, 1945. Citado na página 26.
- FUMERO, F. et al. Rim-one: An open retinal image database for optic nerve evaluation. *Int. Sym. on CBMS*, p. 1 – 6, 07 2011. Citado na página 20.
- HUNDT, A. *Keras-tensorflow implementation of Fully Convolutional Networks for Semantic Segmentation*. [S.l.]: Github, 2017. <https://github.com/aurora95/Keras-FCN>. Citado na página 24.
- JU, J. *A playable implementation of Fully Convolutional Networks with Keras*. [S.l.]: Github, 2016. <https://github.com/JihongJu/keras-fcn>. Citado na página 24.
- LAURETTI, C. R.; FILHO, A. L. Glaucomas. *Medicina (Ribeirão Preto. Online)*, 1997. Citado na página 13.
- LECUN, I.; BENGIO, Y.; HINTON, G. Deep learning. *Nature - International Journey of Science*, 2017. Citado na página 18.
- MONTEIRO, J. G. *História do Glaucoma*. Loures, Portugal: Théa Portugal, 2014. Citado na página 13.
- PAKHOMOV, D. et al. Deep residual learning for instrument segmentation in robotic surgery. *Cornell University Library*, 2017. Citado na página 23.
- PINHEIRO, P. O.; COLLORBERT, R. Recurrent convolutional neural networks for scene labeling. *31st International Conference on Machine Learning, Beijing, China*, 2014. Citado na página 23.
- RAJAPUT, G.; RESHMI, B.; SIDRAMAPPA, C. Automatic localization of fovea center using mathematical morphology in fundus images. *International Journal of Machine Intelligence*, v. 3, n. 4, 2011. Citado 2 vezes nas páginas 22 e 29.
- REIS ALEXANDRE SOARES CASTRO, M. N.; BURGOYNE, C. F.; CHAUHAN, B. C. Optic disc margin anatomy in patients with glaucoma and normal. controls with spectral domain optical coherence tomography. *Ophthalmology - Journal - Elsevier*, v. 119, n. 4, p. 738–747, 2012. Citado na página 13.

- ROSEBROCK, A. *Deep Learning for Computer Vision with Python, 1ª Edição*. [S.l.]: PYIMAGESEARCH, 2017. Citado 3 vezes nas páginas 17, 18 e 19.
- SHELHAMER, E.; LONG, J.; DARRELL, T. Fully convolutional networks for semantic segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2016. Citado 3 vezes nas páginas 15, 23 e 26.
- SRIVASTAVA, R. et al. Using deep learning for robustness to parapapillary atrophy in optic disc segmentation. In: *2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI)*. [S.l.: s.n.], 2015. p. 768–771. ISSN 1945-7928. Citado 2 vezes nas páginas 22 e 29.
- STAAL, J. et al. Ridge-based vessel segmentation in color images of the retina. *IEEE Transactions on Medical Imaging*, 2010. Citado na página 22.
- TAN, J. H. et al. Segmentation of optic disc, fovea and retinal vasculature using a single convolutional neural network. *ARXIV*, 2017. Citado 2 vezes nas páginas 22 e 29.
- THEANO, D. T. *Deep Learning 0.1 Documentation*. [S.l.], 2013. Disponível em: <http://deeplearning.net/tutorial/fcn_2D_seg.html>. Acesso em: 11.9.2018. Citado na página 19.
- WEINREB, R. N.; KHAW, P. T. Primary open-angle glaucoma. *THE LANCET*, 2004. Citado na página 13.
- ZHANG YI-XIN, H.-B. H.; WEI, S.-H. Clinical characteristics of nonglaucomatous optic disc cupping. *Experimental and Therapeutic Medicine*, v. 1, n. 7, p. 995–999, 2014. Citado na página 13.
- ZOU, K. et al. Statistical validation of image segmentation quality based on a spatial overlap index. *Academic radiology*, 2004. Citado na página 27.



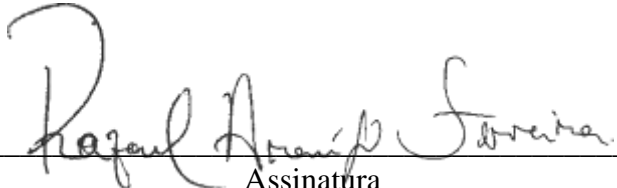
**TERMO DE AUTORIZAÇÃO PARA PUBLICAÇÃO DIGITAL NA BIBLIOTECA
“JOSÉ ALBANO DE MACEDO”**

Identificação do Tipo de Documento

- () Tese
- () Dissertação
- (**X**) Monografia
- () Artigo

Eu, **Rafael Araújo Ferreira**, 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 **Deteccção e Segmentação do Disco Óptico em Imagens da Retina** 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, 12 de dezembro de 2018.


Assinatura