

UNIVERSIDADE FEDERAL DO AMAZONAS
FACULDADE DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

EXPLORANDO TÉCNICAS DE APRENDIZADO PROFUNDO NO
DIAGNÓSTICO DE DOENÇAS INFECCIOSAS

ANNE DE SOUZA OLIVEIRA

MANAUS
2024

UNIVERSIDADE FEDERAL DO AMAZONAS
FACULDADE DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

ANNE DE SOUZA OLIVEIRA

EXPLORANDO TÉCNICAS DE APRENDIZADO PROFUNDO NO
DIAGNÓSTICO DE DOENÇAS INFECCIOSAS

Tese apresentada ao Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal do Amazonas, como requisito parcial para obtenção do título de Doutora em Engenharia Elétrica na área de concentração Controle e Automação de Sistemas.

Orientadora: Profa. Dra. Marly Guimarães Fernandes Costa
Coorientador: Prof. Dr. Cícero Ferreira Fernandes Costa Filho

MANAUS
2024

Ficha Catalográfica

Ficha catalográfica elaborada automaticamente de acordo com os dados fornecidos pelo(a) autor(a).

Oliveira, Anne de Souza
O48e Explorando Técnicas de Aprendizado Profundo no Diagnóstico de
Doenças Infecciosas / Anne de Souza Oliveira . 2024
109 f.: il. color; 31 cm.

Orientadora: Marly Guimarães Fernandes Costa
Coorientador: Cícero Ferreira Fernandes Costa Filho
Tese (Doutorado em Engenharia Elétrica) - Universidade Federal
do Amazonas.

1. Redes neurais convolutivas. 2. Aprendizado profundo. 3.
Malária. 4. Covid-19. I. Costa, Marly Guimarães Fernandes. II.
Universidade Federal do Amazonas III. Título



Ministério da Educação
Universidade Federal do Amazonas
Coordenação do Programa de Pós-Graduação em Engenharia Elétrica

FOLHA DE APROVAÇÃO

Poder Executivo Ministério da Educação
Universidade Federal do Amazonas
Faculdade de Tecnologia
Programa de Pós-graduação em Engenharia Elétrica

Pós-Graduação em Engenharia Elétrica. Av. General Rodrigo Octávio Jordão Ramos, nº 3.000 - Campus Universitário, Setor Norte - Coroado, Pavilhão do CETELI. Fone/Fax (92) 99271-8954 Ramal:2607. E-mail: ppgee@ufam.edu.br

ANNE DE SOUZA OLIVEIRA

EXPLORANDO TÉCNICAS DE APRENDIZADO PROFUNDO NO DIAGNÓSTICO DE DOENÇAS INFECCIOSAS

Tese apresentada ao Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal do Amazonas, como requisito parcial para obtenção do título de Doutora em Engenharia Elétrica na área de concentração Controle e Automação de Sistemas.

Aprovada em 17 de dezembro de 2024.

BANCA EXAMINADORA

Profª. Dra. Marly Guimarães Fernandes Costa- Presidente
Profª. Dra Maria da Graça Vale Barbosa Guerra - Membro Titular 1 - Externo
Prof. Dr. Celso Barbosa Carvalho - Membro Titular 2 - Interno
Prof. Dr. Jorge Augusto de Oliveira Guerra - Membro Titular 3 - Externo
Prof. Dr. Thiago Brito Bezerra - Membro Titular 4 - Interno

Documento assinado eletronicamente

Aprovada em 17 de dezembro de 2024.

Manaus, 27 de janeiro de 2025.



Documento assinado eletronicamente por **Thiago Brito Bezerra, Professor do Magistério Superior**, em 06/02/2025, às 08:30, conforme horário oficial de Manaus, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Celso Barbosa Carvalho, Professor do Magistério Superior**, em 06/02/2025, às 08:32, conforme horário oficial de Manaus, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Marly Guimarães Fernandes Costa, Professor do Magistério Superior**, em 06/02/2025, às 08:37, conforme horário oficial de Manaus, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Maria das Graças Vale Barbosa Guerra, Usuário Externo**, em 06/02/2025, às 09:16, conforme horário oficial de Manaus, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Jorge Augusto de Oliveira Guerra, Usuário Externo**, em 06/02/2025, às 09:18, conforme horário oficial de Manaus, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



A autenticidade deste documento pode ser conferida no site https://sei.ufam.edu.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **2426649** e o código CRC **5E53BAE7**.

Av. General Rodrigo Octávio Jordão Ramos, nº 3.000 - Bairro Coroado Campus Universitário, Setor Norte -
Telefone: 99271-8954
CEP 69080-900 Manaus/AM - Pavilhão do CETELI. E-mail: ppgee@ufam.edu.br

Referência: Processo nº 23105.050636/2024-13

SEI nº 2426649

Criado por [31183646291](#), versão 2 por [31183646291](#) em 27/01/2025 16:37:24.

AGRADECIMENTOS

Às deusas e aos deuses africanos, por terem me guiado sempre que precisei.

À minha família e amigos, por serem fonte de amor. Eu amo vocês.

À minha orientadora, Profa. Dra. Marly Guimarães Fernandes Costa, e ao meu coorientador, Prof. Dr. Cícero Ferreira Fernandes Costa Filho, por compartilharem seus conhecimentos e contribuírem de forma significativa para o desenvolvimento deste trabalho.

À Fundação de Amparo à Pesquisa do Estado do Amazonas (FAPEAM), no período de 01/03/2020 a 28/02/2023; à Coordenação de Aperfeiçoamento do Pessoal de Nível Superior (CAPES), no período de 01/05/2023 a 29/02/2024; e ao projeto SUPER, em dois momentos: o primeiro, no período de 04/05/2020 a 21/07/2021, e o segundo, no período de 10/10/2022 a 29/02/2024, pelo fornecimento de bolsa de estudos, que permitiu dedicação exclusiva ao desenvolvimento da tese.

À Universidade Federal do Amazonas e, em especial, ao Centro de Pesquisa e Desenvolvimento de Tecnologia Eletrônica e da Informação (CETELI), por propiciar a infraestrutura necessária à realização deste trabalho.

Este trabalho foi desenvolvido com o apoio do Governo do Estado do Amazonas por meio Fundação de Amparo à Pesquisa do Estado do Amazonas, com a concessão de bolsa de estudo.

O presente trabalho foi realizado com o apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

Esta pesquisa, realizada no âmbito do Projeto Samsung-UFAM de Ensino e Pesquisa (SUPER), de acordo com o Artigo 39 do Decreto nº10.521/2020, foi financiada pela Samsung Eletrônica da Amazônia Ltda, nos termos da Lei Federal nº8.387/1991, através do convênio 001/2020 firmado com a UFAM e FAEPI, Brasil.

RESUMO

A Organização Mundial da Saúde (OMS) define doenças infecciosas como aquelas causadas por microrganismos patogênicos, como bactérias, vírus, parasitas e fungos. Essas doenças podem ter altos índices de mortalidade e sérias repercussões globais, devido à natureza rápida e inesperada de sua disseminação. Um exemplo é a malária, uma doença infecciosa potencialmente fatal, causada por protozoários transmitidos às pessoas pela picada das fêmeas de mosquitos do gênero *Anopheles*. Outro exemplo é a COVID-19, causada pelo coronavírus SARS-CoV-2, que também teve uma rápida e inesperada disseminação, gerando uma crise mundial de saúde. Este estudo propõe métodos automáticos para a detecção de parasitas da malária em imagens de microscopia ótica e a segmentação de área pulmonar afetada pela COVID-19 em imagens tomográficas. Para a malária foi desenvolvido um método para imagens microscópicas de gota espessa, utilizando classificadores de pixels e uma rede neural profunda de 34 camadas. A metodologia também emprega componentes HSV da imagem como variáveis de entrada para perceptron multicamadas e árvore de decisão. O conjunto de dados foi segmentado em diferentes grupos baseados nos tamanhos dos parasitas, permitindo a avaliação de métricas de desempenho como acurácia, sensibilidade, especificidade, precisão e *F1-score*. Os melhores resultados foram obtidos em imagens com parasitas maiores, com taxas de precisão de 91,71% e 93,14%. Em imagens com parasitas menores, as taxas de precisão foram mais baixas, alcançando 76,58% e 71,58%. Os resultados são comparáveis aos de estudos publicados anteriormente. Em relação à COVID-19, foi desenvolvido um método para ser usado com imagens de tomografias computadorizadas e foram comparadas três abordagens de segmentação automática: 1) segmentação das áreas afetadas a partir de duas estratégias de particionamento de dados (estratégias de exame e *slice*); 2) segmentação automática e concordância interobservador entre dois grupos de radiologistas; e 3) desempenho de segmentação utilizando métodos manuais e automáticos de segmentação pulmonar. Foram usados dois conjuntos de dados e duas arquiteturas de redes neurais convolutivas (U-Net e CNN2). As métricas Dice para segmentação automática da COVID-19 foram de $73,01 \pm 9,47\%$ e $84,66 \pm 5,41\%$ para as estratégias de exame e *slice*, respectivamente. Com a segmentação manual dos pulmões, as métricas Dice foram $74,47 \pm 9,94\%$ e $85,35 \pm 5,41\%$, respectivamente. A estratégia de *slice* obteve um desempenho ligeiramente melhor que a de exame. Não foram observadas diferenças estatisticamente significativas entre a segmentação automática e a concordância interobservador em um grupo de sete tomografias computadorizadas. Os resultados destacam a eficácia dos métodos propostos, com bom desempenho nas métricas de avaliação.

Palavras-chave: redes neurais convolutivas, aprendizado profundo, malária, COVID-19, imagem de gota espessa, tomografia computadorizada.

ABSTRACT

The World Health Organization (WHO) defines infectious diseases as those caused by pathogenic microorganisms such as bacteria, viruses, parasites, and fungi. These diseases can have high mortality rates and serious global repercussions due to the rapid and unexpected nature of their spread. One example is malaria, a potentially fatal infectious disease caused by protozoa transmitted to people through the bite of female mosquitoes of the genus *Anopheles*. Another example is COVID-19, caused by the SARS-CoV-2 coronavirus, which also spread rapidly and unexpectedly, creating a global health crisis. This study proposes automated methods for detecting malaria parasites in optical microscopy images and segmenting lung areas affected by COVID-19 in medical tomography images. For malaria, a method was developed for microscopic images of thick blood smear, utilizing pixel classifiers and a 34-layer deep neural network. The methodology also uses HSV components of the image as input variables for multilayer perceptron and decision tree classifiers. The dataset was segmented into different groups based on parasite sizes, allowing the evaluation of performance metrics such as accuracy, sensitivity, specificity, precision, and F1-score. The best results were obtained in images with larger parasites, with accuracy rates of 91.71% and 93.14%. For images with smaller parasites, accuracy rates were lower, reaching 76.58% and 71.58%. The results are comparable to those of previously published studies. Regarding COVID-19, a method was developed to be used with computed tomography (CT) images, and three automatic segmentation approaches were compared: 1) segmentation of affected areas using two data partitioning strategies (examination and slice strategies); 2) automatic segmentation and interobserver agreement between two groups of radiologists; and 3) segmentation performance using both manual and automatic lung segmentation methods. Two datasets and two convolutional neural network architectures (U-Net and CNN2) were used. The Dice metrics for automatic COVID-19 segmentation were $73.01 \pm 9.47\%$ and $84.66 \pm 5.41\%$ for the examination and slice strategies, respectively. For manual lung segmentation, the Dice metrics were $74.47 \pm 9.94\%$ and $85.35 \pm 5.41\%$, respectively. The slice strategy showed slightly better performance than the examination strategy. No statistically significant differences were observed between automatic segmentation and interobserver agreement in a group of seven CT scans. The results highlight the effectiveness of the proposed methods, showing good performance in evaluation metrics.

Keywords: convolutional neural networks, deep learning, malaria, COVID-19, thick blood smear image, computed tomography.

LISTA DE ILUSTRAÇÕES

Figura 1	– Exemplos de (a) Lâmina de esfregaço. (b) Imagem microscópica de esfregaço. (c) Lâmina de gota espessa. (d) Imagem microscópica de gota espessa. . . .	17
Figura 2	– Exemplos de parasitas corados com coloração de Giemsa em gota espessa. (a) Trofozoíto anelar ou imaturo. (b) Trofozoíto amebóide ou maduro. (c) Esquizonte. (d) Gametócito.	17
Figura 3	– Quantidade de casos de COVID-19 registrados globalmente às 16:37 do dia 18 de Outubro de 2023.	18
Figura 4	– Exemplo de uma arquitetura de rede neural convolutiva.	19
Figura 5	– Histogramas do número de parasitas por imagem em cada conjunto. (a) Conjunto 1. (b) Conjunto 2. (c) Conjunto 3. (d) Conjunto 4.	32
Figura 6	– Fluxograma da proposta para detecção de parasitas da malária em imagens microscópicas de gota espessa.	33
Figura 7	– Apresentação de uma imagem microscópica de gota espessa de sangue utilizando o modelo de cor HSV mostrando os objetos em primeiro plano (WBCs, RBCs, parasitas) envolvidos em regiões com tonalidade (componente H) diferentes daquela do fundo da imagem. Além disso, verifica-se fragmentos de células RBC na imagem.	33
Figura 8	– Distribuição, em função das componentes do espaço HSV, dos valores dos pixels extraídos de objetos em primeiro plano, pixels de regiões próximas a objetos em primeiro plano e pixels pertencentes ao fundo: (a) saturação versus tonalidade (S x H); (b) intensidade versus tonalidade (V x H); (c) intensidade versus saturação (V x S).	34
Figura 9	– Perfis dos componentes H, S e V dos pixels de uma região próxima a um parasita e o próprio parasita, versus o fundo. (a) Perfil H. (b) Perfil S. (c) Perfil V.	35
Figura 10	– (a) Uma imagem de gota espessa. (b) Resultado da detecção em primeiro plano. (c) Resultado da extração de regiões próximas a objetos em primeiro plano. (d) Imagem de fusão. (e) <i>Patches</i> correspondentes na imagem original. (f) Alguns <i>patches</i> ampliados.	36
Figura 11	– Número de <i>patches</i> com área menor que 1.600 pixels (40 × 40) e com área maior que 1.600 pixels em cada um dos conjuntos de imagens: (a) conjunto 1; (b) conjunto 2; (c) conjunto 3; (d) conjunto 4.	37
Figura 12	– Arquitetura da CNN1 e CNN2 com 34 camadas e 4 blocos de extração de características.	39
Figura 13	– Desempenho do erro médio quadrático durante o treinamento do MLP para classificação de objetos em primeiro plano.	40

Figura 14 –	Árvore de decisão para classificação de objetos em primeiro plano. (b) arquitetura da árvore (a variável x_2 corresponde ao componente S) (a) Erro de validação cruzada de 10 partes como função do tamanho mínimo da folha. (b) Arquitetura da árvore (a variável x_2 corresponde ao componente S).	40
Figura 15 –	(a) Imagem de dois parasitas próximos um do outro. (b) Regiões segmentadas próximas aos dois parasitas unidos. (c) Regiões separadas usando operação de <i>watershed</i> .	41
Figura 16 –	Imagens de <i>patches</i> extraídos. (a) Pequenos <i>patches</i> positivos. (b) Pequenos <i>patches</i> negativos. (c) Grandes <i>patches</i> negativos.	41
Figura 17 –	Formas de segmentação da COVID-19. (a) Segmentação de uma etapa. (b) Segmentação de duas etapas.	46
Figura 18 –	Gráfico mostrando, para cada tomografia computadorizada, o número total de <i>slices</i> (em azul), o número de <i>slices</i> contendo os pulmões (em amarelo) e o número de <i>slices</i> contendo regiões com COVID-19 (em vermelho). (a) Conjunto de dados 1; (b) Conjunto de dados 2.	53
Figura 18 –	Gráfico mostrando, para cada tomografia computadorizada, o número total de <i>slices</i> (em azul), o número de <i>slices</i> contendo os pulmões (em amarelo) e o número de <i>slices</i> contendo regiões com COVID-19 (em vermelho). (a) Conjunto de dados 1; (b) Conjunto de dados 2 (continuação).	54
Figura 19 –	Métodos adotados neste trabalho.	54
Figura 20 –	Etapas realizadas no pré-processamento das imagens.	54
Figura 21 –	Exemplos de <i>slices</i> de tomografias computadorizadas após pré-processamento. De cima para baixo: (a), (b) e (c) <i>slices</i> ; (d), (e) e (f) máscaras pulmonares; (g), (h) e (i) máscaras de COVID-19.	55
Figura 22 –	Conjuntos de treinamento e teste obtidos com a estratégia de exame.	56
Figura 23 –	Conjuntos de treinamento e teste obtidos com a estratégia de <i>slice</i> .	56
Figura 24 –	Arquitetura de rede neural convolutiva usada para segmentação semântica de pulmões e COVID-19 - U-Net.	57
Figura 25 –	Arquitetura de rede neural convolutiva usada para segmentação semântica da COVID-19 – CNN2.	58
Figura 26 –	Comparação das áreas dos pulmões e COVID-19 segmentadas automaticamente com as áreas segmentadas pelos radiologistas no conjunto de teste. (a) Estratégia de exame. (b) Estratégia de <i>slice</i> .	62
Figura 27 –	A primeira linha mostra a segmentação pulmonar e da COVID-19 com a estratégia de exame. A segunda linha mostra a segmentação pulmonar e da COVID-19 com a estratégia de <i>slice</i> . A segmentação dos radiologistas (conjunto de dados 1) é mostrada em azul, enquanto a segmentação automática é mostrada em vermelho.	64

Figura 28 – (a) Imagem original. (b) Concordância entre radiologistas. A segmentação dos radiologistas no conjunto de dados 1 é mostrada em azul, enquanto a segmentação dos radiologistas no conjunto de dados 2 é mostrada em amarelo.	64
Figura 29 – (a) <i>Slice</i> com áreas de baixo contraste. (b) A segmentação dos radiologistas (conjunto de dados 1) é mostrada em azul, enquanto a segmentação automática é mostrada em vermelho.	65

LISTA DE TABELAS

Tabela 1 – Resumo dos trabalhos publicados na última década sobre detecção do parasita da malária em imagens de gota espessa de sangue.	26
Tabela 2 – Resumo dos trabalhos publicados na última década sobre detecção do parasita da malária em imagens de gota espessa de sangue (continuação).	27
Tabela 2 – Resumo dos trabalhos publicados na última década sobre detecção do parasita da malária em imagens de gota espessa de sangue (continuação).	28
Tabela 2 – Resumo dos trabalhos publicados na última década sobre detecção do parasita da malária em imagens de gota espessa de sangue (continuação).	29
Tabela 2 – Resumo dos trabalhos publicados na última década sobre detecção do parasita da malária em imagens de gota espessa de sangue (continuação).	30
Tabela 3 – Número de parasitas em cada conjunto.	31
Tabela 4 – Área do <i>patch</i> por conjunto de imagens (em número de pixels).	37
Tabela 5 – Número de <i>patches</i> para treinamento e validação.	38
Tabela 6 – Resultados para classificação de <i>patches</i> de parasitas (Média±DP).	42
Tabela 7 – Resultados da classificação de parasita na imagem completa (Média±DP) nos diversos conjuntos de imagens microscópicas de gota espessa.	42
Tabela 8 – Resumo dos artigos publicados nos últimos quatro anos sobre segmentação da COVID-19 em imagens de TC com métodos de segmentação de uma etapa.	47
Tabela 9 – Resumo dos artigos publicados nos últimos quatro anos sobre segmentação da COVID-19 em imagens de TC com segmentação de duas etapas.	50
Tabela 10 – Resumo dos artigos publicados nos últimos quatro anos sobre segmentação da COVID-19 em imagens de TC com segmentação de duas etapas (continuação).	51
Tabela 10 – Resumo dos artigos publicados nos últimos quatro anos sobre segmentação da COVID-19 em imagens de TC com segmentação de duas etapas (continuação).	52
Tabela 11 – Hiperparâmetros usados para treinamento dos classificadores.	58
Tabela 12 – Resultados da segmentação pulmonar no conjunto de teste.	60
Tabela 13 – Resultados da segmentação da COVID-19 no conjunto de teste usando segmentação pulmonar automática.	61
Tabela 14 – Resultados da segmentação da COVID-19 no conjunto de teste usando segmentação pulmonar manual.	62
Tabela 15 – Comparação entre a segmentação da COVID-19 com a estratégia de <i>slice</i> e a concordância interobservador.	63

LISTA DE ABREVIATURAS E SIGLAS

ACM	<i>Association for Computing Machinery</i>
ACWE	<i>Active Contour Without Edge</i>
ADAM	<i>Adaptive Moment Estimation</i>
AUC	<i>Area Under the Curve</i>
CNN	<i>Convolutional Neural Networks</i>
CPU	<i>Central Processing Unit</i>
CV	<i>Cross-Validation</i>
DP	Desvio Padrão
DT	<i>Decision Tree</i>
FkM	<i>Fast k-Means Clustering Algorithm</i>
FPS	<i>Frames Per Second</i>
HBM	<i>High Bandwidth Memory</i>
HSV	<i>Hue, Saturation, Value</i>
IA	Inteligência Artificial
IEEE	<i>Institute of Electrical and Electronics Engineers</i>
IGMS	<i>Iterative Global Minimum Screening</i>
INPA	Instituto Nacional de Pesquisa da Amazônia
IoU	<i>Intersection over Union</i>
MGCS	<i>Modified Global Contrast Stretching</i>
MLP	<i>Multi Layer Perceptron</i>
OMS	Organização Mundial da Saúde
PCR	<i>Polymerase Chain Reaction</i>
PD	Particionamento de Dados
RAM	<i>Random Access Memory</i>

RBC	<i>Red Blood Cells</i>
RCNN	<i>Regional Convolutional Neural Networks</i>
RDT	<i>Rapid Diagnostic Test</i>
ReLU	<i>Rectified Linear Unit</i>
RGB	<i>Red, Green, Blue</i>
RMSProp	<i>Root Mean Square Propagation</i>
ROI	<i>Region Of Interest</i>
RT-PCR	<i>Reverse Transcription Polymerase Chain Reaction</i>
SGDM	<i>Stochastic Gradient Descent with Momentum</i>
SSD	<i>Single Shot Detection</i>
SVM	<i>Support Vector Machine</i>
TC	Tomografia Computadorizada
TPU	<i>Tensor Processing Unit</i>
VGG16	<i>Visual Geometry Group 16</i>
WBC	<i>White Blood Cells</i>
YOLO	<i>You Only Look Once</i>

SUMÁRIO

1	INTRODUÇÃO	15
1.1	Objetivo da pesquisa	21
1.2	Organização da tese	21
2	MÉTODO AUTOMÁTICO PARA DETECÇÃO DE PARASITAS DA MALÁRIA	22
2.1	Revisão da literatura	22
2.2	Materiais e Métodos	31
2.2.1	Materiais	31
2.2.2	Métodos	31
2.2.2.1	Detecção de objetos em primeiro plano	33
2.2.2.2	Detecção de regiões para delimitar os parasitas	34
2.2.2.3	Processo de fusão	35
2.2.2.4	Detecção de parasitas	36
2.2.2.5	Configuração experimental	38
2.3	Resultados	38
2.3.1	Resultados para detecção de objetos em primeiro plano	39
2.3.2	Resultados para detecção das regiões próximas a objetos em primeiro plano	40
2.3.3	Resultados para classificação de <i>patches</i> de parasitas	41
2.3.4	Resultados para classificação de parasitas em imagens completas	42
2.4	Discussão	43
3	MÉTODO AUTOMÁTICO PARA DETECÇÃO DE ÁREAS AFETADAS PELA COVID-19	45
3.1	Revisão da literatura	45
3.2	Materiais e Métodos	53
3.2.1	Materiais	53
3.2.2	Métodos	53
3.2.2.1	Pré-processamento	54
3.2.2.2	Estratégias de PD para treinamento, validação e teste	56
3.2.2.3	Classificadores	57
3.2.2.4	Segmentação pulmonar e da COVID-19	59
3.2.3	Ambiente de simulação e métricas de avaliação	59
3.3	Resultados	60
3.3.1	Resultados da segmentação pulmonar automática	60
3.3.2	Resultados da segmentação da COVID-19 automática	60

3.3.2.1	Com segmentação pulmonar automática	60
3.3.2.2	Com segmentação pulmonar manual	61
3.3.3	Áreas segmentadas das tomografias computadorizadas	62
3.3.4	Comparação da segmentação da COVID-19 com a concordância interobservador	63
3.3.5	Exemplos de segmentação pulmonar e da COVID-19 automáticas	63
3.4	Discussão	64
4	CONCLUSÃO E TRABALHOS FUTUROS	67
4.1	Trabalhos futuros	67
	REFERÊNCIAS	69
	APÊNDICE A – ARTIGO PUBLICADO NA REVISTA <i>BIOMEDICAL</i>	
	<i>SIGNAL PROCESSING AND CONTROL</i> EM 2022 . .	74
	APÊNDICE B – ARTIGO PUBLICADO NA REVISTA <i>DIAGNOSTICS</i>	
	EM 2024	90

1 INTRODUÇÃO

A Organização Mundial da Saúde (OMS) define doenças infecciosas como aquelas causadas por microrganismos patogênicos, como bactérias, vírus, parasitas e fungos. Essas doenças podem ter altos índices de mortalidade e sérias repercussões globais, devido à natureza rápida e inesperada de sua disseminação (WHO, 2024a).

A malária é uma doença infecciosa parasitária potencialmente fatal, causada por protozoários transmitidos às pessoas pela picada das fêmeas de mosquitos do gênero *Anopheles*. A incidência da malária diminuiu em todo o mundo entre 2000 e 2023, mas a prevalência aumentou. De acordo com a Organização Mundial da Saúde (OMS) (WHO, 2024b), em 2000 a prevalência foi de 238 milhões de casos, enquanto em 2023 foi de 263 milhões. A incidência (ou seja, número de casos por 1000 habitantes em risco) diminuiu de 80 em 2000 para 60,4 em 2023.

Globalmente, entre 2000 e 2023, o número de mortes diminuiu de 861.000 em 2000 para 597.000 em 2023. Embora a malária afete todas as faixas etárias, em 2000, 86,7% dos casos ocorreram em crianças menores de 5 anos e 73,7% em 2023 (WHO, 2024b). Cerca de 95% dos casos de malária ocorreram em apenas 29 países. Os seguintes países foram responsáveis por 50,4% dos casos de malária no mundo: Nigéria (25,9%), República Democrática do Congo (12,6%), Uganda (4,8%), Etiópia (3,6%), Moçambique (3,5%) (WHO, 2024b).

Nas Américas, entre 2000 e 2023 a incidência de malária diminuiu 73,7%. Vários países experimentaram redução dos casos. Na Venezuela, que em 2000 teve 35.500 casos, em 2023 teve 135.000. Brasil, Colômbia e Venezuela respondem por 76,8% dos casos na região, com a maioria dos casos devido a espécie *P. vivax*. No Brasil, 99,7% da malária é registrada na região amazônica (GOMES et al., 2020), onde as condições ambientais e socioculturais favorecem a expansão de sua transmissão.

Embora, em países de baixa renda, a incidência, prevalência e número de mortes no mundo devido à malária tenham diminuído entre 2000 e 2023, a malária ainda representa um alto risco para viajantes e migrantes, com casos importados em áreas não endêmicas. Por isso, a OMS recomenda que seu diagnóstico precoce e tratamento rápido sejam as primeiras etapas estabelecidas em qualquer programa de controle (WHO, 2021b).

Os sintomas da malária incluem febre, calafrio e dor de cabeça. No entanto, esse quadro clássico pode ser alterado pelo uso de medicamentos profiláticos ou pela aquisição de imunidade. Na ausência de parâmetros clínicos específicos que permitam confirmar a infecção, é necessário o uso de métodos laboratoriais para o diagnóstico. Os dois métodos mais frequentemente usados para o diagnóstico parasitológico são a microscopia de campo claro e o teste de diagnóstico rápido (RDT). O RDT detecta antígenos ou enzimas específicos do parasita (sejam específicos do gênero ou espécie). A microscopia de campo claro ainda é considerada o padrão ouro com o qual outros métodos tem sido tradicionalmente comparados. A microscopia de campo claro tem as seguintes vantagens sobre outros métodos de diagnóstico: alta sensibilidade; baixos custos

diretos, se a infraestrutura laboratorial para manter o serviço estiver disponível; diferenciação de espécies de *Plasmodium*; determinação de densidades de parasitas e, por fim, por permitir o monitoramento de respostas à terapia (WHO, 2021a).

No entanto, a precisão do diagnóstico depende muito da experiência do microscopista. Treinar e gerenciar microscopistas especializados é caro. Isso é ainda mais verdadeiro em países de baixa renda. Em um estudo para avaliar a acurácia da microscopia da malária em instalações de saúde privadas em Entebbe, Uganda (MUTABAZI et al., 2021), os autores concluíram que a acurácia no diagnóstico da malária usando microscopia foi alta. No entanto, um terço dos pacientes diagnosticados com malária não tinham a doença. A maioria dos erros foi cometida por microscopistas com pouca experiência.

O diagnóstico microscópico da malária é baseado na contagem do número de parasitas (parasitemia) em lâmina de sangue corada. Existem dois tipos de exames microscópicos: esfregaço e gota espessa. No primeiro, uma gota de sangue é espalhada sobre a lâmina de vidro, formando um esfregaço, conforme mostrado na Figura 1(a). No segundo, uma gota de sangue é espalhada sobre uma superfície menor, entre $1,2\text{ cm}^2$ e $1,5\text{ cm}^2$, conforme mostrado na Figura 1(c) (Ethiopian Public Health Institute, 2020; Health Surveillance Secretariat, 2009).

Antes que os filmes sejam avaliados por microscopia, eles são corados para destacar os parasitas. Na preparação do esfregaço, após a deposição de sangue, a lâmina é fixada com álcool metílico por um minuto e então corada com coloração de Giemsa, na proporção de uma gota de corante para 1ml de água tamponada. A Figura 1(b) mostra a aparência do esfregaço sob o microscópio em ampliação de $1000\times$. Nesta figura, um glóbulo vermelho (RBC) e parasitas dentro dos RBCs são mostrados. Outros elementos sanguíneos, como glóbulos brancos (WBC) e plaquetas também podem ser encontrados em imagens de lâminas de esfregaço. Na preparação de gota espessa, as lâminas são inicialmente submetidas a uma solução hipotônica de azul de metileno para desglobalização. Então, a lâmina é corada por 10 minutos em uma solução de coloração de Giemsa. A Figura 1(d) mostra a aparência da gota espessa sob o microscópio em ampliação de $1000\times$. Nesta figura, um WBC e parasitas da malária são mostrados. Diferentemente do esfregaço, os parasitas não estão dentro das células RBCs (Health Surveillance Secretariat, 2009).

Devido à fixação dos glóbulos vermelhos e, portanto, mantendo a forma dos parasitas em seu interior inalterada, o método de esfregaço permite um melhor exame de sua morfologia. Na gota espessa, os elementos sanguíneos (incluindo parasitas, se houver) são mais concentrados ($20\sim 30\times$) do que em uma área igual de uma lâmina de filme fino, permitindo que um volume maior de sangue seja examinado. Como um volume maior ($6\text{ }\mu\text{l}$) de sangue é examinado na gota espessa, ela é mais sensível do que a lâmina de esfregaço, sendo mais adequada para detectar baixos níveis de parasitemia (Ethiopian Public Health Institute, 2020) (Health Surveillance Secretariat, 2009). Assim, o exame de gota espessa é realizado para detectar e quantificar (densidade parasitária) os parasitas da malária no diagnóstico microscópico de rotina.

Existem cinco espécies de protozoários *Plasmodium* que infectam humanos: *Plasmodium*

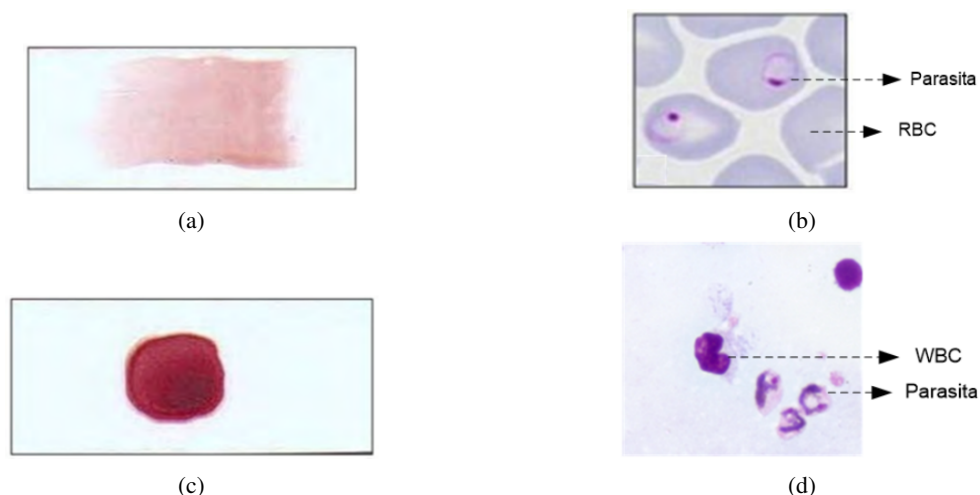


Figura 1 – Exemplos de (a) Lâmina de esfregaço. (b) Imagem microscópica de esfregaço. (c) Lâmina de gota espessa. (d) Imagem microscópica de gota espessa.

falciparum, *Plasmodium vivax*, *Plasmodium ovale*, *Plasmodium malariae* e *Plasmodium knowlesi* (Ethiopian Public Health Institute, 2020). Dentro dos RBCs, dependendo do seu estágio de desenvolvimento e características específicas, os parasitas podem assumir várias formas e tamanhos (comprimento entre $1\ \mu\text{m}$ e $5\ \mu\text{m}$ ou entre $40\ \mu\text{m}$ e $80\ \mu\text{m}$ de comprimento): trofozoíto anelar ou imaturo; trofozoíto amebóide ou maduro; esquizonte e gametócito. A Figura 2 mostra exemplos de parasitas *P. vivax* corados com coloração de Giemsa em gota espessa. Da Figura 2, tiramos algumas conclusões sobre o problema da detecção do parasita da malária em gota espessa: 1. Dependendo do estágio reprodutivo, o parasita tem tamanhos diferentes. Os menores são os trofozoítos imaturos. Os tamanhos maiores são os gametócitos; 2. Um parasita pode ser composto de vários objetos corados (por exemplo, núcleos + citoplasma); 3. Um parasita pode assumir muitas formas. Essas diferentes características representam um desafio na detecção automática de parasitas da malária.

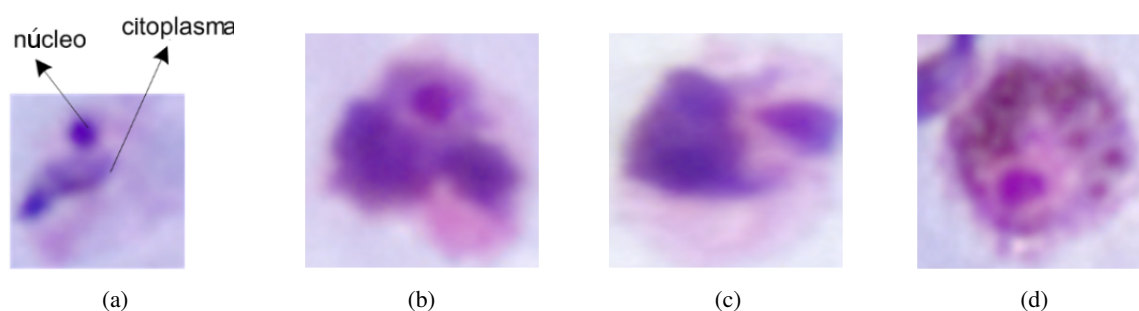


Figura 2 – Exemplos de parasitas corados com coloração de Giemsa em gota espessa. (a) Trofozoíto anelar ou imaturo. (b) Trofozoíto amebóide ou maduro. (c) Esquizonte. (d) Gametócito.

Uma doença infecciosa que teve sérias repercussões globais, devido à natureza rápida e inesperada de sua disseminação foi a COVID-19, causada pelo coronavírus SARS-CoV-2. Nenhuma das variantes do coronavírus anteriores tinha causado doenças graves em humanos nas últimas décadas, mas em 31 de dezembro de 2019, a Organização Mundial da Saúde foi alertada

sobre casos de pneumonia na cidade de Wuhan, China, que foram confirmados em 7 de janeiro de 2020 como pertencentes a variante do SARS-CoV-2 (PAHO, 2020).

Em 11 de março de 2020, a COVID-19 foi caracterizada como uma pandemia devido à distribuição geográfica da doença em vários países ao redor do mundo. A Figura 3 mostra a quantidade de casos de COVID-19 registrados globalmente no dia 18 de outubro de 2023 às 16:37. A gravidade da COVID-19 foi estratificada pela OMS como crítica, grave e não grave (WHO, 2023a) e os sintomas listados foram febre, cansaço e tosse seca, bem como sintomas menos comuns, como perda de paladar ou olfato, congestão nasal, conjuntivite, dor de garganta, dor de cabeça, dor muscular ou articular, erupções cutâneas, náusea ou vômito, diarreia, calafrios ou tontura.

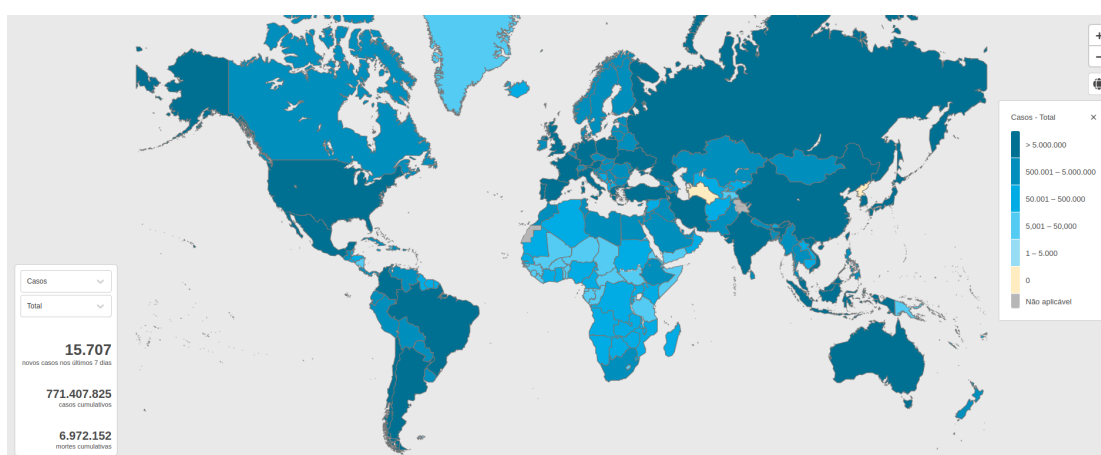


Figura 3 – Quantidade de casos de COVID-19 registrados globalmente às 16:37 do dia 18 de Outubro de 2023.

Fonte: Adaptada de WHO (2023b).

A confirmação do diagnóstico da COVID-19 é realizada com a identificação do RNA viral via Reação em Cadeia da Polimerase de Transcrição Reversa (RT-PCR), considerado o teste padrão-ouro. Entretanto, em situações onde há sintomas da doença, mas o RT-PCR não está disponível; há demora na obtenção dos resultados, ou ainda, os resultados do RT-PCR são negativos, mas há sintomas que sugerem a doença, a confirmação do diagnóstico pode ser feita por meio de exames de imagem do tórax.

Os exames de imagem do tórax incluem radiografia de tórax, tomografia computadorizada (TC) de tórax e ultrassonografia pulmonar. Em uma revisão sistemática realizada em WHO (2020), nos primeiros meses da pandemia de COVID-19, foram encontrados 23 estudos que usaram esses três tipos de imagem para avaliar a acurácia do diagnóstico em relação a um padrão de referência, três dos quais usaram radiografia de tórax, 19 usaram tomografias computadorizadas de tórax e um usou ultrassonografia pulmonar. Entre os 19 estudos de TC, cinco usaram inteligência artificial (IA) para interpretar as imagens. Entre os estudos que usaram IA, dois usaram modelos de aprendizado profundo, sendo que um deles utilizavam redes neurais convolutivas (CNNs). As tarefas realizadas nesses estudos foram em grande parte de classificação, onde o modelo determina se a TC é classificada com ou sem COVID-19.

Na revisão de literatura realizada neste trabalho, também foram encontrados estudos que utilizaram modelos de aprendizado profundo treinados com CNNs para delimitar as regiões dos pulmões afetadas pela COVID-19 em exames de TC. Após a análise dos artigos, foi observado que a segmentação (delimitação) das áreas afetadas é de fundamental importância para o acompanhamento clínico dos pacientes, pois possibilita monitorar a proporção de COVID-19 na região pulmonar em exames de TC. Adicionalmente observou-se que, os estudos que tratam da segmentação da COVID-19 podem ser classificados de acordo com o número de etapas utilizadas nesse processo de segmentação e de acordo com a metodologia utilizada para particionamento dos dados de treinamento e teste do modelo.

As redes neurais convolutivas (CNN) (KUMAR; DESHPANDE, 2018) são um tipo de rede que usam uma operação matemática chamada de convolução para extrair características de uma imagem. Nessa operação, as informações dos dados de entrada e do filtro são multiplicadas e somadas para produzir a saída convoluída.

A Figura 4 apresenta um exemplo de uma arquitetura de rede neural convolutiva. A primeira camada de uma arquitetura de rede neural convolutiva é a camada de entrada, que determina a largura e altura da imagem que será lida, bem como a quantidade de canais da mesma. A segunda camada é a camada de convolução. A camada de convolução converte a imagem lida na camada de entrada em mapas de características. Essa conversão é realizada através de filtros de convolução, *stride* e *padding*. Nas configurações do filtro são determinadas a quantidade de filtros que serão utilizados, bem como a altura e largura de cada filtro. O *stride* determina o número de pixels que serão utilizados para mover o filtro sobre a imagem. O *padding* determina o número de zeros que serão adicionados às bordas da imagem para controlar o tamanho da saída da convolução. Após a camada de convolução é aplicada uma função de ativação.

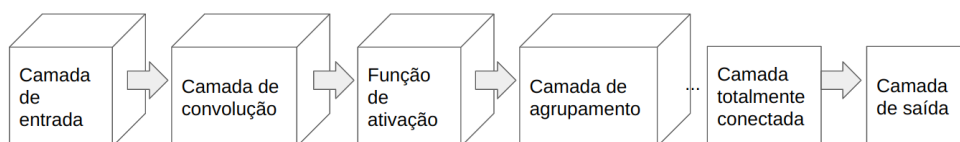


Figura 4 – Exemplo de uma arquitetura de rede neural convolutiva.

A terceira camada de uma arquitetura de rede neural convolutiva é a camada de agrupamento (*pooling*). A camada de agrupamento proporciona a redução da dimensão das características aprendidas, combinando pixels vizinhos de uma área de agrupamento definida por um filtro e um *stride*, em um valor determinado pelo tipo de agrupamento, que pode ser escolhido entre o de média ou máximo. As camadas de convolução, função de ativação e agrupamento podem ser repetidas várias vezes.

A próxima camada de uma arquitetura de CNN é a camada totalmente conectada. Essa camada transforma o mapa de características em um vetor. Cada neurônio nessa camada estará conectado a todos os neurônios da camada anterior. A última camada de uma arquitetura de CNN é a camada de saída. Essa camada também é totalmente conectada, e sua dimensão é igual a quantidade de classes que o conjunto de dados utilizado apresenta.

No aprendizado de máquina, existem os parâmetros do modelo, calculados pelo algoritmo de aprendizado, e os parâmetros que são configurados, os hiperparâmetros (GOODFELLOW; BENGIO; COURVILLE, 2016). A seleção dos hiperparâmetros se baseia na verificação de que o modelo não vá sofrer *underfitting*, ou seja, que o modelo não tenha a capacidade de capturar os detalhes contidos no conjunto de dados de treinamento.

Os hiperparâmetros podem ser divididos em nível de estrutura e topologia de uma rede neural ou em nível de algoritmo de aprendizagem (GÉRON, 2017). Os hiperparâmetros em nível de estrutura e topologia de rede dizem respeito à quantidade de camadas, número de neurônios, funções de ativação, tamanho dos filtros, *stride* e *padding* etc. Os hiperparâmetros em nível de algoritmo de aprendizagem consistem em variáveis externas ao modelo cujo valor não pode ser estimado a partir dos dados, como: taxa de aprendizagem; tamanho do mini lote; momento; decaimento da taxa de aprendizagem e número de épocas de treinamento.

Um conjunto de dados disponibilizado para a etapa de treinamento de uma rede neural convolutiva é uma representação que contém apenas um percentual de amostras de um conjunto de dados completo. A etapa de treinamento consiste em encontrar um modelo apropriado que mantenha a capacidade de generalizar quando uma amostra não conhecida (originária do conjunto de teste) for apresentada. Um modelo sofre *overfitting* quando captura bem os detalhes contidos no conjunto de dados de treinamento, mas erra quando é avaliado no conjunto de dados de teste (GÉRON, 2017).

Para evitar problemas de *overfitting*, algumas técnicas de regularização podem ser usadas. A técnica de regularização L2 evita que os pesos da conexão da rede se tornem muito grandes. A técnica de regularização *dropout* consiste em excluir temporariamente, a cada passo de treinamento, um determinado neurônio que não será ativado no próximo passo (BONACCORSO, 2017). Essa técnica é frequentemente combinada com a técnica L2.

Para dimensionar a capacidade apropriada de um modelo que consiga realizar o diagnóstico automático de uma doença infecciosa, é preciso trabalhar com um modelo suficientemente complexo para evitar problemas de *underfitting*. Isso faz com que o modelo gerado tenha uma boa aproximação no conjunto de dados de treinamento, mas possa apresentar problemas de *overfitting* no conjunto de dados de teste. Por este motivo, explorar técnicas de aprendizado profundo para auxiliar no diagnóstico de doenças infecciosas, como a malária e a COVID-19, é fundamental e desafiador.

Este trabalho explora a aplicação de redes neurais convolutivas tanto para o diagnóstico da malária quanto para a COVID-19, com foco na segmentação de imagens microscópicas e de exames de tomografia computadorizada. A utilização de aprendizado profundo para diagnosticar e monitorar essas doenças infecciosas, especialmente em locais com recursos limitados, pode representar um avanço significativo, permitindo diagnósticos mais rápidos e precisos.

1.1 Objetivo da pesquisa

Propor métodos automáticos baseados em técnicas de aprendizado profundo aplicados a imagens médicas, visando superar as limitações dos métodos tradicionais e melhorar a precisão no diagnóstico da malária e da COVID-19.

Objetivos específicos:

- Propor um método automático para detecção de parasitas da malária *Plasmodium vivax* em imagens microscópicas de gota espessa de sangue;
- Propor um método automático para detecção de áreas de opacidades em vidro fosco afetadas pela COVID-19 em imagens de tomografias computadorizadas.

1.2 Organização da tese

Este trabalho é organizado da seguinte forma:

- **Capítulo 2 - MÉTODO AUTOMÁTICO PARA DETECÇÃO DE PARASITAS DA MALÁRIA**, contendo um fluxo de trabalho desenvolvido para analisar imagens microscópicas de gota espessa de sangue de pacientes infectados com malária, usando técnicas de aprendizado profundo para melhorar o diagnóstico da doença.
- **Capítulo 3 - MÉTODO AUTOMÁTICO PARA DETECÇÃO DE ÁREAS AFETADAS PELA COVID-19**, contendo um fluxo de trabalho desenvolvido para analisar imagens de tomografias computadorizadas de pacientes infectados com COVID-19, usando técnicas de aprendizado profundo para melhorar o diagnóstico da doença.
- **Capítulo 4 - CONCLUSÃO E TRABALHOS FUTUROS**, apresenta as conclusões deste trabalho e os trabalhos futuros.
- **Apêndice A - ARTIGO PUBLICADO NA REVISTA *BIOMEDICAL SIGNAL PROCESSING AND CONTROL* EM 2022**, disponibiliza o primeiro artigo resultante deste trabalho.
- **Apêndice B - ARTIGO PUBLICADO NA REVISTA *DIAGNOSTICS* EM 2024**, disponibiliza o segundo artigo resultante deste trabalho.

2 MÉTODO AUTOMÁTICO PARA DETECÇÃO DE PARASITAS DA MALÁRIA

2.1 Revisão da literatura

Na última década, para amenizar os problemas decorrentes da falta de mão de obra qualificada para o diagnóstico da malária por microscopia de campo claro, um grande esforço tem sido feito na pesquisa e desenvolvimento de soluções automáticas ou semiautomáticas para o diagnóstico da malária por meio da análise de imagens de campo microscópico de lâmina de gota espessa: (NAKASI et al., 2020a; MADHU et al., 2021; PATTANAIK; MITTAL; KHAN, 2020; ABDURAHMAN; FANTE; ALIY, 2021; CHIBUTA; ACAR, 2020; YANG et al., 2020; NAKASI; MWEBAZE; ZAWEDDE, 2021; QUINN et al., 2016; ARIS et al., 2020). A seguir, é feita uma revisão das tendências observadas nos trabalhos publicados na literatura na última década, sobre diagnóstico automático da malária por meio de microscopia, os quais são sumariados na Tabela 2.

Em relação às técnicas utilizadas para identificar os parasitas, são identificados três grupos distintos. No primeiro grupo, são utilizadas apenas técnicas de processamento digital de imagens para detectar os parasitas (ABIDIN; SALAMAH; NUGROHO, 2016; DAVE; UPLA, 2017; AZIF; NUGROHO; WIBIRAMA, 2018; QUINN et al., 2014). O segundo grupo é caracterizado por dois momentos. No primeiro momento, são extraídas características dos objetos. Em um segundo momento, essas características são utilizadas em classificadores como máquinas de vetores de suporte, perceptron multicamada, *k-means*, que diferenciam os parasitas de outras estruturas, como glóbulos brancos, plaquetas e ruído (ARIS et al., 2020; DELAHUNT et al., 2015). Por fim, o terceiro grupo é caracterizado pela utilização de técnicas de aprendizado profundo (MADHU et al., 2021; YANG et al., 2020; HUQ; PERVIN, 2020; PATTANAIK; MITTAL; KHAN, 2020; NAKASI et al., 2020a; ABDURAHMAN; FANTE; ALIY, 2021; CHIBUTA; ACAR, 2020; NAKASI; MWEBAZE; ZAWEDDE, 2021; QUINN et al., 2016; YANG et al., 2019; NAKASI et al., 2020b).

A aparência de imagens microscópicas de gotas espessas pode apresentar características diferentes em termos de cor, baseada no pH da coloração de *Giemsa* (DELAHUNT et al., 2015), devido às diferentes condições de iluminação utilizadas (DAVE, 2017). Para preservar a uniformidade nas imagens, é comum, nos trabalhos do primeiro e segundo grupos, utilizar técnicas de processamento digital de imagem.

Em Dave (2017), as imagens capturadas são convertidas no espaço de cor HSV (do inglês, *Hue, Saturation, Value*) para reduzir variações devido ao brilho e à concentração da solução de coloração, o que, segundo os autores, afeta muito as imagens no espaço de cor RGB (do inglês, *Red, Green, Blue*). Em Aris et al. (2020), para realce de contraste das imagens, é utilizado o algoritmo *Modified Global Contrast Stretching* (MGCS), que destaca e aguça a qualidade da

imagem microscópica. Na sequência, para normalização de cor, foi utilizado um algoritmo *gray world*. Em Abidin, Salamah e Nugroho (2016), os autores melhoram a imagem aplicando técnicas de aumento de contraste (KAUR; CHOUDHARY, 2012) e alongamento escuro (HANIF; MASHOR; MOHAMED, 2011).

Nos artigos do grupo 1, os parasitas da malária são detectados usando apenas técnicas de segmentação emprestadas do processamento digital de imagens, seguidas de operações de filtragem para remover ruído. Em Abidin, Salamah e Nugroho (2016), os parasitas são segmentados usando a técnica de Contorno Ativo Sem Borda (ACWE), uma versão modificada da técnica de contornos ativos, também conhecida como *snakes*. Na sequência, os autores aplicam filtros morfológicos para remover ruído e, finalmente, uma operação de limiarização. Em um conjunto de dados de 50 imagens com 140×140 pixels, os autores relataram uma precisão de 97,57% e uma taxa de falsos positivos de 12,04%.

Em Dave e Upla (2017), os autores usaram 87 imagens de campos microscópicos de 3136×2352 pixels. Cada campo microscópico é capturado em 10 profundidades focais diferentes. O *ground truth* é obtido por meio da anotação manual das coordenadas x-y dos parasitas presentes na imagem realizada por um microscopista especialista no diagnóstico laboratorial de malária. Como um estágio de pré-processamento, a filtragem de ruído é feita usando valores máximos de filtro Laplaciano obtidos das 10 profundidades (ou planos) focais. Para detectar parasitas, os autores primeiro separam o primeiro plano (partículas coradas, ou seja, parasitas, WBCs, plaquetas e ruído) do fundo (partículas não coradas) por meio do histograma de limiar adaptativo da componente de saturação da imagem. Depois, os objetos não parasitas do primeiro plano são eliminados usando a técnica de morfologia matemática denominada de componentes conectados. De acordo com os autores, os parasitas ocupam uma área máxima de 700 pixels. Usando esse limiar de área, objetos que não sejam parasitas são eliminados. Os autores relataram uma acurácia de 90,81% e uma taxa de falsos positivos de 10,12%.

Nos artigos do grupo 2, as seguintes operações são empregadas para detectar os parasitas: segmentação + extração de características + classificador. Em Dave (2017), partículas coradas, ou seja, parasitas em todos os estágios de vida, leucócitos, plaquetas e outros artefatos são segmentados (extraídos) do fundo da imagem (*background*), usando a técnica de limiarização de histograma adaptativo do canal de saturação do espaço de cor HSV. Na sequência, para cada região de interesse (ROI, do inglês *Region Of Interest*), 123 características são extraídas (momentos centrais, momentos invariantes de Hu, momentos de Zernike, área, etc.) e submetidas a uma máquina de vetores de suporte (SVM, do inglês, *Support Vector Machine*). Em um conjunto de dados de 87 imagens de campos microscópicos, de 3136×2352 pixels, coletados de pacientes distintos, com variação na concentração da solução de coloração e densidade do parasita, os autores relataram uma sensibilidade de 86,34% e uma especificidade de 96,60%.

Em Aris et al. (2020), as imagens microscópicas de esfregaços de sangue espessos foram segmentadas e classificadas usando o algoritmo de agrupamento FkM (do inglês, *Fast k-Means Clustering Algorithm*, usando 15 componentes de cor dos seguintes espaços: RGB, CMY, XYZ,

HSV e YUV. No algoritmo FkM, uma combinação do algoritmo *k-means* com valor de função discreta do histograma de nível é empregada para determinar os centros de agrupamento do treinamento do conjunto de imagens. Os componentes de cor extraídos são agrupados em três categorias, que são leucócitos e parasitas da malária (preto), precipitação de coloração (cinza) e fundo (branco). O algoritmo de *watershed* é usado para separar parasitas conectados, e um filtro de ruído descarta objetos menores que 7 pixels e maiores que 200 pixels (WBCs). Usando um conjunto de dados de 100 imagens de 1294×980 pixels de dimensão, os autores relataram uma precisão de 99,74%, uma sensibilidade de 77,30% e uma especificidade de 99,76%.

Uma outra abordagem para detecção e classificação de parasitas em imagens completas envolve a obtenção de *patches*. *Patches* são subimagens extraídas de imagens dos campos microscópicos. É o caso dos artigos do grupo 3, em que técnicas de aprendizado profundo são usadas para classificação de *patches* em positivos (contendo parasita) ou negativos (sem parasita) (MADHU et al., 2021; YANG et al., 2020; HUQ; PERVIN, 2020), classificação de imagens (PATTANAIK; MITTAL; KHAN, 2020) ou para detecção e classificação de parasitas em imagens microscópicas completas (NAKASI et al., 2020a; ABDURAHMAN; FANTE; ALIY, 2021; CHIBUTA; ACAR, 2020; NAKASI; MWEBAZE; ZAWEDDE, 2021; QUINN et al., 2016; YANG et al., 2019; NAKASI et al., 2020b).

Por exemplo, em Yang et al. (2020) os autores usam o método de limiar de Otsu para filtragem de WBC em imagens em tons de cinza, triagem global mínima iterativa (IGMS) para obtenção de *patches* (gerar candidatos a parasitas localizando valores de intensidade mínima em uma imagem em tons de cinza) e CNN para classificação de *patches*. Em um conjunto de dados com 1.819 imagens microscópicas de gota espessa e 84.961 parasitas anotados, os autores relataram uma precisão de 97,26%, uma sensibilidade de 82,73%, uma especificidade de 98,39% e uma precisão de 78,98%.

Outra abordagem para detecção e classificação de parasitas em imagem completa é usar algoritmos de detecção de objetos genéricos de última geração, como *you only look once* (YOLO) (ABDURAHMAN; FANTE; ALIY, 2021; CHIBUTA; ACAR, 2020), *single short detection* (SSD) (NAKASI et al., 2020a; ABDURAHMAN; FANTE; ALIY, 2021; NAKASI; MWEBAZE; ZAWEDDE, 2021), *faster regional convolucional neural network* (Faster RCNN) (NAKASI et al., 2020a; ABDURAHMAN; FANTE; ALIY, 2021; NAKASI; MWEBAZE; ZAWEDDE, 2021; YANG et al., 2019), Retina Net (NAKASI et al., 2020a) e Refine Det (NAKASI et al., 2020a).

Em Abdurahman, Fante e Aliy (2021), os autores declararam que os modelos YOLOv3 e YOLOv4 não são otimizados para detectar objetos pequenos, como parasitas da malária, em imagens microscópicas, então eles modificaram esses modelos aumentando a escala de recursos e adicionando mais camadas de detecção para aprimorar sua capacidade de detectar objetos pequenos sem diminuir notavelmente a velocidade de detecção. Neste trabalho, os autores usaram um banco de 1.182 imagens RGB de 750×750 pixels, capturadas com um smartphone conectado à ocular do microscópio. A anotação é realizada por com caixas delimitadoras ao redor dos parasitas da malária. Os principais resultados obtidos com YOLOV4 CNN são os

seguintes: precisão média - 96,32, precisão - 95, *recall* (sensibilidade) - 94, F1-score: 94, IoU (do inglês, *Intersection over Union*) médio - 62,12 e FPS - 29,60.

Em Nakasi et al. (2020a), utilizando a RetinaNet, e usando 643 imagens de 750×750 pixels, os autores relataram uma precisão de 86,97% e um *recall* (sensibilidade) de 60,86%. Em Chibuta e Acar (2020), com YOLOv3, em um conjunto de dados de 2703 imagens de 1024×768 pixels, os autores relataram uma precisão média de 88,70%.

Nos últimos anos, com o argumento de acelerar o diagnóstico em locais remotos e com poucos recursos, alguns trabalhos têm apresentado soluções usando um dispositivo móvel acoplado à objetiva do microscópio na aquisição das imagens. As imagens microscópicas de filme espesso são então adquiridas e processadas no dispositivo móvel (NAKASI et al., 2020a; CHIBUTA; ACAR, 2020). Em outro trabalho, os autores desenvolveram uma aplicação para diagnóstico de malária baseada em um aplicativo *Web*. Neste caso, as imagens são adquiridas localmente e processadas remotamente. O resultado é então enviado de volta ao ponto de coleta de imagens (NAKASI et al., 2020b).

Nos artigos apresentados anteriormente que não usam algoritmos genéricos de detecção de objetos, observamos as seguintes tendências: a) a segmentação de objetos em primeiro plano (plaquetas, leucócitos, parasitas e alguns artefatos) é geralmente afetada pela aplicação de limiares nos histogramas da componente de saturação da imagem (DAVE; UPLA, 2017; DAVE, 2017) e limiares nos valores de nível de cinza (YANG et al., 2020; QUINN et al., 2014); b) como mostrado na introdução, em imagens microscópicas de gota espessa de sangue, os parasitas são geralmente compostos de um ou mais objetos corados. Para que um parasita não seja detectado mais de uma vez (gerando falsos positivos), os autores adotaram procedimentos como operações de fechamento morfológico (AZIF; NUGROHO; WIBIRAMA, 2018), avaliação de uma vizinhança de um objeto detectado (YANG et al., 2020), dentre outros.

Neste trabalho propomos uma abordagem para detecção de parasitas da malária com técnicas de aprendizado profundo, utilizando a classificação de *patches*, com as seguintes contribuições:

1. Segmentar objetos de primeiro plano usando classificadores de pixel, como MLP, árvores de decisão de regressão logística e componentes HSV;
2. Delimitar um *patch* contendo um parasita usando classificadores de pixel, como MLP, árvores de decisão de regressão logística e componentes do espaço de cor HSV;
3. Separar o parasita da malária de outros objetos de primeiro plano da imagem, tais como leucócitos, plaquetas e artefatos, usando uma rede neural convolutiva (CNN) treinada do zero;
4. Avaliar o desempenho da tarefa de detecção automática do parasita da malária em um conjunto de imagens microscópicas de gota espessa de sangue, contendo parasitas em diferentes fases do seu ciclo de vida.

Tabela 1 – Resumo dos trabalhos publicados na última década sobre detecção do parasita da malária em imagens de gota espessa de sangue.

Referência	Objetivo	Conjunto de dados	Pré-processamento	Ground Truth	Segmentação/ Classificação	Resultados (%)
Nakasi et al. (2020a)	Detecção de parasitas em imagem completa	643 imagens de 750x750 pixels, capturadas com um smartphone acoplado à ocular do microscópio.	Aumento de dados por meio de inversão horizontal e vertical aleatória.	Caixas delimitadoras ao redor dos parasitas da malária.	Modelos de aprendizado profundo pré-treinados: faster-R-CNN (ResNet50, ResNet101, Inception V2), RetinaNet (SSD ResNetFPN), SSD-MobileNetV2, SSD-InceptionV2.	faster-R-CNN(ResNet50): Precisão: 77,91, Sensibilidade: 89,81; RetinaNet: Precisão: 86,97, Sensibilidade: 60,86; SSD-InceptionV2: Precisão: 91,50, Sensibilidade: 37,53
Madhu et al. (2021)	Classificação de <i>patches</i>	Um total de 27.558 imagens estão disponíveis, sendo 13.779 imagens de células normais e as outras 13.779 imagens infectadas com parasitas da malária.	Redução da dimensão das imagens para 128x128 pixels.	RBC positiva (infectada) e RBC negativa (não infectada) anotados manualmente.	Redes de Cápsulas	Acurácia: 98,7±0,108, AUC-ROC: 98,73±0,105, Sensibilidade: 98,50±0,109, Especificidade: 98,97±0,189
Pattanaik, Mittal e Khan (2020)	Classificação completa da imagem	1.182 imagens RGB com 750x750 pixels, capturadas com um smartphone acoplado à ocular do microscópio.	Não aplicado	Imagens classificadas como infectadas com parasitas da malária ou não infectadas.	Rede Neural Artificial de Link Funcional (FLANN) e Autocodificador Esparso Empilhado (SSAE).	Acurácia: 89,10
Abdurahman, Fante e Aliy (2021)	Detecção de parasitas em imagem completa	1.182 imagens de 750x750 pixels, capturadas com um smartphone acoplado à ocular do microscópio.	Não aplicado	Caixas delimitadoras ao redor dos parasitas da malária.	Faster RCNN, SSD, YOLOV3 CNN e YOLOV4 CNN	YOLOV4 CNN: Precisão média: 96,32, Precisão: 95, Sensibilidade: 94, Pontuação F1: 94, IOU média: 62,12, FPS: 29,60. Faster RCNN: Precisão: 92,7; Sensibilidade: 86,9 Pontuação F1: 89,71

Tabela 2 – Resumo dos trabalhos publicados na última década sobre detecção do parasita da malária em imagens de gota espessa de sangue (continuação).

Referência	Objetivo	Conjunto de dados	Pré-processamento	Ground Truth	Segmentação/ Classificação	Resultados (%)
Chibuta e Acar (2020)	Detecção de parasitas em imagem completa	Conjunto de dados A: 2.703 imagens de 1.024x768 pixels, obtidas de uma câmera acoplada ao microscópio. Conjunto de dados B: 1.182 imagens de 750x750 pixels, capturadas com um smartphone acoplado à ocular do microscópio.	Não aplicado	Caixas delimitadoras ao redor dos parasitas da malária.	YOLOV3 CNN	Conjunto de dados A: Precisão média: 88,70. Conjunto de dados B: Precisão média: 90,02
Yang et al. (2020)	Detecção de parasitas em imagem completa	Conjunto de imagens adquiridas de 150 pacientes, incluindo 1.819 imagens de gota espessa e 84.961 parasitas anotados.	Não aplicado	Anotação manual das coordenadas x-y dos parasitas.	Método de Otsu para filtragem de leucócitos, triagem global mínima iterativa (IGMS) para segmentação de parasitas e CNN para classificação de parasitas	Acurácia: 97,26, Sensibilidade: 82,73, Especificidade: 98,39, Precisão: 78,98, AUC: 97,34
Nakasi, Mwebaze e Zawedde (2021)	Detecção de parasitas em imagem completa	903 imagens de 3264x2448 pixels.	Aumento de dados por meio de inversão vertical e horizontal aleatória.	Caixas delimitadoras desenhadas ao redor de parasitas e leucócitos.	Faster R-CNN mais rápido e SSD MobileNet	Precisão média: 55,06, Precisão: 67,20, Recall: 80,20, F1-score: 73,12
Quinn et al. (2016)	Detecção de parasitas em imagem completa	1.182 imagens RGB, capturadas com um smartphone acoplado à ocular do microscópio	Conversão de trofozoítos de malária e marcadores de WBC para o formato TF-Record.	Caixas delimitadoras ao redor dos parasitas da malária.	CNN com uma janela deslizante com o algoritmo de supressão não máxima.	AUC = 100

Tabela 2 – Resumo dos trabalhos publicados na última década sobre detecção do parasita da malária em imagens de gota espessa de sangue (continuação).

Referência	Objetivo	Conjunto de dados	Pré-processamento	Ground Truth	Segmentação/ Classificação	Resultados (%)
Aris et al. (2020)	Detecção de parasitas em imagem completa	100 imagens de 1294x980 pixels	Algoritmo de Alongamento de Contraste Global Modificado (MGCS). Normalização de cores usando o algoritmo <i>Gray world</i> .	Anotação manual das coordenadas x-y dos parasitas.	Algoritmo rápido <i>k-means</i> aplicado em componentes de cores espaciais RGB, HSV, CMYK e YUV.	Acurácia: 99,74, Especificidade: 99,76, Sensibilidade: 77,30
Yang et al. (2019)	Detecção de parasitas em imagem completa	2.967 imagens de 4.032x3.024 pixels, de 200 pacientes, capturadas com um smartphone acoplado à ocular do microscópio	Não aplicado	Anotação manual das coordenadas x-y dos parasitas.	Faster-RCNN	Sensibilidade: 96,84
Abidin, Salamah e Nugroho (2016)	Detecção de parasitas em imagem completa	50 imagens de 140x140 pixels	Filtragem de ruído (filtragem mediana, filtragem gaussiana e passa-baixa) e aprimoramento de imagem (aumento de contraste e alongamento escuro).	Anotação manual das coordenadas x-y dos parasitas.	Ferramentas de processamento de imagem digital para detecção de parasitas: contorno ativo sem borda, depois erosão, dilatação, mascaramento, alongamento de contraste e limiarização.	Acurácia: 97,57, Taxa de falso negativo: 12,04
Dave e Upla (2017)	Detecção de parasitas em imagem completa	87 campos microscópicos com 3.136x2.352 pixels. Cada campo microscópico é capturado em 10 profundidades focais diferentes	Filtragem de ruído usando valores máximos do filtro Laplaciano obtidos de 10 profundidades focais diferentes.	Anotação manual das coordenadas x-y dos parasitas.	Segmentação: limiar do histograma de saturação. Classificador: limiar de área para separar parasitas de glóbulos brancos e plaquetas.	Acurácia gota espessa: 90,81, Acurácia esfregaço: 97,87

Tabela 2 – Resumo dos trabalhos publicados na última década sobre detecção do parasita da malária em imagens de gota espessa de sangue (continuação).

Referência	Objetivo	Conjunto de dados	Pré-processamento	Ground Truth	Segmentação/ Classificação	Resultados (%)
Delahunt et al. (2015)	Detecção de parasitas em imagem completa	183 lâminas positivas e 98 lâminas negativas	Processamento de cores não detalhado.	Anotação manual das coordenadas x-y dos parasitas.	Máquinas de vetores de suporte, redes neurais e árvores de decisão com características morfológicas, de cor e textura dos candidatos.	Sensibilidade: 100, Especificidade: 95
Huq e Pervin (2020)	Classificação de <i>patches</i>	Um total de 27.558 imagens estão disponíveis, com 13.779 imagens sendo imagens de células normais e o restante das 13.779 imagens sendo infectadas com parasitas da malária.	Todas as imagens foram redimensionadas para 224x224 pixels.	Eritrócitos positivos e eritrócitos negativos anotados manualmente.	VGG16 com treinamento adversário para classificação de eritrócitos positivos e eritrócitos negativos anotados manualmente.	Acurácia: 95,9; Precisão: 96,78; Sensibilidade: 95,10; F1-score: 95,59.
Nakasi et al. (2020b)	Detecção de parasitas em imagem completa	643 imagens de 750x750 pixels capturadas com um smartphone acoplado à ocular do microscópio.	Aumento de dados: inversão horizontal e vertical aleatória de imagens selecionadas aleatoriamente.	Caixas delimitadoras ao redor dos parasitas da malária.	Faster RCNN	Acurácia: 93,06; Precisão: 74,61; Sensibilidade: 90,62.
Dave (2017)	Detecção de parasitas em imagem completa	87 imagens de campos microscópicos, com 3.136x2.352 pixels, coletadas de pacientes distintos, com variação na concentração da solução de coloração e na densidade do parasita.	Conversão HSV	Anotação manual das coordenadas x-y dos parasitas.	Segmentação: limiar do histograma de saturação. Classificador: Máquina de vetores de suporte usando 123 características: momentos centrais, momentos invariantes de Hu. Momentos de Zernike, área, etc.	Sensibilidade: 86,34, Especificidade: 96,60

Tabela 2 – Resumo dos trabalhos publicados na última década sobre detecção do parasita da malária em imagens de gota espessa de sangue (continuação).

Referência	Objetivo	Conjunto de dados	Pré-processamento	Ground Truth	Segmentação/ Classificação	Resultados (%)
Azif, Nugroho e Wibirama (2018)	Detecção de parasitas em imagem completa	36 imagens de campos microscópicos, com 1600x1200 pixels. O número médio de parasitas <i>Plasmodium</i> em cada imagem foi de 15 parasitas	Alongamento de contraste e filtragem mediana.	Anotação manual das coordenadas x-y dos parasitas.	Os parasitas são detectados usando uma operação morfológica <i>bottom-hat</i> , limiar de entropia adaptável e métodos de contorno activo sem borda.	Sensibilidade: 98,04, Acurácia: 61,53

2.2 Materiais e Métodos

2.2.1 Materiais

O banco de dados utilizado consiste de quatro conjuntos de 169 imagens, totalizando 676 imagens de tamanho 1.388×1.040 pixels de lâminas de gota espessa. Cada lâmina corresponde a um paciente diferente. A aquisição das imagens foi realizada no Laboratório de Micobacteriologia do Instituto Nacional de Pesquisa da Amazônia – INPA, Brasil.

As imagens foram capturadas no microscópio ZEISS Axio Imager M2, com o software ZEN 2 Pro com aumento de $1000\times$, de quatro lâminas contendo *Plasmodium vivax* fornecidas pela Fundação de Medicina Tropical, Dr. Heitor Vieira Dourado. Cada conjunto de imagens foi obtido de um paciente diferente. A Tabela 3 mostra a média do número de parasitas por imagem em cada conjunto. Conforme mostrado na Tabela 2, a densidade parasitária nos conjuntos 1 e 2 é maior do que nos conjuntos 3 e 4. O paciente correspondente ao conjunto 1, com média de 18,85 parasitas por lâmina, corresponde ao maior nível de infecção, enquanto o paciente correspondente ao conjunto 4, com média de 5,21 parasitas por lâmina, corresponde ao menor nível de infecção. A Figura 5 mostra histogramas do número de parasitas por imagem em cada conjunto. Conforme mostrado, o conjunto com maior densidade parasitária é o conjunto 1, enquanto o com menor densidade parasitária é o conjunto 4.

Tabela 3 – Número de parasitas em cada conjunto.

Estatística	Conjunto de imagens			
	1	2	3	4
Média	18,85	14,18	12,76	5,21
Desvio padrão	4,99	4,98	4,27	3,47

2.2.2 Métodos

A Figura 6 mostra um diagrama de blocos da abordagem adotada neste trabalho para detectar parasitas da malária em lâminas de gota espessa de sangue coradas pelo método Giemsa. O método proposto usa aprendizado profundo para classificação de *patches* (sub-imagens obtidas das imagens dos campos microscópicos) em *patches* positivos (infectados) e negativos (não infectados).

Um problema que surge na detecção de parasitas é que, não raramente, eles são compostos de mais de um objeto corado. Assim, é necessário agrupar estes objetos em uma único *patch*, para não gerar falsos positivos. Neste trabalho, um classificador de pixels usa os componentes do modelo de cor HSV para detectar a região próxima aos objetos do primeiro plano. Embora o processo de desglobalização quebre os RBCs, uma região com uma tonalidade diferente do fundo ainda envolve os parasitas da malária. A Figura 7 mostra uma imagem microscópica de gota espessa de sangue convertida para o modelo de cor HSV mostrando os objetos em primeiro

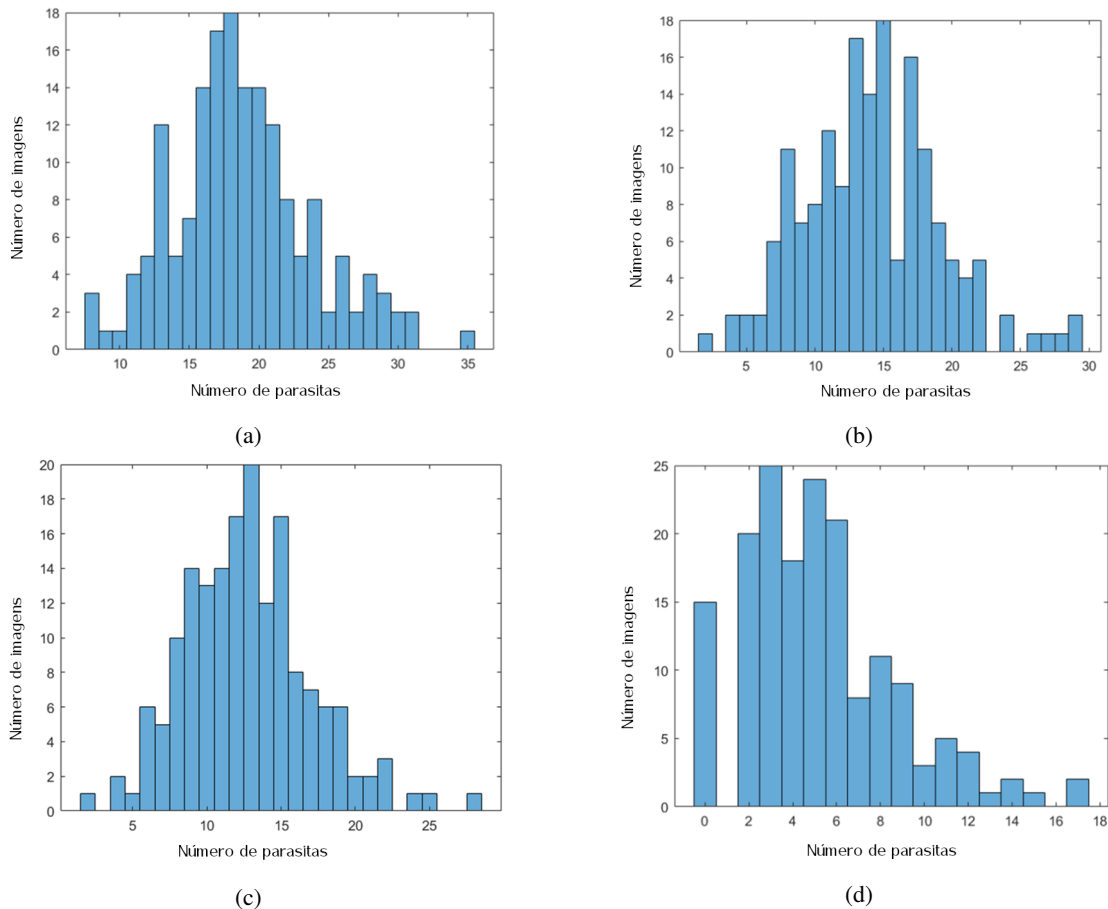


Figura 5 – Histogramas do número de parasitas por imagem em cada conjunto. (a) Conjunto 1. (b) Conjunto 2. (c) Conjunto 3. (d) Conjunto 4.

plano envolvidos em regiões com tonalidade (componente H) diferente daquela do fundo da imagem. Como visto, uma região, com características de HSV diferentes do fundo, envolve os parasitas e leucócitos.

No método proposto, mostrado na Figura 6, os *patches* são formados após um processo de fusão a partir dos resultados dos dois classificadores de pixels: classificador de objetos em primeiro plano e classificador de regiões próximas a objetos em primeiro plano. Nesse processo de fusão, uma intersecção entre os objetos em primeiro plano e as regiões detectadas pelo segundo classificador é realizada.

No desenvolvimento deste trabalho, observou-se que o tamanho dos *patches* contendo parasitas varia muito. Normalizar todos os *patches* para o mesmo tamanho causa muita distorção. Assim, foram treinadas duas CNNs distintas de classificação de *patches*. A primeira foi treinada com *patches* menores que 1.600 pixels, que correspondem a trofozoítos imaturos. A segunda, com *patches* maiores que 1.600 pixels, que correspondem a trofozoítos ameboides ou maduros, esquizontes e gametócitos, bem como a leucócitos e plaquetas, que geralmente são maiores que parasitas (DAVE; UPLA, 2017).

A seguir, apresenta-se cada bloco mostrado na Figura 6.

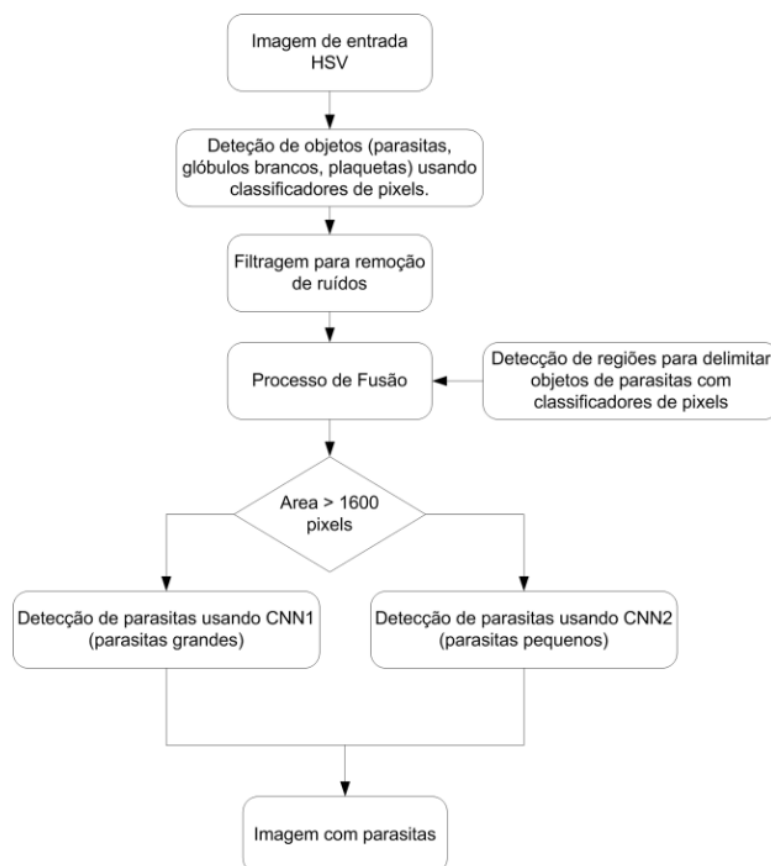


Figura 6 – Fluxograma da proposta para detecção de parasitas da malária em imagens microscópicas de gota espessa.

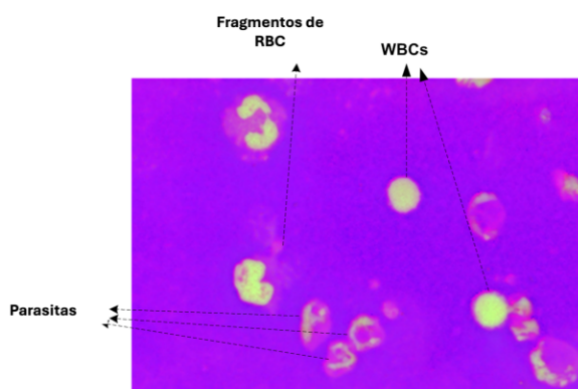


Figura 7 – Apresentação de uma imagem microscópica de gota espessa de sangue utilizando o modelo de cor HSV mostrando os objetos em primeiro plano (WBCs, RBCs, parasitas) envolvidos em regiões com tonalidade (componente H) diferentes daquela do fundo da imagem. Além disso, verifica-se fragmentos de células RBC na imagem.

2.2.2.1 Detecção de objetos em primeiro plano

A Figura 8 mostra os valores das componentes do espaço de cor HSV dos pixels encontrados em uma imagem microscópica de gota espessa: pixels de objetos em primeiro plano, pixels de regiões próximas a objetos em primeiro plano e fragmentos de RBCs, e pixels de fundo. Para tal, extraímos informações de 10.000 pixels de objetos em primeiro plano (5.000 de parasitas e 5.000 de WBCs e plaquetas), 5.000 pixels de regiões próximas a objetos em primeiro plano e

5.000 pixels de fundo de um subconjunto de 5 imagens. Conforme mostrado na Figura 8, os pixels em primeiro plano têm alta saturação (componente S) e baixa intensidade (componente V). Os pixels em segundo plano têm alta intensidade, baixa saturação e baixa tonalidade. Os pixels de regiões próximas aos objetos em primeiro plano têm alta tonalidade e baixa saturação.

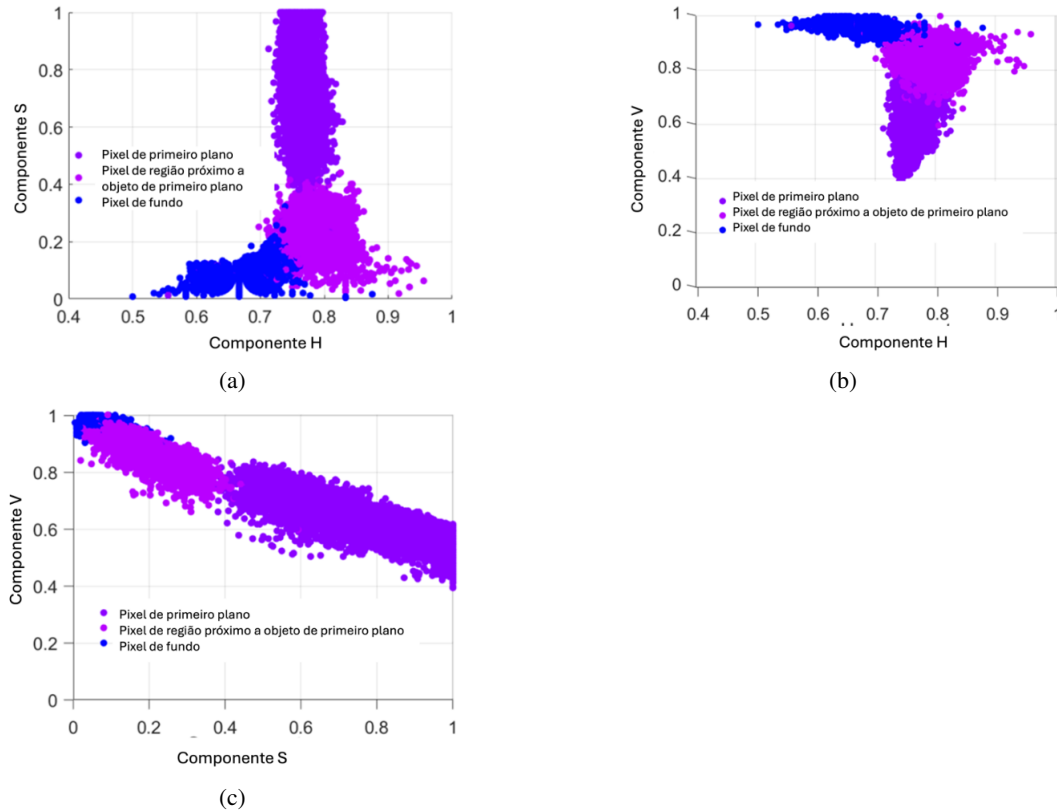


Figura 8 – Distribuição, em função das componentes do espaço HSV, dos valores dos pixels extraídos de objetos em primeiro plano, pixels de regiões próximas a objetos em primeiro plano e pixels pertencentes ao fundo: (a) saturação versus tonalidade (S x H); (b) intensidade versus tonalidade (V x H); (c) intensidade versus saturação (V x S).

Para segmentar os objetos em primeiro plano, avaliamos dois classificadores binários, um perceptron multicamadas (MLP, do inglês *Multi Layer Perceptron*) e uma árvore de decisão (DT, do inglês *Decision Tree*). O desempenho de ambos classificadores é mostrado na seção de resultados. A entrada de ambos os classificadores são os componentes H, S e V do espaço de cores HSV.

2.2.2.2 Detecção de regiões para delimitar os parasitas

A Figura 9 mostra a imagem HSV de um *patch* contendo um parasita, e suas respectivas imagens componentes H, S e V. Observa-se que tais imagens componentes estão divididas em três regiões, a da esquerda e a da direita são regiões pertencentes ao fundo da imagem. Já a parte central engloba uma área próxima ao parasita e o próprio parasita. Por fim é mostrado o perfil de intensidade das três componentes H, S e V dos pixels de uma linha horizontal que passa na região central da imagem (linha horizontal preta). Tais perfis denotam que a parte central

(parasita e região próxima a ele) apresentam baixa intensidade (V) e saturação (S) e tonalidades (H) altas, quando comparados aos pixels de fundo.

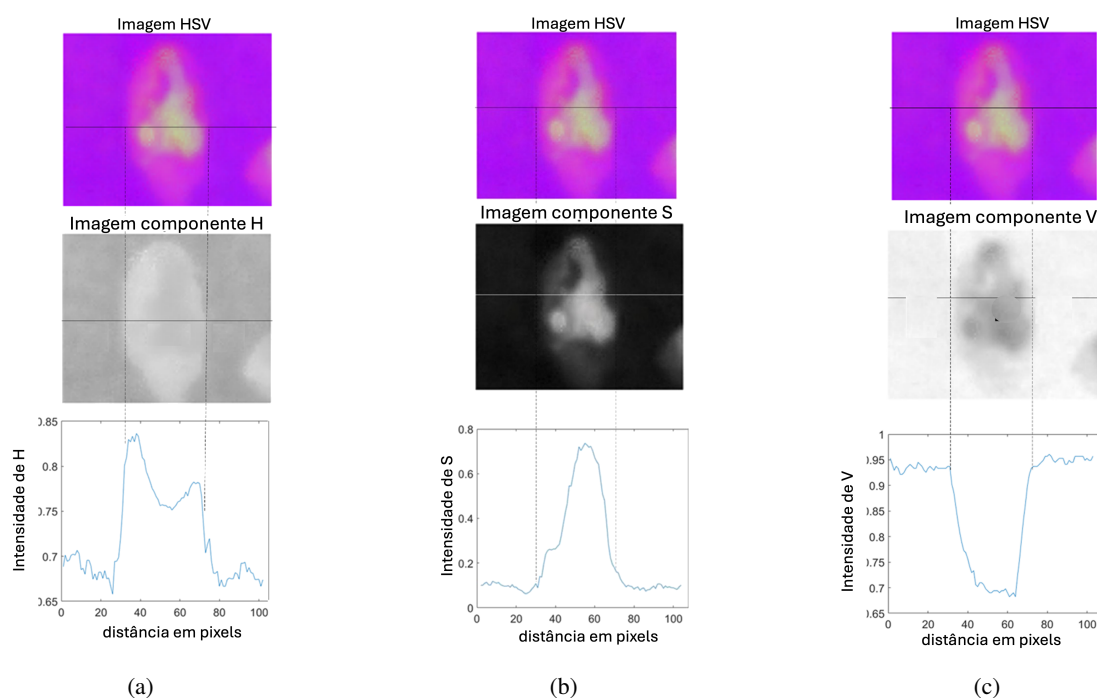


Figura 9 – Perfis dos componentes H, S e V dos pixels de uma região próxima a um parasita e o próprio parasita, versus o fundo. (a) Perfil H. (b) Perfil S. (c) Perfil V.

Para a tarefa de delimitação das regiões com os parasitas, também avaliamos dois classificadores binários, um perceptron multicamadas (MLP) e uma árvore de decisão (DT). A entrada de ambos os classificadores são os componentes H, S e V do espaço de cores HSV. Os resultados serão apresentados na seção de resultados.

2.2.2.3 Processo de fusão

O processo de fusão consiste em criar a intersecção dos objetos em primeiro plano com as regiões detectadas perto deles. Apenas as regiões que cruzam um objeto em primeiro plano são mantidas. A Figura 10 ilustra esse procedimento. A Figura 10(a) mostra uma imagem microscópica de gota espessa de sangue. A Figura 10(b) mostra os objetos em primeiro plano detectados usando o classificador descrito na seção 2.2.2.1 - Detecção de objetos em primeiro plano. A Figura 10(c) mostra as regiões detectadas perto dos objetos em primeiro plano usando o classificador descrito na seção 2.2.2.2 - Detecção de regiões para delimitar os parasitas. A Figura 10(d) mostra a imagem de fusão. A Figura 10(e) mostra os *patches* na imagem original correspondentes às regiões da imagem de fusão. Esses *patches* são aqueles submetidos a uma CNN que separa os parasitas da malária de outros objetos em primeiro plano. A Figura 10(f) mostra alguns *patches* ampliados. Como mostrado, embora alguns parasitas tenham mais de uma região corada, todos eles estão incluídos no mesmo *patch*, minimizando falsos positivos. O *patch* na linha inferior e na coluna da direita é um WBC.

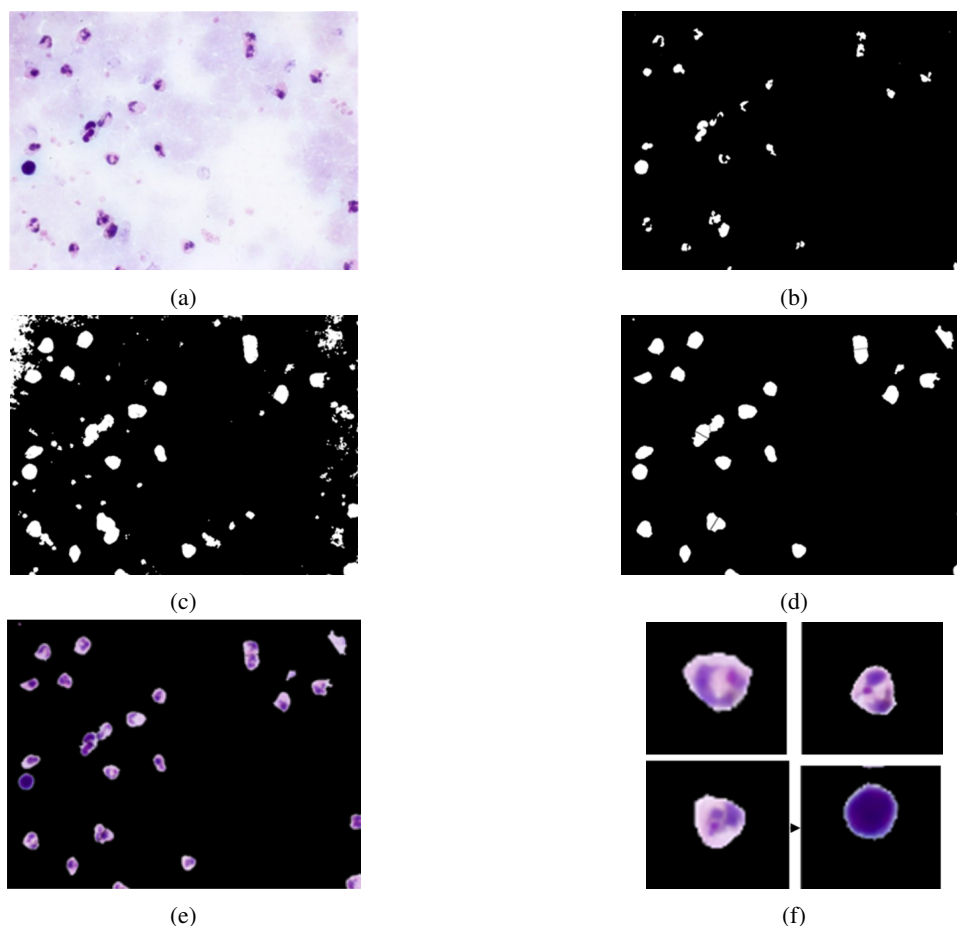


Figura 10 – (a) Uma imagem de gota espessa. (b) Resultado da detecção em primeiro plano. (c) Resultado da extração de regiões próximas a objetos em primeiro plano. (d) Imagem de fusão. (e) *Patches* correspondentes na imagem original. (f) Alguns *patches* ampliados.

2.2.2.4 Detecção de parasitas

Os *patches* mostrados na Figura 10(e), correspondentes a objetos em primeiro plano, são submetidos a uma CNN que discrimina parasitas da malária de leucócitos, plaquetas e ruído. Os *patches* para treinamento das CNNs foram extraídos de 564 imagens, enquanto 92 imagens (23 imagens de cada conjunto de imagens) foram usadas para avaliar o método de detecção de parasitas em imagens completas.

O tamanho dos *patches* varia. Existem *patches* com áreas menores que 700 pixels e *patches* com áreas maiores que 3.500 pixels. A Figura 11 mostra gráficos do número de *patches* com áreas menores que 1.600 pixels (40×40 pixels) e com áreas maiores que 1.600 pixels em cada conjunto de imagens. A Tabela 4 mostra o valor médio da área do *patch* em todos os quatro conjuntos de imagens estudados. Conforme observado, nos conjuntos 1 e 2, há *patches* maiores, enquanto nos conjuntos 3 e 4 há *patches* menores. Os conjuntos 1 e 2 correspondem a imagens de gota espessa de pacientes contendo mais parasitas em estágios avançados de desenvolvimento (esquizontes e gametócitos), com tamanhos maiores, enquanto os conjuntos 3 e 4 correspondem a imagens de gota espessa de pacientes contendo mais parasitas em estágios menos avançados de desenvolvimento (trofozoítos em anel ou imaturos). A escolha desses conjuntos de imagens

permite comparar o desempenho das CNNs com *patches* maiores e menores.

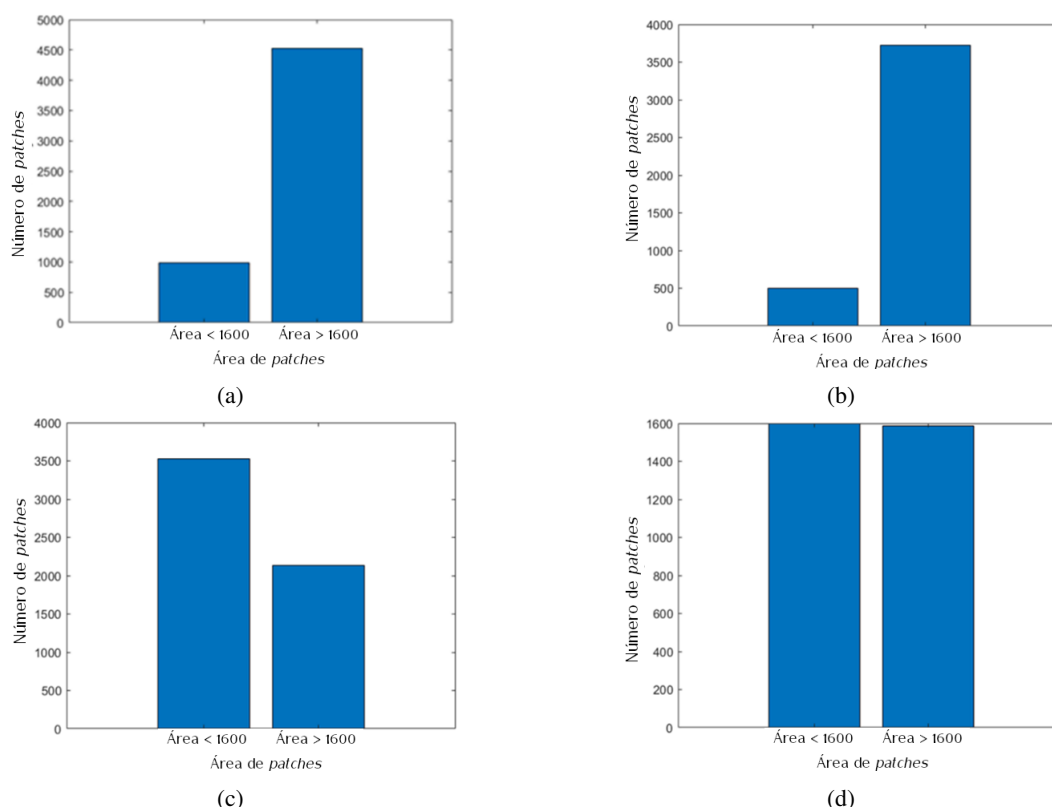


Figura 11 – Número de *patches* com área menor que 1.600 pixels (40×40) e com área maior que 1.600 pixels em cada um dos conjuntos de imagens: (a) conjunto 1; (b) conjunto 2; (c) conjunto 3; (d) conjunto 4.

Tabela 4 – Área do *patch* por conjunto de imagens (em número de pixels).

Conjunto	Área do <i>patch</i>	
	Média	Desvio padrão
1	2.481,7	1.691,4
2	2.801,9	1.478,7
3	1.487,2	3.191,0
4	1.920,9	2.351,0

Para a entrada das CNNs, foi necessário padronizar os tamanhos dos *patches*. Adotamos duas abordagens de padronização. Na primeira, um *patch* com área <1600 pixels é redimensionado para 40×40 pixels e um *patch* com área >1600 pixels é redimensionado para 100×100 pixels. Para classificar os *patches* menores, foi usada uma rede CNN1. Para classificar os *patches* maiores, foi usada uma rede CNN2. A diferença entre CNN1 e CNN2 está apenas na primeira camada, que recebe imagens de tamanhos diferentes. Na segunda abordagem, todos os *patches* são redimensionados para 100×100 pixels, e a rede CNN2 é usada. O motivo para avaliar dois tamanhos de *patches* diferentes é que percebemos que quando *patches* menores são redimensionados para 100×100 pixels, eles ficam muito distorcidos, então, visamos comparar o desempenho com diferentes tamanhos de *patches*.

O número de *patches* positivos (contendo parasita) e negativos (sem parasita) usados para treinamento e validação da CNN1 (*patches* pequenos) e CNN2 (*patches* grandes) é mostrado na Tabela 5. Oitenta e cinco por cento dos *patches* foram usados para treinamento, enquanto 15%, para validação.

Tabela 5 – Número de *patches* para treinamento e validação.

Tamanho	Tipo	
	Positivo	Negativo
Pequeno	5.345	5.702
Grande	4.778	4.425

A arquitetura da CNN1 e CNN2 é mostrada na Figura 12. Há 34 camadas e 4 blocos de extração de características, compostos de camadas CONV-BATCH-RELU-CONV-BATCH-RELU. Em todos os blocos, o tamanho do *kernel* foi ajustado para [3,3]. A profundidade de cada camada CONV no bloco 1, bloco 2, bloco 3 e bloco 4 são 64, 128, 256 e 512, respectivamente. De acordo com Aggarwal (2021), a camada de normalização de lote é um método recente para resolver os problemas de desaparecimento e explosão de gradiente, que causam redução ou aumento da magnitude dos gradientes de ativação nas camadas sucessivas. Além disso, essa camada pode reduzir o efeito do deslocamento de covariáveis internas, que afeta o treinamento de redes neurais profundas.

Para reduzir a representação do mapa, três camadas de *maxpooling* são inseridas entre os blocos de extração de características. As camadas de *maxpooling* extraem valores máximos de uma área de 2×2 pixels, com um passo de [2,2]. Após cada operação de *maxpooling*, a representação do mapa é dividida por 2. No final do quarto bloco de características de extração, para melhorar a generalização, há uma camada de *dropout*, com um fator de 0,5.

Vários métodos de otimização foram avaliados: propagação da raiz quadrada média (RMSProp), descida de gradiente estocástico com momento (SGDM) e ADAM. Os melhores resultados foram obtidos com o método ADAM. O número de épocas foi fixado em 100. Com o método ADAM, uma taxa de aprendizado inicial de 3×10^{-4} e um fator de decaimento do gradiente quadrado de 0,99 foram empregados. O tamanho do *minibatch* foi 128.

2.2.2.5 Configuração experimental

As simulações foram feitas no ambiente MATLAB, em um computador com processador Intel Core i5-9300H CPU 2.40GHz, 8GB RAM, unidade gráfica processador NVIDIA GeForce GTX 1650 (4GB RAM).

2.3 Resultados

Nesta seção, são apresentados quatro tipos de resultados: resultados para detecção de objetos em primeiro plano, resultados para detecção de regiões próximas a objetos em primeiro plano,

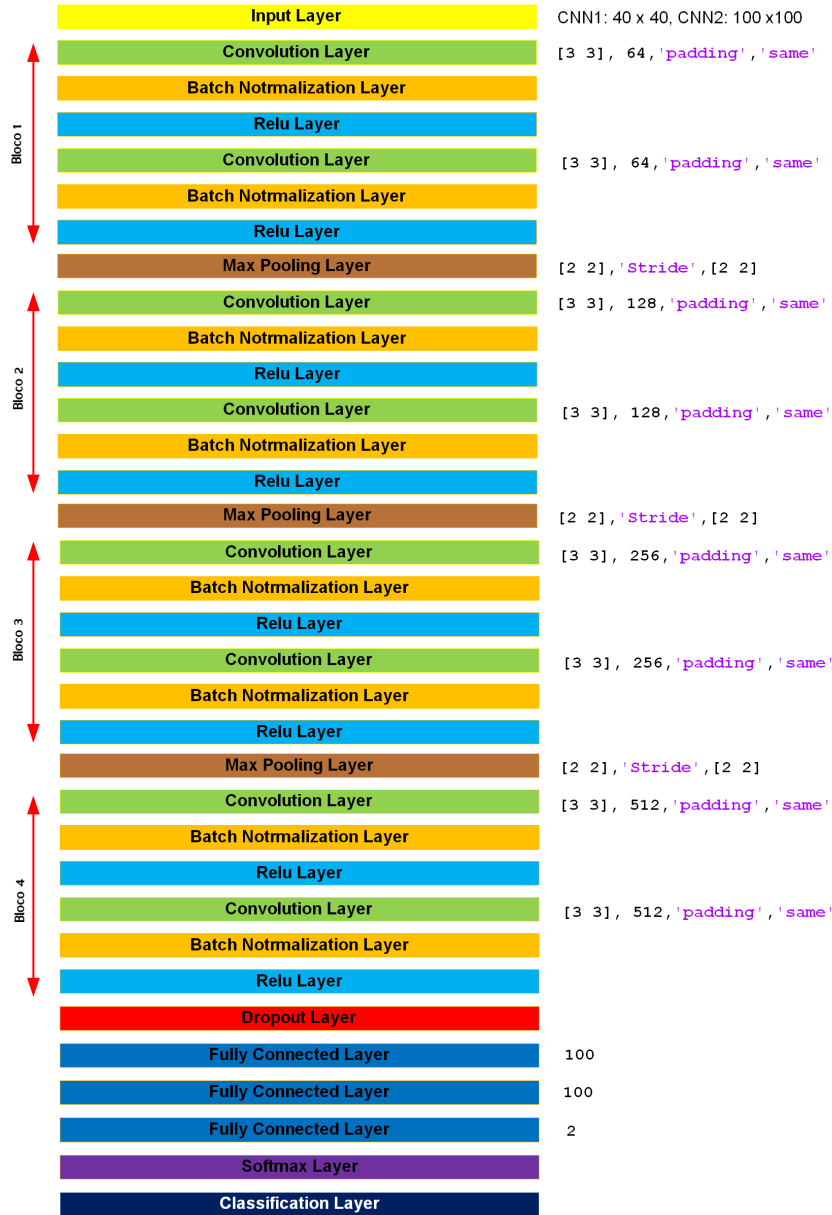


Figura 12 – Arquitetura da CNN1 e CNN2 com 34 camadas e 4 blocos de extração de características.

resultados para classificação de *patches* de parasitas e, finalmente, resultados para detecção de parasitas da malária em imagens completas.

2.3.1 Resultados para detecção de objetos em primeiro plano

Para detectar objetos em primeiro plano, a arquitetura do MLP foi $3 \times 4 \times 1$. Os dados foram divididos em 70%, 15%, 15% para treinamento, validação e teste, respectivamente. O critério de parada foi 16 épocas. A Figura 13 mostra o desempenho do erro médio quadrático durante o treinamento. Foi obtida uma perda de 0,6% com validação cruzada de 10 partes. Em relação ao classificador de árvore, simulamos árvores com diferentes valores de tamanho mínimo de folha. A Figura 14(a) mostra o erro de classificação como uma função do tamanho mínimo de folha. Um tamanho mínimo de folha de 45 foi usado na árvore ótima. A árvore ótima é mostrada

na Figura 14(b) (a variável x_2 corresponde ao componente S do espaço HSV). Foi obtida uma perda de 0,78% com validação cruzada de 10 partes. O classificador MLP foi selecionado para segmentar os objetos de fundo, pois apresenta o menor valor de perda.

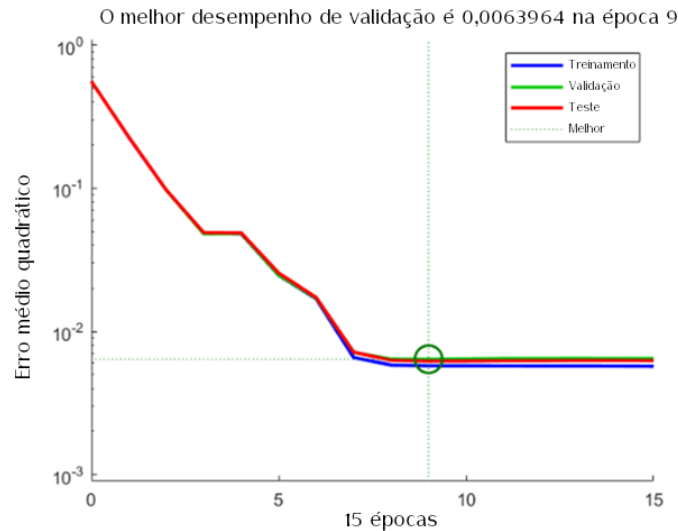


Figura 13 – Desempenho do erro médio quadrático durante o treinamento do MLP para classificação de objetos em primeiro plano.

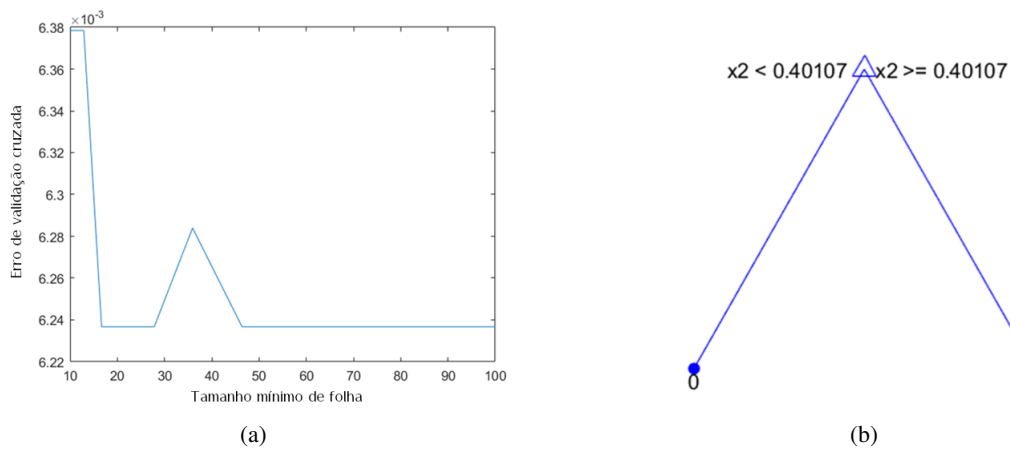


Figura 14 – Árvore de decisão para classificação de objetos em primeiro plano. (b) arquitetura da árvore (a variável x_2 corresponde ao componente S) (a) Erro de validação cruzada de 10 partes como função do tamanho mínimo da folha. (b) Arquitetura da árvore (a variável x_2 corresponde ao componente S).

2.3.2 Resultados para detecção das regiões próximas a objetos em primeiro plano

Para detectar as regiões próximas aos objetos em primeiro plano, a arquitetura MLP também foi $3 \times 4 \times 1$. Os dados também foram divididos em 70%, 15%, 15% para treinamento, validação e teste, respectivamente. O critério de parada foi de 15 épocas. Uma perda de 0,95% com validação cruzada de 10 partes foi obtida no conjunto de teste. Em relação ao classificador de árvore, simulamos árvores com diferentes valores de tamanho mínimo de folha. Um tamanho

mínimo de folha de 45 foi usado na árvore ótima. Foi obtida uma perda de 1,33% com validação cruzada de 10 partes. O classificador MLP foi selecionado para segmentar as regiões próximas aos objetos em primeiro plano, pois apresenta o menor valor de perda.

Às vezes, quando há dois parasitas próximos um do outro, como mostrado na Figura 15, as regiões próximas a cada parasita são unidas. Dois procedimentos são aplicados para resolver esse problema. O primeiro é uma operação de *watershed*. Se, após o *watershed*, o tamanho da região for maior que 4.000 pixels, nós a dividimos em duas partes iguais. Na Figura 15(a) há dois parasitas, um na parte superior e outro na parte inferior. Após o classificador, apenas uma região é detectada, como mostrado na Figura 15(b). A separação de uma região em duas regiões é feita com o procedimento *watershed*. O resultado é mostrado na Figura 15(c).

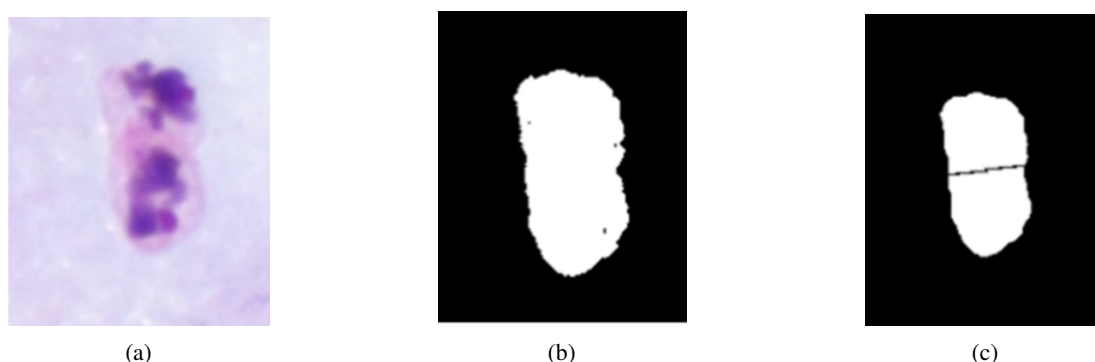


Figura 15 – (a) Imagem de dois parasitas próximos um do outro. (b) Regiões segmentadas próximas aos dois parasitas unidos. (c) Regiões separadas usando operação de *watershed*.

2.3.3 Resultados para classificação de *patches* de parasitas

A Figura 16 mostra alguns exemplos de *patches* positivos e negativos de parasitas detectados. Como visto, um parasita pode ter mais de um objeto corado. Felizmente, pelo método proposto neste trabalho, todos eles estão incluídos em apenas um *patch*.

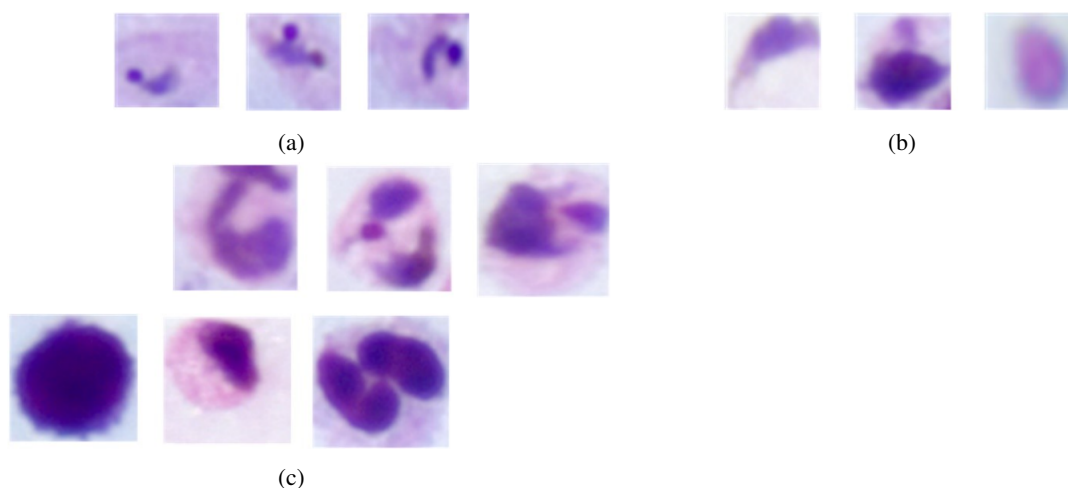


Figura 16 – Imagens de *patches* extraídos. (a) Pequenos *patches* positivos. (b) Pequenos *patches* negativos. (c) Grandes *patches* negativos.

A Tabela 6 mostra os resultados para classificar *patches* de parasitas. Três resultados são mostrados: CNN1 + *patches* pequenos com tamanho 40x40 pixels, CNN2 + *patches* grandes com tamanho 100x100 pixels e CNN2 + todos os *patches* com tamanho 100x100 pixels. Um processo de validação cruzada de 10 partes foi usado. As seguintes métricas são calculadas: acurácia (acur), sensibilidade (sens), especificidade (espe), precisão (prec), F1-score e área sob a curva ROC (AUC), juntamente com o desvio padrão de cada uma. Como visto na Tabela 6, os resultados das duas primeiras linhas, com CNNs dedicadas a *patches* pequenos e grandes, são melhores do que os resultados quando todas os *patches* são convertidos para 100x100 pixels e classificados com a CNN2. Portanto, o método proposto usa as duas CNNs, CNN1 e CNN2, para classificação de *patches*.

Tabela 6 – Resultados para classificação de *patches* de parasitas (Média±DP).

Rede	Patches	Acur. (%)	Sens. (%)	Espe. (%)	Prec. (%)	F1-score (%)	AUC
CNN1	Pequenos	90,10±0,20	85,78±0,17	96,01±0,13	85,88±0,19	87,89±0,14	0,961±0,12
CNN2	Grandes	88,91±0,26	87,75±0,16	89,90±0,22	87,82±0,11	87,90±0,2	0,952±0,16
CNN2	Todos os <i>patches</i>	87,80±0,15	76,79±0,21	95,91±0,25	76,84±0,15	87,91±0,17	0,939±0,11

2.3.4 Resultados para classificação de parasitas em imagens completas

Oitenta imagens dos 4 conjuntos de imagens, 20 imagens por conjunto, foram usadas para avaliar a detecção de parasitas em imagens completas. Um processo de validação cruzada de 10 partes foi usado. Cada parte consiste de oito imagens, duas de cada conjunto (duas imagens do conjunto 1, duas imagens do conjunto 2, duas imagens do conjunto 3 e duas imagens do conjunto 4). As seguintes métricas foram avaliadas: acurácia (acur), sensibilidade (sens), especificidade (espe), precisão (prec) e F1-score, juntamente com o desvio padrão de cada uma. A Tabela 7 mostra esses valores.

Tabela 7 – Resultados da classificação de parasita na imagem completa (Média±DP) nos diversos conjuntos de imagens microscópicas de gota espessa.

Conjunto	Acur. (%)	Sens. (%)	Espe. (%)	Prec. (%)	F1-score (%)
1	91,59±0,79	77,23±3,17	97,21±0,57	91,71±1,40	83,81±1,85
2	92,66±0,61	80,93±1,92	97,53±0,37	93,14±1,23	86,59±1,36
3	87,53±1,45	55,46±5,05	95,70±0,41	76,58±2,05	64,25±3,88
4	91,52±1,19	65,55±2,23	95,71±1,11	71,55±4,36	68,33±2,14

O tempo necessário para avaliar uma imagem de 1.044×1.388 pixels foi de aproximadamente 6s. Como um diagnóstico de malária requer avaliação de 160 imagens, isso implica um tempo de 16 minutos.

2.4 Discussão

Este trabalho apresenta um método para detectar parasitas da malária usando CNNs para classificação de *patches*. Podemos dividi-lo em quatro etapas: a detecção de objetos de primeiro plano, a delimitação dos *patches*, a classificação dos *patches* e, por fim, a detecção dos parasitas da malária em uma imagem de campo microscópico de esfregaço de gota espessa.

No método proposto, objetos em primeiro plano são detectados usando classificadores de pixels. Como mencionado anteriormente, em outros trabalhos da literatura, a detecção em primeiro plano é geralmente efetuada aplicando limiares nos histogramas de canais de saturação (DAVE; UPLA, 2017; DAVE, 2017) e limiares nos valores de nível de cinza (YANG et al., 2020; QUINN et al., 2014). Este trabalho avalia dois classificadores: o MLP e uma árvore de decisão, usando os componentes do espaço de cor HSV como dados de entrada. Embora a árvore de decisão tenha tido um desempenho ótimo com um limiar simples aplicado no componente S, coincidindo com o proposto em (DAVE; UPLA, 2017; DAVE, 2017), o melhor desempenho foi obtido com o MLP, usando os três componentes do espaço HSV.

Sobre a delimitação de *patches*, também propomos classificadores de pixels usando componentes do espaço HSV para obter uma região próxima aos objetos em primeiro plano. Em outros trabalhos os autores adotaram procedimentos como operações de fechamento morfológico (AZIF; NUGROHO; WIBIRAMA, 2018) ou avaliação de uma vizinhança de um objeto detectado (YANG et al., 2020). Provavelmente o método proposto neste trabalho contribuiu para obter uma baixa taxa de falsos positivos, o que se reflete nos altos valores de especificidade obtidos na Tabela 7.

Em relação à classificação de *patches*, a abordagem adotada de classificar *patches* pequenos e *patches* grandes com diferentes CNNs mostrou resultados interessantes. A Tabela 6 mostra que a CNN1, treinada com *patches* pequenos (área < 1600 pixels), e a CNN2, treinada com *patches* grandes (área > 1600 pixels), obtiveram AUC de 0,961 e 0,952, respectivamente. Esses valores de AUC são melhores do que o valor de 0,939, obtido com a CNN2 treinada com todos os *patches* redimensionados para 100x100 pixels. Acreditamos que redimensionar *patches* pequenos para 100x100 pixels gera muita distorção, tornando o reconhecimento de CNN mais difícil.

A Tabela 7 mostra que a sensibilidade nos conjuntos 3 e 4 é menor do que a sensibilidade nos conjuntos 1 e 2. As seguintes razões explicam esse fato: 1) A sensibilidade da CNN1, treinada com *patches* pequenos, é menor do que a sensibilidade da CNN2, treinado com *patches* grandes; 2) os conjuntos 1 e 2, como mostrado na Figura 11, têm mais *patches* grandes do que *patches* pequenos, enquanto os conjuntos 3 e 4 têm mais *patches* pequenos. *Patches* pequenos são mais difíceis de classificar porque podem ser confundidos com ruído.

Este trabalho apresentou dois tipos de resultados: resultados referentes a classificações de *patches* e resultados referentes a classificações de parasitas em imagens completas. Os resultados publicados anteriormente na literatura apresentaram valores para as métricas diferentes, provavelmente, porque utilizaram conjuntos de dados diferentes. Por exemplo, quanto à classificação de parasitas em imagens completas, em Abdurahman, Fante e Aliy (2021), Nakasi,

Mwebaze e Zawedde (2021), dois trabalhos desenvolvidos utilizando o estado da arte de algoritmos de detecção de objetos genéricos, *faster regional convolutional neural network* (*Faster RCNN*), obtiveram precisões de 92,7% e 67,2%, respectivamente. No primeiro, foi utilizado um conjunto de dados com 1.182 imagens de 750×750 pixels, enquanto no segundo foi utilizado um banco de dados com 903 imagens de 3.264×2.448 pixels. Outro exemplo de detecção de parasitas em imagem completa, adotando a abordagem utilizada neste trabalho, determinação de *patches* e classificação com aprendizado profundo, é encontrado em Yang et al. (2020). Neste artigo, os autores obtiveram uma precisão de 78,98%. Em Chibuta e Acar (2020), também utilizando o estado da arte de algoritmos de detecção de objetos genéricos, uma rede *you only look once* (YOLO), os autores obtiveram precisões médias entre 64,5% e 83,2% (CHIBUTA; ACAR, 2020). Em nosso trabalho, foram obtidos valores de precisão variando entre 71,55% e 91,71%. Os menores valores são obtidos com imagens contendo parasitas menores.

Em relação aos valores de sensibilidades e especificidades, em Aris et al. (2020) os autores, utilizando o algoritmo rápido *k-means*, obtiveram resultados semelhantes aos obtidos neste trabalho: altos valores de especificidade, e valores medianos de sensibilidade, na faixa entre 55,28% e 79,71%. Neste trabalho as sensibilidades variaram entre 55,46% e 80,93%. Os menores valores são obtidos com imagens contendo parasitas menores. Em Yang et al. (2020), utilizando técnicas de aprendizado profundo, como *Faster RCNN* e YOLO, os autores concluíram que essas redes de detecção de objetos não funcionam bem para objetos muito pequenos como parasitas, com tamanho médio de 44×44 pixels, resultando em muitos falsos negativos. Assim, pode-se afirmar que o método apresentado neste trabalho para detecção de parasitas da malária tem métricas de precisão, sensibilidade e especificidade comparáveis aos métodos do estado da arte para detecção de objetos.

3 MÉTODO AUTOMÁTICO PARA DETECÇÃO DE ÁREAS AFETADAS PELA COVID-19

3.1 Revisão da literatura

Neste trabalho foi realizada uma pesquisa bibliográfica sobre como o aprendizado de máquina profundo estava sendo empregado em tarefas de segmentação de infecções pulmonares causadas pela COVID-19 nos últimos quatro anos. As bases de dados selecionadas foram PubMed, IEEE, ScienceDirect e ACM e a expressão geral usada na consulta foi: ((machine learning) AND (deep learning)) AND ((image segmentation) AND (computerized tomography)) AND ((lung infection) AND (Coronavirus OR COVID-19)), com as adaptações apropriadas para cada base. O número total de artigos retornados sem repetições foi 56.

Na fase de triagem, 31 artigos foram mantidos e 25 foram excluídos, de acordo com os seguintes critérios estabelecidos: 1. O artigo foi incluído se realizasse treinamento de redes neurais convolutivas para segmentar a COVID-19; 2. O artigo foi incluído se citasse as bases de dados utilizadas; 3. O artigo foi incluído se citasse as divisões utilizadas nas bases de dados; 4. O artigo foi incluído se citasse as redes utilizadas para realizar a(s) tarefa(s) de segmentação; 5. O artigo foi incluído se citasse a métrica Dice.

Na fase de análise, 22 artigos foram mantidos e 9 foram excluídos. Os critérios estabelecidos foram: 1. O artigo foi incluído se citasse a arquitetura de rede usada para segmentar os pulmões, COVID-19 ou ambas; 2. O artigo foi incluído se citasse o tipo de método adotado, automático ou iterativo; 3. O artigo foi incluído se tivesse como objetivo treinar um modelo para segmentar a COVID-19.

Após a análise dos artigos, foi observado que os estudos podem ser classificados de acordo com o número de etapas utilizadas no processo de segmentação e de acordo com a metodologia utilizada para particionamento dos dados. Com relação ao número de etapas usadas no processo de segmentação, foram identificadas duas abordagens na literatura. Na primeira abordagem, mostrada na Figura 17(a), a segmentação da COVID-19 é realizada em uma única etapa. Esta abordagem é referida neste trabalho como segmentação de uma etapa. Na segunda abordagem, mostrada na Figura 17(b), a segmentação da área afetada pela COVID-19 é precedida por uma etapa de segmentação pulmonar. Esta abordagem é referida neste trabalho como segmentação de duas etapas.

Com relação à metodologia usada para particionamento de dados (PD), também foram identificadas duas estratégias. Na primeira estratégia, os dados são divididos por tomografias computadorizadas completas, que são usadas exclusivamente para treinar ou para testar o classificador. Na segunda estratégia, os dados são divididos por *slices*, onde uma porcentagem dos *slices* de uma tomografia computadorizada são usados para treinar, e a porcentagem dos *slices* restantes, para testar o classificador. A primeira estratégia é chamada neste trabalho de

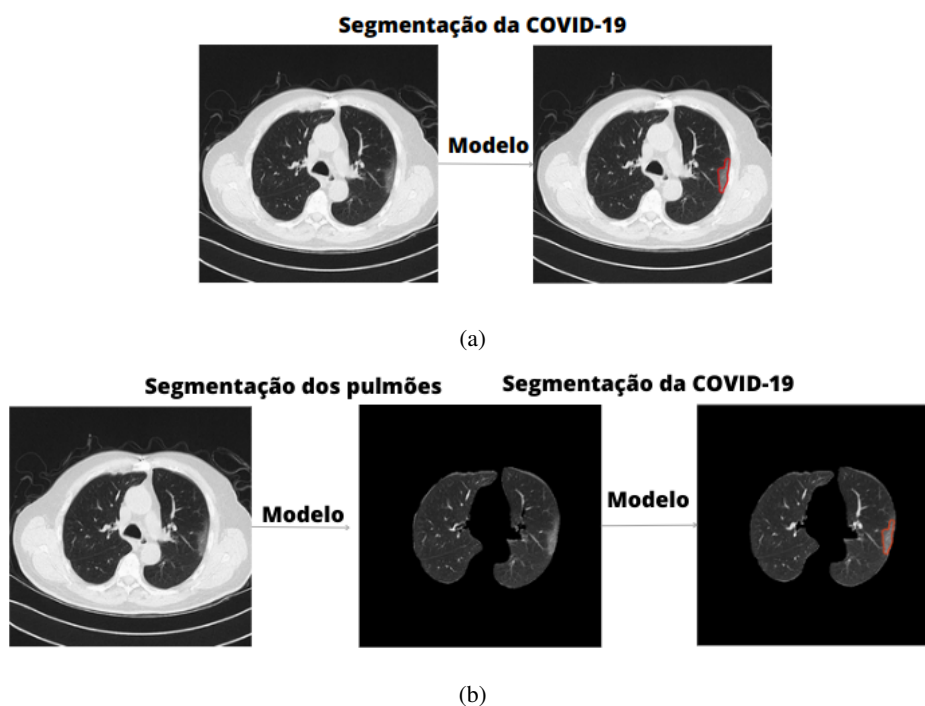


Figura 17 – Formas de segmentação da COVID-19. (a) Segmentação de uma etapa. (b) Segmentação de duas etapas.

estratégia de exame, e a segunda é chamada de estratégia de *slice*.

Na Tabela 8 estão os artigos que usaram a segmentação de uma etapa, separados conforme as estratégias de PD utilizadas. É importante ressaltar que, entre estes estudos, Kitrungrotsakul et al. (2021) e Shan et al. (2021) adotaram a estratégia de exame para particionamento de dados e apresentaram valores para a métrica Dice de 92,75% e 91,60%, respectivamente. Ambos usaram métodos interativos para obter modelos de segmentação da COVID-19. Os demais estudos usaram métodos automáticos para segmentar a COVID-19, com o trabalho de Wang et al. (2020) alcançando uma métrica Dice de 80,29%.

Entre os estudos que adotaram a estratégia de *slice* para particionamento de dados, todos usaram métodos automáticos, com o maior valor para a métrica Dice alcançado no trabalho de Ranjbarzadeh et al. (2021), 92%, seguido pelo trabalho de Amyar et al. (2020), 88%. Os estudos em Kitrungrotsakul et al. (2021), Amara et al. (2022), Zhang, Ren e Wei (2021) e Geng et al. (2024), empregaram um dos conjuntos de dados usados neste trabalho: COVID-19 CT Lung and Infection Segmentation Dataset (JUN et al., 2020). Este conjunto de dados é chamado neste trabalho de COVID-19 CT Seg.

Tabela 8 – Resumo dos artigos publicados nos últimos quatro anos sobre segmentação da COVID-19 em imagens de TC com métodos de segmentação de uma etapa.

Estratégia	Referência	Classificador	Conjunto de dados	Treinamento	Validação	Teste	Dice (%)
Exame	Kitrungrotsakul et al. (2021)	Attention-RefNet	Proprietário	149 exames	29 exames	32 exames	92,75
			COVID-19-CT-Seg	3 pastas com exames	1 pasta com exames	1 pasta com exames	91,37
			MICCAI	3 pastas com exames	1 pasta com exames	1 pasta com exames	81,22
	Shan et al. (2021)	VB-Net	Proprietário	249 exames	-	300 exames	91,60±10,00
	Wang et al. (2020)	COPLE-Net	Proprietário	378 exames	50 exames	130 exames	80,29±11,14
	Zhou et al. (2020)	U-Net	Harbin	4-fold CV com exames		1-fold CV com exames	78,30
	Li et al. (2020)	U-Net com backbone ResNet 34	Proprietário	531 exames	-	30 exames	Radiologista com 20 anos de experiência 74,00±0,28
							Radiologista com 15 anos de experiência 76,00±0,29
	Liu et al. (2021)	nCoVSegNet	LIDC-IDRI	875 exames	$\frac{1}{4}$	-	-
			MosMedData	40 exames	-	10 exames	68,43
			Coronacases	-	-	10 exames	68,94
	Hoang-Thi et al. (2021)	U-Net	Proprietário		50 exames	130 exames	67,00

Tabela 8 – Resumo dos artigos publicados nos últimos quatro anos sobre segmentação da COVID-19 em imagens de TC com métodos de segmentação de uma etapa (continuação).

Estratégia	Referência	Classificador	Conjunto de dados	Treinamento	Validação	Teste	Dice (%)
Slice	Ranjbarzadeh et al. (2021)	RNC	COVID-CT	70%	10%	20%	92,00
	Amyar et al. (2020)	Baseada na U-Net	COVID-CT COVID-19-CT <i>Henri Becquerel Cancer Center</i>	1.069 <i>slices</i>	150 <i>slices</i>	150 <i>slices</i>	88,00
	Amara et al. (2022)	O-Net	COVID-19-CT-Seg MosMedData	70%	-	30%	86,00
	Zhang, Ren e Wei (2021)	QC-HC U-Net	MSD	80%	-	20%	85,31
	Geng et al. (2024)	STCNet	COVID-19-CT-Seg	80%	-	20%	
			CC-CCII Segmentation	70%	-	30%	82,78
			COVID-19-CT-Seg	70%	-	30%	79,92
	Wu et al. (2021)	JSC	<i>COVID-CS</i>	2.794 <i>slices</i>	-	1.061 <i>slices</i>	78,50
	Fan et al. (2020)	Semi-Inf-Net	COVID-19 CT	45 <i>slices</i>	5 <i>slices</i>	50 <i>slices</i>	73,90
			COVID-19 CT Collection	1.600 <i>slices</i>	-	-	
	Fung et al. (2021)	SSInf-Net	COVID-19 ICTCF	698 <i>slices</i> 6.654 <i>slices</i>	114 <i>slices</i> -	117 <i>slices</i> -	63,00

Na Tabela 10 estão os artigos que usaram a segmentação de duas etapas, separados conforme as estratégias de PD utilizadas. Todos estes trabalhos implementaram métodos de segmentação automática. Entre os estudos que adotaram a estratégia de exame para particionamento de dados, destacamos os de Gong et al. (2021), Ma et al. (2021) e Wu et al. (2020), que obtiveram valores de métrica Dice de 71,30%, 67,30% e 63,20%, respectivamente. Entre os estudos que adotaram a estratégia de *slice* para particionamento de dados, destacamos os de Ilhan et al. (2023), Paluru et al. (2021), Kang et al. (2021) e Saood e Hatem (2021), com métricas Dice de 85,00%, 79,80%, 78,06% e 74,90%, respectivamente.

Os estudos em Paluru et al. (2021), Kang et al. (2021), Ma et al. (2021) e Wu et al. (2020) empregaram os conjuntos de dados usados neste trabalho: COVID-19 CT Seg (JUN et al., 2020) e/ou MedSeg Covid Dataset 2 (Artificial Intelligence AS, 2020). Este último é referido neste trabalho como MedSeg. Ma et al. (2021) usaram nnU-Net para segmentar os pulmões esquerdo e direito e detectar infecção por COVID-19. Em Wu et al. (2020), a arquitetura U-Net foi usada para segmentar regiões infectadas com COVID-19. Paluru et al. (2021) propuseram a arquitetura Anam-Net e forneceram as porcentagens de regiões normais e infectadas por COVID-19. Em Kang et al. (2021), a arquitetura RCNN-ResNeSt-200 foi usada no primeiro módulo, dos três propostos pelos autores, para treinar e avaliar modelos de segmentação dos pulmões e COVID-19.

Dentre os estudos listados nas Tabelas 8 e 10, foi observado que: nenhum estudo comparou a segmentação automática da COVID-19 por meio das duas estratégias de PD, exame e *slice*; que apenas Li et al. (2020) apresentou uma comparação do desempenho da segmentação automática da COVID-19 com o desempenho da segmentação realizada por dois grupos de radiologistas no mesmo conjunto de imagens e também realizou a concordância interobservador entre esses dois grupos; e que dois estudos utilizaram métodos interativos para segmentar a COVID-19, ou seja, contaram com intervenção humana no processo de segmentação, como visto em Kitrungrotsakul et al. (2021) e Shan et al. (2021).

Diante do que foi descrito acima, as contribuições deste trabalho são as seguintes:

- Comparar o desempenho de técnicas de aprendizado profundo com a segmentação de duas etapas para COVID-19 em imagens de tomografia computadorizada (TC), em ambas as estratégias de particionamento de dados, exame e *slice*;
- Comparar o desempenho de técnicas de aprendizado profundo com a segmentação de duas etapas para COVID-19 em imagens de tomografia computadorizada (TC), com o nível de concordância de segmentação realizado por dois grupos distintos de radiologistas no mesmo conjunto de dados;
- Comparar o desempenho de técnicas de aprendizado profundo com a segmentação de duas etapas para COVID-19 em imagens de tomografia computadorizada (TC), por meio de segmentação pulmonar manual e automática.

Tabela 9 – Resumo dos artigos publicados nos últimos quatro anos sobre segmentação da COVID-19 em imagens de TC com segmentação de duas etapas.

Estratégia	Referência	Classificador	Conjunto de dados	Treinamento	Validação	Teste	Dice (%)
Exame	Gong et al. (2021)	Pulmões	Proprietário	64 exames	23 exames	77 exames	COVID-19
		Modelo de outro autor		58 exames	14 exames	87 exames	Instituto 1
		COVID-19					71,30
		RNC 3D densa					Instituto 2
	Ma et al. (2021)	Pulmões	COVID-19-CT-Seg	4-fold CV com 16 exames	-	1-fold CV com 4 exames	Instituto 3
		nnU-Net					64,10
		COVID-19					Pulmão direito
		nnU-Net					87,90±9,30
	Wu et al. (2020)	Pulmões	Proprietário	COVID-19	COVID-19	COVID-19	Pulmão esquerdo
		Modelo de outro autor					85,80±10,50
		COVID-19					COVID-19
		U-Net					67,30±22,30
			MosMed	-	-	50 exames	COVID-19
							58,80±20,60
							COVID-19
							Infecção
			Proprietário	COVID-19	-	-	63,20
							Consolidação
							62,80
			Proprietário	COVID-19	-	-	

Tabela 10 – Resumo dos artigos publicados nos últimos quatro anos sobre segmentação da COVID-19 em imagens de TC com segmentação de duas etapas (continuação).

Estratégia	Referência	Classificador	Conjunto de dados	Treinamento	Validação	Teste	Dice (%)
Slice	Ilhan et al. (2023)	Pulmões Modelo de outro autor COVID-19 U-Net	Segmentation dataset nr. 2	-	-	COVID-19 Consolidação 9 exames	
			COVID-19 CT	4-fold CV com 80 slices	-	1-fold CV com 20 slices	COVID-19 85,00
	Paluru et al. (2021)	Pulmões Modelo de outro autor COVID-19 Anam-Net	COVID-19-CT-Seg (exp. 2)	3-fold CV	-	1-fold CV com 545 slices	COVID-19 79,80
			COVID-19 CT	270 slices	-	-	COVID-19 75,50
			Segmentation dataset nr. 2	-	-	704 slices	
			-	Modelo treinado no exp. 2	-	-	COVID-19 66,40
			Segmentation dataset nr. 2	-	-	704 slices	

Tabela 10 – Resumo dos artigos publicados nos últimos quatro anos sobre segmentação da COVID-19 em imagens de TC com segmentação de duas etapas (continuação).

Estratégia	Referência	Classificador	Conjunto de dados	Treinamento	Validação	Teste	Dice (%)
	Kang et al. (2021)	Pulmões RCNN-ResNeSt-200 COVID-19 RCNN-ResNeSt-200	NSCLC COVID-19-CT-Seg	50.756 <i>slices</i>	-	1.222 <i>slices</i>	Pulmões 97,18
			MosMed MSD COVID-19-CT-Seg	5.854 <i>slices</i>	-	1.117 <i>slices</i>	COVID-19 78,06
	Saood e Hatem (2021)	Pulmões Modelo de outro autor COVID-19 SegNet	COVID-19 CT	72 <i>slices</i>	10 <i>slices</i>	18 <i>slices</i>	COVID-19 74,90

3.2 Materiais e Métodos

3.2.1 Materiais

O conjunto de dados 1, COVID-19 CT Seg, contém 20 tomografias computadorizadas, 3.520 *slices* avaliados em dois estágios: avaliação inicial por dois radiologistas e revisão por um radiologista mais experiente (JUN et al., 2020). Um total de 1.844 *slices* são positivos para a COVID-19. A Figura 18(a) mostra o número de *slices* por tomografia computadorizada presente neste conjunto de dados. Para cada tomografia computadorizada, máscaras pulmonares e máscaras de COVID-19 estavam disponíveis. O número de *slices* com COVID-19 variou de 2 a 216 entre as tomografias computadorizadas. O tamanho original das imagens varia de 630x630 pixels ou 512x512 pixels, dependendo da tomografia. Todas as regiões da imagem associadas à COVID-19 foram rotuladas com opacidades em vidro fosco.

O conjunto de dados 2, MedSeg, contém nove tomografias computadorizadas (Artificial Intelligence AS, 2020). Este conjunto de dados tem 829 *slices* avaliados por um radiologista, dos quais 373 *slices* são positivos. A Figura 18(b) mostra o número de *slices* por tomografia computadorizada presente neste conjunto de dados. Ele contém sete tomografias computadorizadas em comum com o conjunto de dados 1.

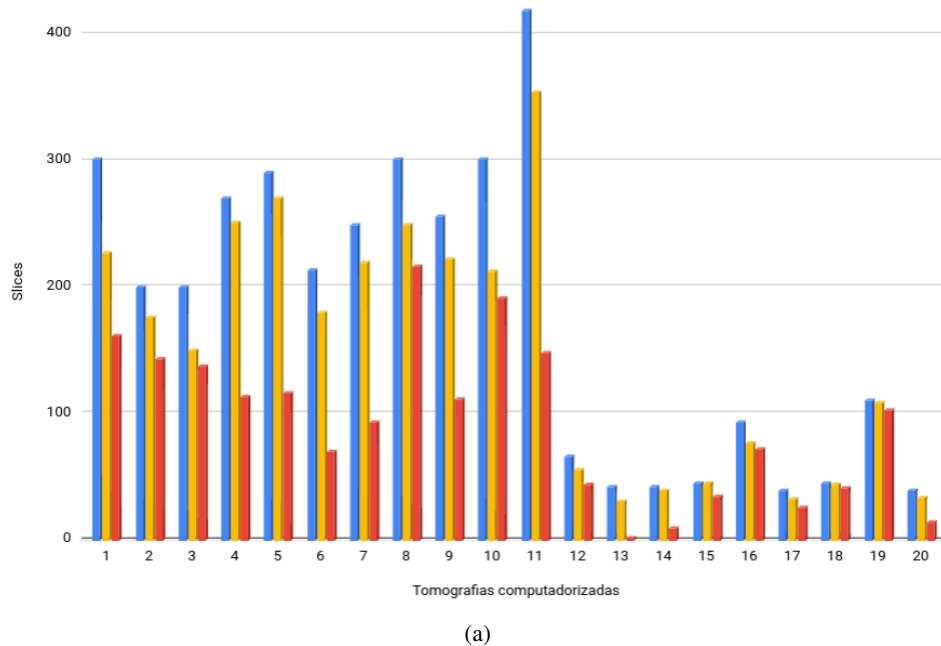
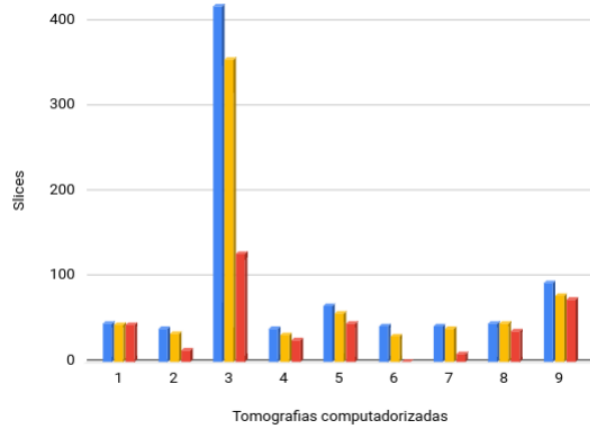


Figura 18 – Gráfico mostrando, para cada tomografia computadorizada, o número total de *slices* (em azul), o número de *slices* contendo os pulmões (em amarelo) e o número de *slices* contendo regiões com COVID-19 (em vermelho). (a) Conjunto de dados 1; (b) Conjunto de dados 2.

3.2.2 Métodos

Este trabalho implementa uma segmentação de duas etapas. A Figura 19 ilustra os métodos realizados para avaliar a segmentação de áreas de COVID-19 em tomografias computadorizadas.



(b)

Figura 18 – Gráfico mostrando, para cada tomografia computadorizada, o número total de *slices* (em azul), o número de *slices* contendo os pulmões (em amarelo) e o número de *slices* contendo regiões com COVID-19 (em vermelho). (a) Conjunto de dados 1; (b) Conjunto de dados 2 (continuação).

Com relação ao particionamento dos dados foram avaliadas as estratégias de exame e *slice*.

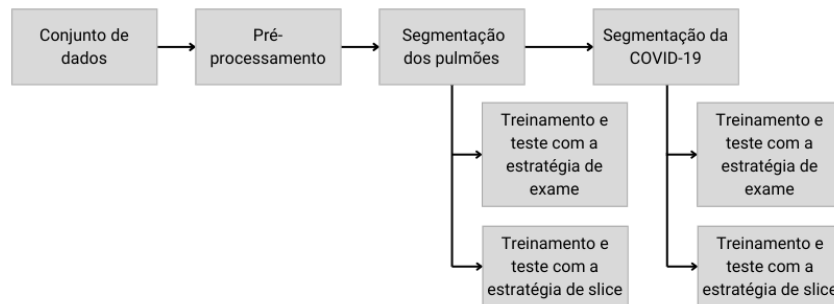


Figura 19 – Métodos adotados neste trabalho.

3.2.2.1 Pré-processamento

As etapas de pré-processamento são mostradas na Figura 20. Na primeira etapa, um radiologista do nosso grupo de pesquisa avaliou todas as tomografias computadorizadas de ambos os conjuntos de dados e relatou que duas tomografias computadorizadas do conjunto de dados 1 eram idênticas, com a única diferença que uma delas continha mais *slices* do que a outra. Portanto, um dos exames foi removido e o conjunto de dados 1 ficou com 19 tomografias computadorizadas.

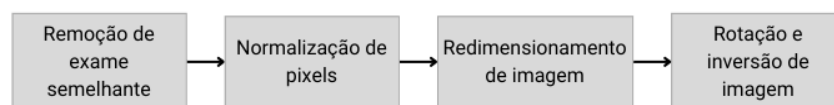


Figura 20 – Etapas realizadas no pré-processamento das imagens.

As intensidades dos pixels de cada *slice*, em ambos os conjuntos de dados, foram normalizados para o intervalo de 0-255, de acordo com o processo de normalização MAX-MIN mostrado em (1) (GONZALEZ; WOODS, 2010):

$$f_n = \frac{f_o - f_{\min}}{f_{\max} - f_{\min}} \quad (1)$$

em que:

f_n : valor do pixel normalizado;

f_o : valor do pixel original;

$f_{\min}$: valor mínimo do pixel no *slice*;

$f_{\max}$: valor máximo do pixel no *slice*.

Todas as imagens foram redimensionadas para 512 x 512 pixels. Nesta etapa, uma tomografia computadorizada do conjunto de dados 1, que apresentou uma proporção diferente entre altura e largura, foi removida. Portanto, o conjunto de dados 1 ficou com 18 tomografias computadorizadas. Para que todas as tomografias computadorizadas tivessem o tórax voltado para baixo, as operações de rotação e inversão foram aplicadas na última etapa de pré-processamento. A Figura 21 mostra exemplos de *slices* de tomografias computadorizadas após todas as modificações realizadas.

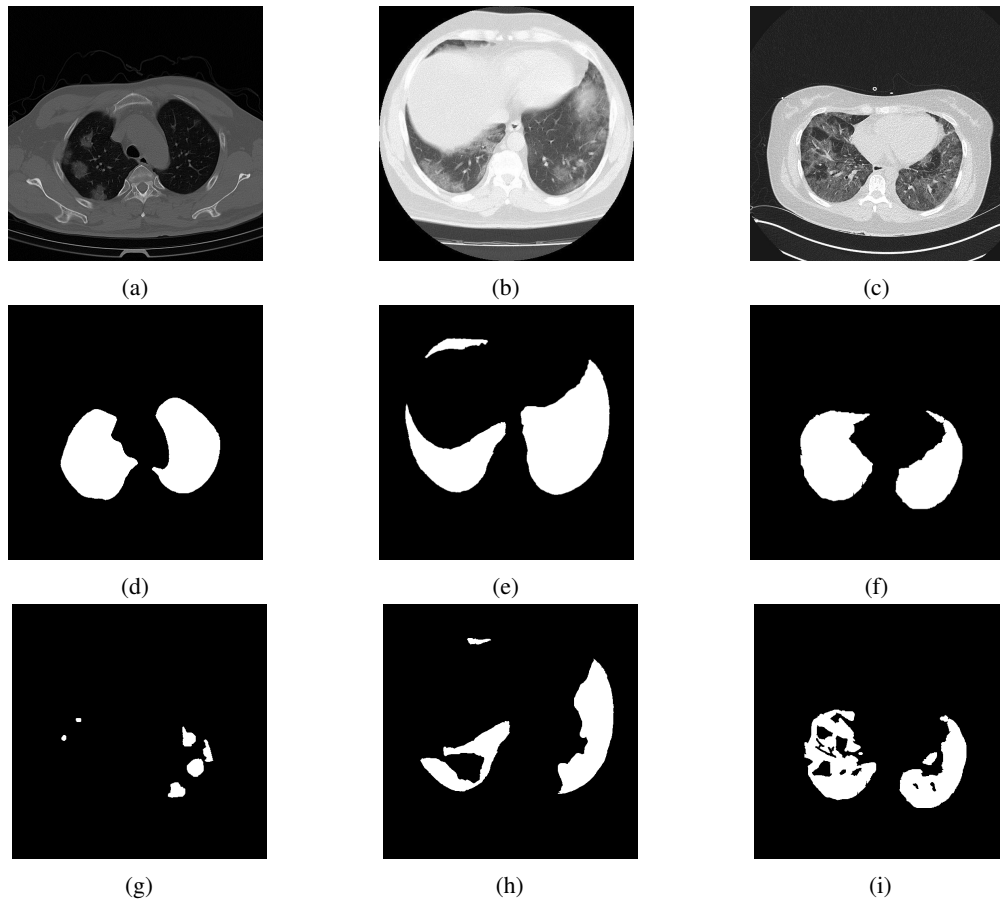


Figura 21 – Exemplos de *slices* de tomografias computadorizadas após pré-processamento. De cima para baixo: (a), (b) e (c) *slices*; (d), (e) e (f) máscaras pulmonares; (g), (h) e (i) máscaras de COVID-19.

3.2.2.2 Estratégias de PD para treinamento, validação e teste

Para dividir os conjuntos de dados em treinamento e teste, as duas estratégias de PD, por exame e *slice*, foram usadas. Na estratégia de exame, que usa tomografias computadorizadas completas exclusivamente para treinar ou para testar um classificador, foi usado um processo aleatório para selecionar as tomografias computadorizadas de ambos os conjuntos. A Figura 22 mostra como os conjuntos de treinamento e teste ficaram após a aplicação da estratégia de exame.

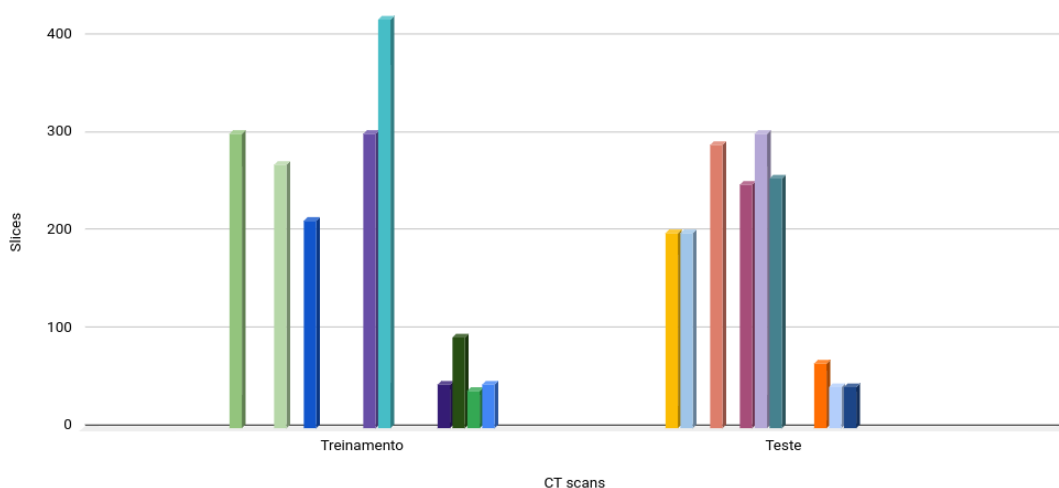


Figura 22 – Conjuntos de treinamento e teste obtidos com a estratégia de exame.

Na estratégia de *slice*, que usa uma porcentagem dos *slices* de uma tomografia computadorizada para treinar, e a porcentagem dos *slices* restantes para testar um classificador, foi usado um processo intercalado de 50% para selecionar os *slices* de cada tomografia computadorizada. A Figura 23 mostra como os conjuntos de treinamento e teste ficaram após a aplicação da estratégia de *slice*.

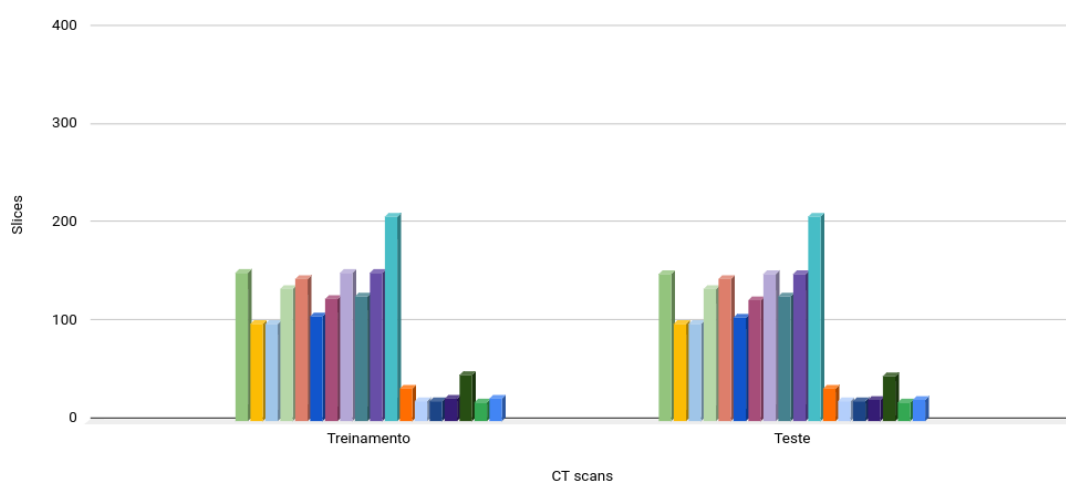


Figura 23 – Conjuntos de treinamento e teste obtidos com a estratégia de *slice*.

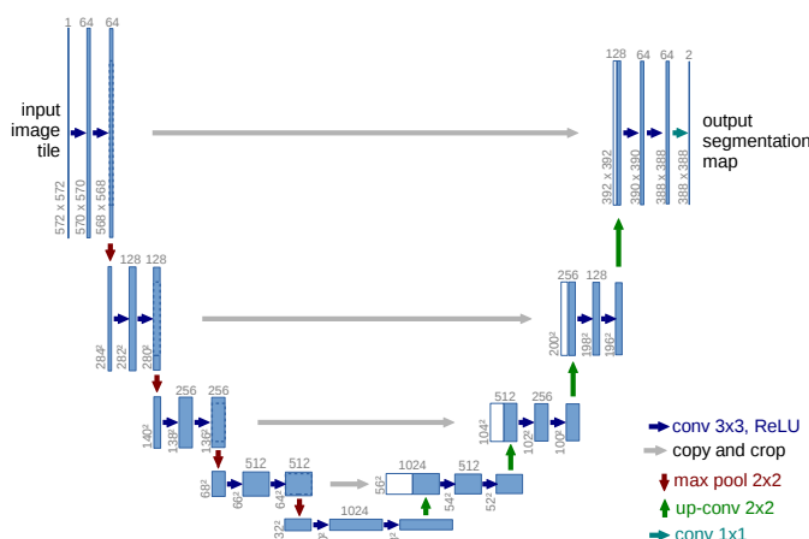
Todos os conjuntos de treinamento foram divididos em tempo de execução para criar um

conjunto de validação. Este procedimento foi realizado durante a fase de treinamento do modelo e a proporção adotada foi de 75:15.

3.2.2.3 Classificadores

Duas redes neurais convolutivas foram usadas neste trabalho para segmentação semântica. A U-Net (RONNEBERGER; FISCHER; BROX, 2015), Figura 24, consiste de um codificador com 4 estágios de subamostragem e um decodificador com 4 estágios de sobreamostragem. Esta arquitetura conecta camadas do codificador com camadas do decodificador.

Cada estágio de subamostragem é composto por 2 camadas de convolução com *kernel* 3×3 seguido por uma camada de *maxpooling* 2×2. Cada estágio de sobreamostragem é composto por uma camada de deconvolução 2×2, uma concatenação com o mapa de características correspondente do codificador e 2 camadas de convolução com *kernel* 3×3. Em alguns dos experimentos em nosso trabalho, para melhorar a generalização, uma camada *dropout* foi inserida entre a camada *Conv2D* com 1024 filtros e a camada *Conv2DTranspose* com 512 filtros.



Fonte: Ronneberger, Fischer e Brox (2015).

Figura 24 – Arquitetura de rede neural convolutiva usada para segmentação semântica de pulmões e COVID-19 - U-Net.

A CNN2, Figura 25, foi proposta pelo nosso grupo de pesquisa em Miyagawa et al. (2018), para segmentar o lúmen em imagens de tomografia de coerência óptica intravascular. Esta arquitetura é dividida em 4 blocos de subamostragem e 4 blocos de sobreamostragem.

Cada etapa da subamostragem é composta por camadas de convolução-batch-ReLU, seguida por uma camada de convolução 1x1 e uma camada *maxpooling* 2×2. Entre a última etapa de subamostragem e o início da primeira etapa de sobreamostragem, há uma camada de dropout. Cada etapa da sobreamostragem é composta por camadas de convolução transposta-batch-ReLU, seguida por camadas de convolução-batch-ReLU e uma camada de convolução 1x1. Na saída,

3.2.2.4 Segmentação pulmonar e da COVID-19

Para segmentação pulmonar, apenas a arquitetura U-Net foi empregada. Para cada estratégia PD, 8 experimentos foram realizados, combinando dropout (D), regularização L_2 (L) e aumento de dados (A). Para segmentação da COVID-19, ambas as arquiteturas U-Net e CNN2 foram empregadas. Para cada estratégia, 8 experimentos também foram realizados usando a mesma combinação mencionada. A segmentação da COVID-19 foi obtida por meio de segmentação pulmonar automática e manual. Para segmentação pulmonar manual, os rótulos pulmonares dos radiologistas foram usados.

3.2.3 Ambiente de simulação e métricas de avaliação

O ambiente usado para escrever os códigos para treinamento e teste dos modelos foi um serviço de nuvem hospedado pela Google, o *Google Colab Pro* (Google, 2020b). O *software* usado para implementar os códigos foi o *Keras* (Keras, 2020), uma API de aprendizado profundo escrita em *Python*. O *kernel* de aprendizado de máquina usado com o *Keras* foi o *Tensorflow 2.12.0* (Tensorflow, 2020), uma plataforma de código aberto. A versão do *Python* usada foi a 3.9.16. Todos os códigos e modelos gerados foram armazenados no *Google Drive*, cuja a capacidade de armazenamento foi ampliada com o *Google One* (Google, 2020a).

Este trabalho usou o *Google Colab Pro* com a seguinte configuração: capacidade de armazenamento, 100 GB; memória RAM, 35 GB; processador, Intel(R) Xeon(R) CPU @ 2,30 GHz; sistema operacional, Linux com distribuição Ubuntu 20.04.3 LTS; unidade de processamento de tensor, TPU v2; e memória RAM da unidade de processamento de tensor, 16 GB de HBM (*High Bandwidth Memory*).

Para avaliar o desempenho do modelo, foram empregadas as seguintes métricas: acurácia (2), acurácia global (3), Dice (4) e Jaccard (5).

$$Acurácia = \frac{\frac{VP}{VP+FN} + \frac{VN}{VN+FP}}{2} \quad (2)$$

$$Acurácia\ global = \frac{VP + VN}{VP + VN + FP + FN} \quad (3)$$

$$Dice = \frac{2VP}{2VP + FP + FN} \quad (4)$$

$$Jaccard = \frac{VP}{VP + FP + FN} \quad (5)$$

em que:

VP: verdadeiros positivos;

VN: verdadeiros negativos;

FP: falsos positivos;

FN: falsos negativos.

Para avaliar se as diferenças entre os valores médios de duas métricas foram estatisticamente significativas, foi utilizado o teste t de Student, com nível de significância de 1%.

3.3 Resultados

Nesta seção, apresentamos resultados referentes aos seguintes tópicos: segmentação pulmonar, segmentação de áreas de COVID-19 com segmentação pulmonar automática, segmentação de áreas de COVID-19 com segmentação pulmonar manual e uma comparação entre segmentação automática da COVID-19 e concordância interobservador.

3.3.1 Resultados da segmentação pulmonar automática

Para ambas as estratégias de PD, os melhores resultados no conjunto de validação foram obtidos com U-Net + D + L (estratégia de exame, Dice = 99,79%; estratégia de *slice*, Dice = 98,64%).

Com os melhores modelos do conjunto de validação, os resultados para segmentação pulmonar no conjunto de teste foram os apresentados na Tabela 12. Os resultados obtidos com a estratégia de *slice* foram melhores do que os resultados obtidos com a estratégia de exame. No entanto, não houve diferenças estatisticamente significativas. Para acurácia, o valor $t = 2,17$, $p > 0,01$ ($p = 0,019797$), IC de 95%: exame - [96,293 – 98,907], *slice* - [98,395 – 99,125]. Para acurácia global, o valor $t = 2,07$, $p > 0,01$ ($p = 0,024445$), IC de 95%: exame - [98,135 – 99,885], *slice* - [99,572 – 99,748]. Para Dice, o valor $t = 2,42$, $p > 0,01$ ($p = 0,011382$), IC de 95%: exame - [91,886 – 98,694], *slice* - [97,907 – 98,673]. Para Jaccard, o valor $t = 2,44$, $p > 0,01$ ($p = 0,010892$), IC de 95%: exame - [91,805 – 98,495], *slice* - [97,719 – 98,541]. Para garantir a validade do teste t de Student, aplicamos o teste Shapiro-Wilk às acurácias, acurácias globais, valores de Dice e Jaccard, e obtivemos $W(27) = 0,77$, $p < 0,01$ ($p = 4,38 \text{ e-}5$); $W(27) = 0,44$, $p < 0,01$ ($p = 4,12 \text{ e-}9$); $W(27) = 0,53$, $p < 0,01$ ($p = 3,83 \text{ e-}8$) e $W(27) = 0,55$, $p < 0,01$ ($p = 5,70 \text{ e-}8$), respectivamente.

Tabela 12 – Resultados da segmentação pulmonar no conjunto de teste.

Estratégia	Acurácia (%)	Acurácia global (%)	Dice (%)	Jaccard (%)
Exame	97,60±2,00	99,01±1,34	95,29±5,21	95,15±5,12
<i>Slice</i>	98,76±0,79	99,66±0,19	98,29±0,83	98,13±0,89

3.3.2 Resultados da segmentação da COVID-19 automática

3.3.2.1 Com segmentação pulmonar automática

Para a rede U-NET, os melhores resultados no conjunto de validação para a estratégia de exame foi obtido com U-NET (Dice = 87,30%) e para a estratégia de *slice* foi obtido com U-NET

+ A + D (Dice = 86,61%). Para a rede CNN2, os melhores resultados no conjunto de validação para a estratégia de exame foi obtido com CNN2 + D (Dice = 86,06%) e para a estratégia de *slice* foi obtido com CNN2 + A (Dice = 84,22%).

Para ambas as estratégias de PD, os melhores resultados no conjunto de validação foram obtidos com a arquitetura U-Net. A Tabela 13 mostra os resultados para a segmentação da COVID-19 no conjunto de teste com esta arquitetura. Os resultados obtidos com a estratégia de *slice* foram melhores do que os resultados obtidos com a estratégia de exame. As diferenças entre os valores de Dice e Jaccard foram estatisticamente significativas: valor $t = 3,93$, $p < 0,01$ ($p = 0,00033$), IC de 95%: exame - [66,823 – 79,197], *slice* - [82,161 – 87,159] e valor $t = 4,02$, $p < 0,01$ ($p = 0,000308$), IC de 95%: exame - [75,121 – 82,699], *slice* - [84,885 – 88,535], respectivamente. No entanto, as diferenças entre os valores de acurácia e acurácia global não foram estatisticamente significativas: valor $t = 0,98$, $p > 0,01$ ($p = 0,167464$), IC de 95%: exame - [85,031 – 92,649], *slice* - [89,003 – 92,477] e valor $t = 0,67$, $p > 0,01$ ($p = 0,255797$), IC de 95%: exame - [99,289 – 99,891], *slice* - [99,538 – 99,842], respectivamente. Para garantir a validade do teste t de Student, aplicamos o teste de Shapiro-Wilk às acurácias, acurácias globais, valores de Dice e Jaccard, e obtivemos $W(27) = 0,88$, $p < 0,01$ ($p = 0,007$); $W(27) = 0,77$, $p < 0,01$ ($p = 7,61 \text{ e-}5$); $W(27) = 0,87$, $p < 0,01$ ($p = 0,005$) e $W(27) = 0,90$, $p < 0,05$ ($p = 0,023$), respectivamente.

Tabela 13 – Resultados da segmentação da COVID-19 no conjunto de teste usando segmentação pulmonar automática.

Estratégia	Acurácia (%)	Acurácia global (%)	Dice (%)	Jaccard (%)
Exame	88,84±5,83	99,59±0,46	73,01±9,47	78,91±5,80
<i>Slice</i>	90,74±3,76	99,69±0,33	84,66±5,41	86,71±3,95

3.3.2.2 Com segmentação pulmonar manual

Os resultados para segmentação da COVID-19 mostrados na Tabela 14, referentes às estratégias de exame e *slice*, são semelhantes aos mostrados na Tabela 13 : os valores de Dice e Jaccard são estatisticamente significativos, enquanto os valores de acurácia e acurácia global não são estatisticamente significativos. Para acurácia, o valor $t = 0,37$, $p > 0,01$ ($p = 0,355785$), IC de 95%: exame - [86,792 – 94,828], *slice* - [89,800 – 93,320]; para acurácia global o valor $t = 0,54$, $p > 0,01$; ($p = 0,296566$), IC de 95%: exame - [97,559 – 101,741], *slice* - [98,426 – 101,014]; para Dice, o valor $t = 3,43$, $p < 0,01$ ($p = 0,001132$), IC de 95%: exame - [67,976 – 80,964], *slice* - [82,606 – 88,094]; e para Jaccard, o valor $t = 3,49$, $p < 0,01$ ($p = 0,000981$). Para garantir a validade do teste t de Student, aplicamos o teste de Shapiro-Wilk às acurácias, acurácias globais, valores de Dice e Jaccard, e obtivemos $W(27) = 0,92$, $p < 0,05$ ($p = 0,04$); $W(27) = 0,81$, $p < 0,01$ ($p = 2,80 \text{ e-}4$); $W(27) = 0,89$, $p < 0,05$ ($p = 0,013$) e $W(27) = 0,88$, $p < 0,05$ ($p = 0,011$), respectivamente.

Tabela 14 – Resultados da segmentação da COVID-19 no conjunto de teste usando segmentação pulmonar manual.

Estratégia	Acurácia (%)	Acurácia global (%)	Dice (%)	Jaccard (%)
Exame	90,81±6,15	99,65±0,32	74,47±9,94	79,90±6,04
<i>Slice</i>	91,56±3,81	99,72±0,28	85,35±5,94	87,29±4,36

Para cada estratégia de PD, as métricas mostradas na Tabela 14 são melhores do que aquelas mostradas na Tabela 13 com segmentação pulmonar automática. No entanto, não há diferenças estatisticamente significativas para a estratégia de exame (valor t da acurácia = 0,69, $p > 0,01$ ($p = 0,260415$); valor t da acurácia global = 0,32, $p > 0,01$ ($p = 0,374759$); valor t de Dice = 0,29, $p > 0,01$ ($p = 0,384441$); e valor t de Jaccard = 0,34, $p > 0,01$ ($p = 0,371246$)) ou para a estratégia de *slice* (valor t da acurácia = 0,64, $p > 0,01$ ($p = 0,265343$); valor t da acurácia global = 0,25, $p > 0,01$ ($p = 0,400395$); valor t de Dice = 0,35, $p > 0,01$ ($p = 0,362962$); e valor t de Jaccard = 0,40, $p > 0,01$ ($p = 0,344803$)).

3.3.3 Áreas segmentadas das tomografias computadorizadas

As áreas segmentadas automaticamente dos pulmões e da COVID-19 nos conjuntos de teste, expressas como porcentagens da área da tomografia computadorizada, são comparadas com a mesma área segmentada por radiologistas. A Figura 26 mostra os gráficos de caixas para as segmentações dos pulmões e COVID-19, comparadas com as segmentações feitas pelos radiologistas.

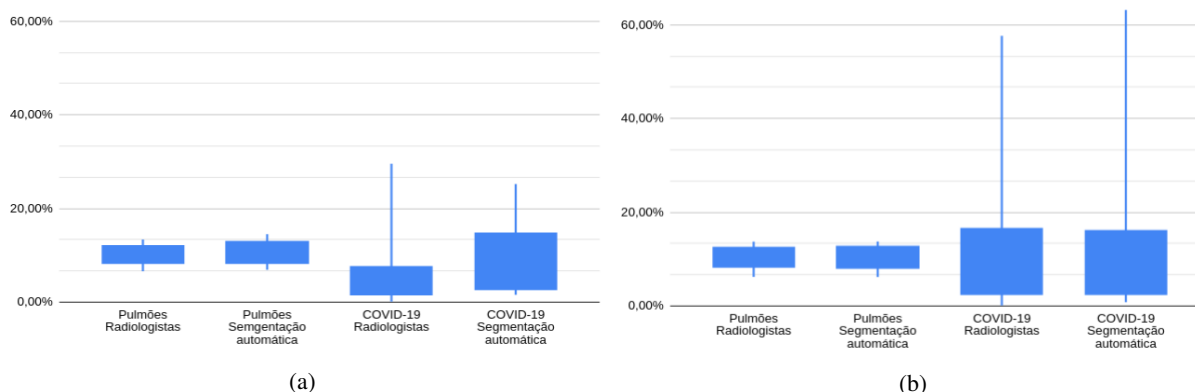


Figura 26 – Comparação das áreas dos pulmões e COVID-19 segmentadas automaticamente com as áreas segmentadas pelos radiologistas no conjunto de teste. (a) Estratégia de exame. (b) Estratégia de *slice*.

Conforme mostra a Figura 26, a porcentagem de áreas segmentadas com a estratégia de *slice* é mais semelhante à porcentagem da área segmentada pelos radiologistas. Esses resultados podem ser explicados pelo uso da Tabela 13, que compara a segmentação automática da COVID-19 por meio da estratégia de exame e *slice*. Conforme mostrado nesta tabela, as métricas para a estratégia de *slice* são maiores do que as métricas para a estratégia de exame, e as diferenças são estatisticamente significativas para os valores de Dice e Jaccard.

3.3.4 Comparação da segmentação da COVID-19 com a concordância inter-observador

Os resultados dos melhores modelos para segmentação de COVID-19 para a estratégia de *slice* foram comparados com a concordância interobservador. Para essa comparação, foram usadas as 7 tomografias computadorizadas comuns ao conjunto de dados 1 e 2.

Conforme mostra a Tabela 15, o método automático usando a estratégia de *slice* apresentou melhores métricas do que a concordância interobservador em ambos os conjuntos de dados. No entanto, as diferenças entre os resultados do método automático em ambos os conjuntos de dados e a concordância interobservador não foram estatisticamente significativas. Neste caso, o teste de Shapiro-Wilk mostrou que o teste t de Student não pode ser aplicado. Portanto, foi aplicado o teste de Mann-Whitney. Para o conjunto de dados 1 x interobservador, obtivemos: acurácia: $U = 16$, $p > 0,01$ ($p = 0,15386$); acurácia global: $U = 20$, $p > 0,01$ ($p = 0,30503$); Dice: $U = 13$, $p > 0,01$ ($p = 0,07927$); Jaccard: $U = 13$, $p > 0,01$ ($p = 0,07927$). Para o conjunto de dados 2 x interobservador, obtivemos: acurácia: $U = 6$, $p > 0,01$ ($p = 0,01072$); acurácia global: $U = 24,5$, $p > 0,01$ ($p = 0,47608$); Dice: $U = 17$, $p > 0,01$ ($p = 0,18673$); Jaccard: $U = 13$, $p > 0,01$ ($p = 0,07927$).

Tabela 15 – Comparação entre a segmentação da COVID-19 com a estratégia de *slice* e a concordância inter-observador.

Método	Acurácia (%)	Acurácia global (%)	Dice (%)	Jaccard (%)
Método automático no conjunto de dados 1	89,64±4,68	99,51±0,43	83,21±7,52	85,67±5,45
Método automático no conjunto de dados 2	88,53±5,93	99,61±0,32	80,78±9,45	84,13±6,54
Concordância interobservador	88,08±5,16	99,52±0,43	77,69±7,53	81,76±4,89

3.3.5 Exemplos de segmentação pulmonar e da COVID-19 automáticas

A Figura 27 mostra um exemplo de segmentação pulmonar e da COVID-19 em um *slice* de uma tomografia computadorizada utilizando a U-Net, treinada com as estratégias de exame e *slice*. Conforme mostrado, observa-se uma grande concordância entre as áreas segmentadas por ambas as estratégias.

A Figura 28 mostra um exemplo de segmentação da COVID-19 realizada por dois grupos de radiologistas em um *slice* de uma tomografia computadorizada. Conforme mostrado, há uma leve diferença entre as segmentações realizadas pelos dois grupos.

Em áreas de baixo contraste em imagens de TC, podem ocorrer variações de segmentação, conforme mostrado na Figura 29. Nestes casos, a dificuldade em distinguir áreas com intensidades de pixels semelhantes compromete a identificação de áreas afetadas pela COVID-19,

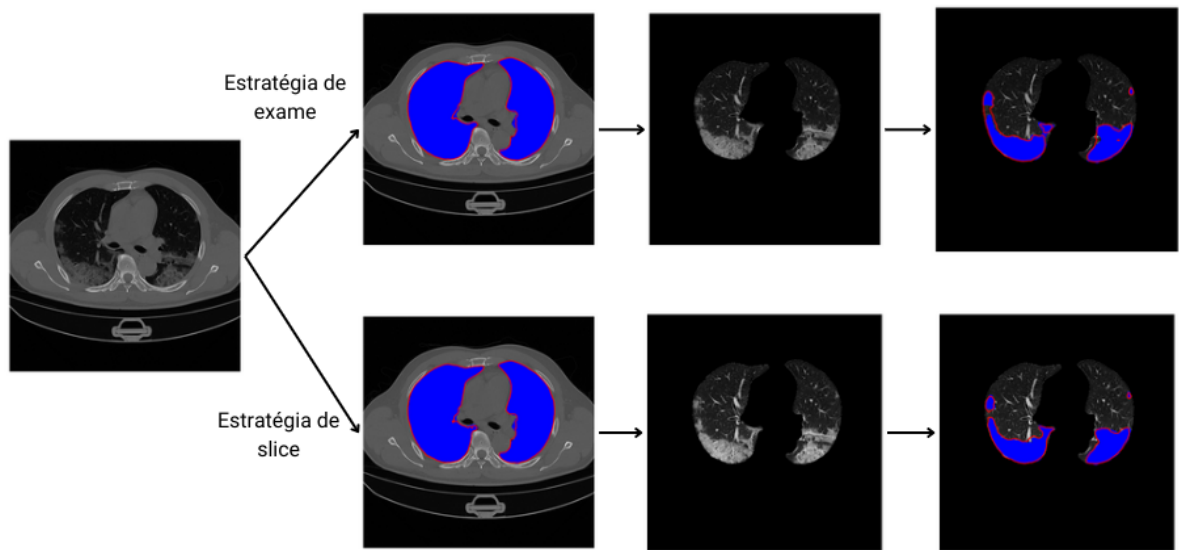


Figura 27 – A primeira linha mostra a segmentação pulmonar e da COVID-19 com a estratégia de exame. A segunda linha mostra a segmentação pulmonar e da COVID-19 com a estratégia de *slice*. A segmentação dos radiologistas (conjunto de dados 1) é mostrada em azul, enquanto a segmentação automática é mostrada em vermelho.

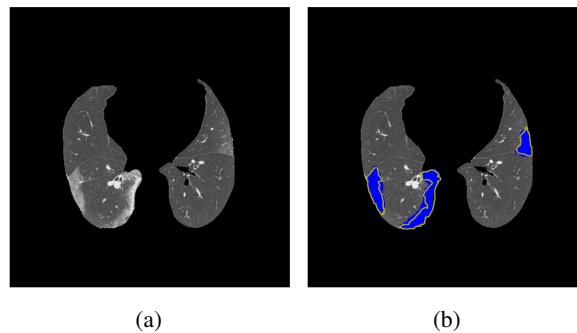


Figura 28 – (a) Imagem original. (b) Concordância entre radiologistas. A segmentação dos radiologistas no conjunto de dados 1 é mostrada em azul, enquanto a segmentação dos radiologistas no conjunto de dados 2 é mostrada em amarelo.

como também foi identificado por autores que utilizaram a segmentação de uma etapa e a estratégia de exame (SHAN et al., 2021; WANG et al., 2020), a segmentação de uma etapa e a estratégia de *slice* (FAN et al., 2020), a segmentação de duas etapas e a estratégia de exame (MA et al., 2021) e a segmentação de duas etapas e a estratégia de *slice* (ILHAN et al., 2023).

O tempo necessário para avaliar uma imagem de 512×512 pixels é de aproximadamente 0,08s para a estratégia de exame e 0,07s para a estratégia de *slice*. Para uma avaliação de uma tomografia computadorizada com 200 imagens, por exemplo, isso implica um tempo de 16 segundos usando a estratégia de exame e 14 segundos usando a estratégia de *slice*.

3.4 Discussão

Na Tabela 10, que apresenta os resultado de uma revisão da literatura sobre segmentação de duas etapas, também é possível observar que os melhores valores da métrica Dice são obtidos com

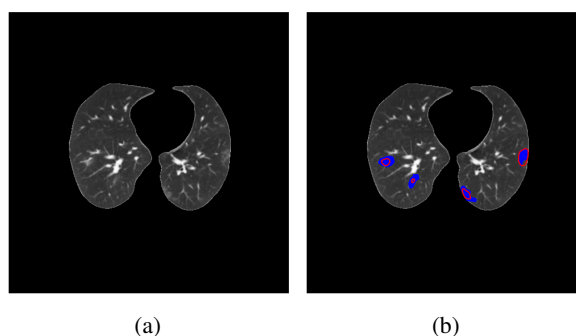


Figura 29 – (a) *Slice* com áreas de baixo contraste. (b) A segmentação dos radiologistas (conjunto de dados 1) é mostrada em azul, enquanto a segmentação automática é mostrada em vermelho.

a estratégia de *slice* em comparação com a estratégia de exame. As Tabelas 13 e 14 demonstram que os resultados obtidos neste estudo com segmentação de duas etapas, utilizando ambas as estratégias de PD, são competitivos com resultados previamente apresentados na literatura.

Por exemplo, para a segmentação da COVID-19 com a estratégia de exame, os melhores valores para a métrica Dice obtidos na literatura para a segmentação de duas etapas foram 71,3% e 67,3% para Gong et al. (2021) e Ma et al. (2021), respectivamente. Neste trabalho, com a segmentação automática dos pulmões, foi obtido um Dice de 73,01%. Por outro lado, para a segmentação da COVID-19 com a estratégia de *slice*, os melhores valores para a métrica Dice obtidos na literatura para segmentação de duas etapas foram 85,0% e 79,8% em Ilhan et al. (2023) e Paluru et al. (2021), respectivamente. Neste trabalho, com a segmentação automática dos pulmões, foi obtido um Dice de 84,66%.

Em relação à segmentação pulmonar, os resultados obtidos neste trabalho também são compatíveis com o estado da arte. O melhor valor para a métrica Dice obtido na literatura para segmentação pulmonar foi de 97,18% (KANG et al., 2021) com a estratégia de *slice*. Considerando a estratégia de exame os valores foram de 87,90% para o pulmão direito e 85,80% para o pulmão esquerdo em Ma et al. (2021). Neste trabalho, obtivemos valores de Dice de 95,29% com a estratégia de exame e 98,29% com a estratégia de *slice*.

A rede U-Net também foi usada em vários outros estudos anteriores sobre a segmentação de áreas de COVID-19: Amyar et al. (2020), Zhou et al. (2020), Li et al. (2020), Hoang-Thi et al. (2021) e Wu et al. (2021). Neste estudo, embora os conjuntos de dados usados sejam pequenos, creditamos os bons resultados obtidos à combinação da arquitetura U-Net com técnicas para melhorar a generalização: dropout, regularização L_2 e aumento de dados. Várias combinações dessas três técnicas foram usadas para obter os melhores resultados.

Neste estudo, nós comparamos o desempenho dos modelos profundos propostos para segmentação da COVID-19, quando a estratégia de *slice* foi usada, com a concordância inter-observador. Na literatura, apenas o estudo Li et al. (2020) comparou a segmentação da COVID-19 com métodos automáticos e a concordância interobservador. Para a estratégia de exame, os autores obtiveram Dices de $0,74 \pm 0,28$ e $0,76 \pm 0,29$ para a segmentação da região infectada pela COVID-19, comparando os resultados do método automático com os de dois radiologistas

experientes, respectivamente. Esses valores foram próximos à concordância interobservador, que foi de $0,79 \pm 0,25$.

Neste trabalho, usando a estratégia de *slice*, conforme apresentado na Tabela 15, obtivemos um Dice de $83,21 \pm 7,52$ e $80,78 \pm 9,45$ para as segmentações da região infectada por COVID-19, realizadas pelo método automático e pelos dois grupos de radiologistas, nos conjuntos de dados 1 e 2, respectivamente. A concordância interobservador apresentou um valor de Dice de $77,69 \pm 7,53$. No entanto, as diferenças entre os valores de Dice obtidos pelo método de segmentação em ambos os conjuntos de dados e a concordância interobservador não foram estatisticamente significativas.

Os valores de Dice dos métodos automáticos obtidos neste trabalho, com a estratégia de *slice*, são maiores que os valores obtidos em Li et al. (2020) com a estratégia de exame. O valor de Dice interobservador obtido neste trabalho é próximo ao obtido em Li et al. (2020). No entanto, o desvio padrão é maior. Devido ao pequeno número de tomografias computadorizadas comuns nos dois conjuntos de dados, não foi possível fazer uma comparação via estratégia de exame.

A comparação dos resultados apresentados na Tabela 14 mostrou que, para a segmentação automática da COVID-19, quando a rede U-Net e a segmentação manual dos pulmões foram utilizadas, os valores obtidos foram melhores para a estratégia de *slice* do que para a estratégia de exame, com as métricas Dice e Jaccard com diferenças estatisticamente significativas no nível de significância de 1%. A avaliação dos resultados apresentados nas Tabelas 13 e 14 mostrou que, para as mesmas métricas mencionadas, não houve diferenças estatisticamente significativas no nível de significância de 1% quando foi utilizada a segmentação manual ou automática dos pulmões.

O método demonstrou capacidade de identificar e quantificar alterações pulmonares. Os resultados sugerem que ele pode ser aplicado em outras condições pulmonares com características radiológicas semelhantes, como opacidades em vidro fosco e consolidações pulmonares, comuns em infarto pulmonar, pneumonias virais e pneumonite por hipersensibilidade. A segmentação oferece uma forma de quantificar o envolvimento pulmonar que pode ser útil para monitorar a progressão de doenças e avaliar respostas ao tratamento. Em cenários de alta demanda por exames e sistemas de saúde sobrecarregados, o método pode ser integrado a sistemas de triagem, ajudando a otimizar o fluxo de trabalho e priorizar exames com alterações significativas. A capacidade do método de segmentar e quantificar áreas de alteração pulmonar, com diferenças estatisticamente não significativas em comparação à segmentação realizada por dois grupos de radiologistas, sugere seu potencial como uma ferramenta valiosa, com aplicação além do contexto da COVID-19. Isso pode contribuir para a melhoria do diagnóstico e manejo de diversas doenças pulmonares.

4 CONCLUSÃO E TRABALHOS FUTUROS

O presente trabalho teve como objetivo propor métodos automáticos baseados em técnicas de aprendizado profundo para auxiliar no diagnóstico da malária e da COVID-19. No geral, as previsões dos modelos foram competitivas com as encontradas na literatura.

O capítulo 2 apresentou um método automático para detecção de parasitas da malária em imagens microscópicas de gota espessa, usando classificadores de pixels para obter objetos em primeiro plano e delimitar objetos com parasitas. Para ambos os processos, os atributos de cor através dos componentes HSV foram usados. Para classificação dos *patches*, foi utilizada uma rede CNN com 34 camadas. Os resultados mostraram que diferentes valores para as métricas são obtidos com imagens contendo diferentes tamanhos de parasitas ou diferentes estágios de desenvolvimento. Os melhores valores para as métricas foram obtidos com imagens contendo tamanhos grandes de parasitas, provavelmente porque parasitas pequenos são facilmente confundidos com ruído. Os resultados obtidos são comparáveis aos de outros trabalhos previamente publicados na literatura.

O capítulo 3 apresentou um método automático para detecção de áreas afetadas pela COVID-19 em imagens de tomografia computadorizada. Ao utilizar a U-NET, a segmentação da COVID-19 obteve resultados ligeiramente melhores com a estratégia de *slice* em comparação à estratégia de exame. Para as métricas Dice e Jaccard, observou-se uma diferença estatisticamente significativa no nível de 1%. A comparação entre a segmentação automática da COVID-19 realizada com a U-NET e a concordância interobservador, com base em um conjunto de 7 tomografias computadorizadas utilizando a estratégia de *slices*, não revelou diferenças estatisticamente significativas. Além disso, quando a U-Net foi aplicada, não foram encontradas diferenças estatisticamente significativas em termos de acurácia, acurácia global, Dice e Jaccard, entre as segmentações manuais e automáticas dos pulmões.

Os resultados obtidos neste trabalho demonstram um avanço significativo em relação às metodologias encontradas na literatura, evidenciando a eficácia das abordagens propostas para a análise de imagens de gota espessa e tomografias computadorizadas. Embora os conjuntos de dados utilizados sejam relativamente pequenos, os métodos desenvolvidos apresentam grande potencial para aplicação em cenários clínicos. Além disso, a metodologia adotada pode ser expandida para a interpretação de imagens de outras doenças, contribuindo para o aprimoramento de ferramentas diagnósticas. A seção 4.1 apresenta os trabalhos futuros para a ampliação e aplicação deste estudo.

4.1 Trabalhos futuros

Os resultados apresentados neste trabalho destacam o potencial das técnicas de aprendizado profundo na criação de métodos automáticos para auxiliar no diagnóstico de doenças infecciosas.

Como direções para trabalhos futuros, seria interessante explorar outras técnicas e arquiteturas de aprendizado profundo, bem como outros conjuntos de dados. Especificamente, sugerimos as seguintes tarefas:

- Realizar uma avaliação clínica do método proposto para detecção de parasitas da malária, descrito no capítulo 2, por meio de uma comparação com os resultados de técnicos experientes em microscopia;
- Aplicar a estratégia de exame, apresentada no capítulo 3, em conjuntos de dados avaliados por dois grupos de radiologistas e comparar: 1. Os resultados obtidos pelos métodos automáticos com os resultados de cada grupo de radiologistas; 2. Os resultados dos métodos automáticos com a concordância interobservador;
- Utilizar conjunto de dados maiores do que os descritos nos capítulos 2 e 3, para realizar uma comparação com os resultados obtidos pelos métodos propostos neste trabalho.

REFERÊNCIAS

- ABDURAHMAN, F.; FANTE, K. A.; ALIY, M. Malaria parasite detection in thick blood smear microscopic images using modified yolov3 and yolov4 models. *BMC Bioinformatics*, v. 22, p. 112, 2021. Disponível em: <<https://bmcbioinformatics.biomedcentral.com/articles/10.1186/s12859-021-04062-2>>.
- ABIDIN, S.; SALAMAH, U.; NUGROHO, A. Segmentation of malaria parasite candidates from thick blood smear microphotographs image using active contour without edge. In: *2016 1st International Conference on Biomedical Engineering (IBIOMED)*. [S.l.]: IEEE, 2016. p. 1–6.
- AGGARWAL, C. C. *Neural Networks and Deep Learning*. [S.l.: s.n.], 2021.
- AMARA, K. et al. COVIR: A virtual rendering of a novel NN architecture O-Net for COVID-19 Ct-scan automatic lung lesions segmentation. *Comput Graph*, v. 104, 2022. ISSN 23146141.
- AMYAR, A. et al. Multi-task deep learning based CT imaging analysis for COVID-19 pneumonia: Classification and segmentation. *Computers in Biology and Medicine*, v. 126, 2020.
- ARIS, T. A. et al. Fast k-means clustering algorithm for malaria detection in thick blood smear. In: *2020 IEEE 10th International Conference on System Engineering and Technology (ICSET)*. IEEE, 2020. p. pp. 268–273. Disponível em: <<https://ieeexplore.ieee.org/document/9356443>>.
- Artificial Intelligence AS. *COVID-19*. 2020. Disponível em: <<http://medicalsegmentation.com/covid19/>>.
- AZIF, F. M.; NUGROHO, H. A.; WIBIRAMA, S. Adaptive threshold determination based on entropy in active contour without edge method for malaria parasite candidate detection. *2018 4th International Conference on Science and Technology (ICST)*, p. 1–5, 2018.
- BONACCORSO, G. *Machine Learning Algorithms*. Packt, 2017. ISBN 9781785889622. Disponível em: <<https://www.safaribooksonline.com>>.
- CHIBUTA, S.; ACAR, A. C. Real-time malaria parasite screening in thick blood smears for low-resource setting. *Journal of Digital Imaging*, v. 33, p. 763–775, 2020. Disponível em: <<https://doi.org/10.1007/s10278-019-00284-2>>.
- DAVE, I. R. Image analysis for malaria parasite detection from microscopic images of thick blood smear. In: *2017 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET)*. [S.l.]: IEEE, 2017. p. 1303–1307.
- DAVE, I. R.; UPLA, K. P. Computer aided diagnosis of malaria disease for thin and thick blood smear microscopic images. In: *2017 4th International Conference on Signal Processing and Integrated Networks (SPIN)*. [S.l.]: IEEE, 2017. p. 561–565.
- DELAHUNT, C. B. et al. Automated microscopy and machine learning for expert-level malaria field diagnosis. In: *2015 IEEE Global Humanitarian Technology Conference (GHTC)*. [S.l.]: IEEE, 2015. p. 393–399.
- Ethiopian Public Health Institute. *Manual for the Laboratory Diagnoses of Malaria*. [S.l.]: Federal Ministry of Health, 2020.

- FAN, D. P. et al. Inf-Net: Automatic COVID-19 Lung Infection Segmentation from CT Images. *IEEE Transactions on Medical Imaging*, v. 39, n. 8, p. 2626–2637, 2020. ISSN 1558254X.
- FUNG, D. L. X. et al. Self-supervised deep learning model for COVID-19 lung CT image segmentation highlighting putative causal relationship among age, underlying disease and COVID-19. *Computers in Biology and Medicine*, v. 19, 2021.
- GENG, P. et al. STCNet: Alternating CNN and improved transformer network for COVID-19 CT image segmentation. *Biomedical Signal Processing and Control*, v. 93, 2024. ISSN 1746-8094.
- GÉRON, A. *Hands-On Machine Learning with Scikit-Learn and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems*. O'Reilly Media, 2017. ISBN 9781491962244. Disponível em: <<https://www.safaribooksonline.com>>.
- GOMES, E. C. de S. et al. Outbreak of autochthonous cases of malaria in coastal regions of Northeast Brazil: the diversity and spatial distribution of species of Anopheles. *Parasites Vectors*, v. 13, n. 1, p. 621, 2020. ISSN 1756-3305.
- GONG, K. et al. A multi-center study of COVID-19 patient prognosis using deep learning-based CT image analysis and electronic health records. *European Journal of Radiology*, v. 139, 2021.
- GONZALEZ, R. C.; WOODS, R. E. *Processamento Digital de Imagens*. [S.l.]: Pearson Education do Brasil, 2010. ISBN 978-85-7605-401-6.
- GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. *Deep Learning*. MIT Press, 2016. Disponível em: <<http://www.deeplearningbook.org>>.
- Google. *Google One*. 2020. Disponível em: <https://one.google.com/about?hl=pt_BR>.
- Google. *Olá, este é o Colaboratory*. 2020. Disponível em: <<https://colab.research.google.com/notebooks/welcome.ipynb?hl=pt-BR>>.
- HANIF, N.; MASHOR, M. Y.; MOHAMED, Z. Image enhancement and segmentation using dark stretching technique for plasmodium falciparum for thick blood smear. In: *2011 IEEE 7th International Colloquium on Signal Processing and its Applications*. [S.l.]: IEEE, 2011.
- Health Surveillance Secretariat. *Manual for the Laboratory Diagnoses of Malaria*. Brasília: Ministry of Health, 2009.
- HOANG-THI, T. N. et al. Deep learning for lung disease segmentation on CT: Which reconstruction kernel should be used? *Diagnostic and Interventional Imaging*, Elsevier Masson SAS, v. 102, n. 11, p. 691–695, 2021. ISSN 22115684. Disponível em: <<https://doi.org/10.1016/j.diii.2021.10.001>>.
- HUQ, A.; PERVIN, M. T. Robust deep neural network model for identification of malaria parasites in cell images. In: *2020 IEEE Region 10 Symposium (TENSYP)*. [S.l.]: IEEE, 2020. p. 1456–1459.
- ILHAN, A. et al. COVID-19 Lung CT image segmentation using localization and enhancement methods with U-Net. *Procedia Computer Science*, v. 218, p. 1660–1667, 2023. ISSN 1877-0509.
- JUN, M. et al. *COVID-19 CT Lung and Infection Segmentation Dataset*. Zenodo, 2020. Disponível em: <<https://doi.org/10.5281/zenodo.3757476>>.

KANG, M. et al. Quantitative Assessment of Chest CT Patterns in COVID-19 and Bacterial Pneumonia Patients: a Deep Learning Perspective. *Journal of Korean Medical Science*, v. 36, n. 5, p. 1–14, 2021. ISSN 1598-6357.

KAUR, J.; CHOUDHARY, A. Comparison of several contrast stretching techniques on acute leukemia images. *International Journal of Engineering and Innovative Technology (IJEIT)*, v. 2, n. 1, p. 332–335, 2012.

Keras. *Keras*. 2020. Disponível em: <<https://keras.io/>>.

KITRUNGROTSAKUL, T. et al. Attention-RefNet: Interactive Attention Refinement Network for Infected Area Segmentation of COVID-19. *IEEE Journal of Biomedical and Health Informatics*, IEEE, v. 25, n. 7, p. 2363–2373, 2021. ISSN 21682208.

KUMAR, M.; DESHPANDE, A. *Artificial Intelligence for Big Data*. Packt, 2018. ISBN 9781788472173. Disponível em: <<https://www.safaribooksonline.com>>.

LI, Z. et al. From community-acquired pneumonia to COVID-19: a deep learning-based method for quantitative analysis of COVID-19 on thick-section CT scans. *European Radiology*, v. 30, n. 12, p. 6828–6837, 2020. ISSN 14321084.

LIU, J. et al. COVID-19 lung infection segmentation with a novel two-stage cross-domain transfer learning framework. *Medical Image Analysis*, Elsevier B.V., v. 74, p. 102205, 2021. ISSN 13618423. Disponível em: <<https://doi.org/10.1016/j.media.2021.102205>>.

MA, J. et al. Toward data-efficient learning: A benchmark for COVID-19 CT lung and infection segmentation. *Medical Physics*, v. 48, n. 3, p. 1197–1210, 2021. ISSN 00942405.

MADHU, G. et al. Imperative dynamic routing between capsules network for malaria classification. *Computers, Materials and Continua*, v. 68, n. 1, p. 903–919, 2021. ISSN 1546-2218. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1546221821009292>>.

MIYAGAWA, M. et al. Lumen segmentation in optical coherence tomography images using convolutional neural network. In: *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. [S.l.: s.n.], 2018. p. 600–603.

MUTABAZI, T. et al. Assessment of the accuracy of malaria microscopy in private health facilities in Entebbe Municipality, Uganda: a cross-sectional study. *Malaria Journal*, v. 20, n. 1, p. 250, 2021. ISSN 1475-2875.

NAKASI, R.; MWEBAZE, E.; ZAWEDDE, A. Mobile-aware deep learning algorithms for malaria parasites and white blood cells localization in thick blood smears. *Algorithms*, v. 14, n. 1, p. 17, 2021. Disponível em: <<https://doi.org/10.3390/a14010017>>.

NAKASI, R. et al. A new approach for microscopic diagnosis of malaria parasites in thick blood smears using pre-trained deep learning models. *SN Applied Sciences*, v. 2, n. 7, p. 1255, 2020. ISSN 2523-3971.

NAKASI, R. et al. A web-based intelligence platform for diagnosis of malaria in thick blood smear images: a case for a developing country. In: *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. [S.l.]: IEEE, 2020. p. 4238–4244.

PAHO. *COVID-19: An Informative Guide*. 2020.

PALURU, N. et al. Anam-Net: Anamorphic Depth Embedding-Based Lightweight CNN for Segmentation of Anomalies in COVID-19 Chest CT Images. *IEEE Transactions on Neural Networks and Learning Systems*, v. 32, n. 3, p. 932–946, 2021. ISSN 21622388.

PATTANAIK, P. A.; MITTAL, M.; KHAN, M. Z. Unsupervised deep learning cad scheme for the detection of malaria in blood smear microscopic images. *IEEE Access*, v. 8, p. 94936–94946, 2020.

QUINN, J. A. et al. Automated blood smear analysis for mobile malaria diagnosis. *Mobile Point-of-Care Monitors and Diagnostic Device*, p. 115–132, 2014.

QUINN, J. A. et al. Deep convolutional neural networks for microscopy-based point of care diagnostics. In: *Machine Learning for Healthcare Conference*. [s.n.], 2016. p. 271–281. Disponível em: <<https://jquinn.air.ug/publication/2016-deep-convolutional-neural-networks-for>>.

RANJBARZADEH, R. et al. Lung Infection Segmentation for COVID-19 Pneumonia Based on a Cascade Convolutional Network from CT Images. *BioMed Research International*, v. 2021, 2021. ISSN 23146141.

RONNEBERGER, O.; FISCHER, P.; BROX, T. *U-Net: Convolutional Networks for Biomedical Image Segmentation*. arXiv, 2015. Disponível em: <<https://arxiv.org/abs/1505.04597>>.

SAOOD, A.; HATEM, I. COVID-19 lung CT image segmentation using deep learning methods: U-Net versus SegNet. *BMC Medical Imaging*, BioMed Central, v. 21, n. 1, p. 1–10, 2021. ISSN 14712342. Disponível em: <<https://doi.org/10.1186/s12880-020-00529-5>>.

SHAN, F. et al. Abnormal lung quantification in chest CT images of COVID-19 patients with deep learning and its application to severity prediction. *Medical Physics*, v. 48, n. 4, p. 1633–1645, 2021. ISSN 00942405.

Tensorflow. *Tensorflow*. 2020. Disponível em: <<https://www.tensorflow.org/>>.

WANG, G. et al. A Noise-Robust Framework for Automatic Segmentation of COVID-19 Pneumonia Lesions from CT Images. *IEEE Transactions on Medical Imaging*, v. 39, n. 8, p. 2653–2663, 2020. ISSN 1558-254X.

WHO. *Use of chest imaging in COVID-19: a rapid advice guide*. 2020. Disponível em: <<https://www.who.int/publications/i/item/use-of-chest-imaging-in-covid-19>>.

WHO. *Guideline WHO Guidelines for malaria*. 2021. 225 p. Disponível em: <<http://apps.who.int/bookorders>>.

WHO. *World health statistics 2021: monitoring health for the SDGs, sustainable development goals*. 2021. Disponível em: <<https://www.who.int/publications/i/item/9789240027053>>.

WHO. *Therapeutics and COVID-19: living guideline*. 2023. Disponível em: <<https://app.magicapp.org/#/guideline/nBkO1E>>.

WHO. *WHO COVID-19 dashboard*. 2023. Disponível em: <<https://data.who.int/dashboards/covid19/cases?n=c>>.

WHO. *Infectious diseases*. 2024. Disponível em: <<https://www.emro.who.int/health-topics/infectious-diseases/index.html>>.

WHO. *World malaria report 2024: addressing inequity in the global malaria response*. 2024. Disponível em: <https://cdn.who.int/media/docs/default-source/malaria/world-malaria-reports/world-malaria-report-2024-spreadview.pdf?sfvrsn=3ccb3695_3>.

WU, D. et al. Severity and Consolidation Quantification of COVID-19 from CT Images Using Deep Learning Based on Hybrid Weak Labels. *IEEE Journal of Biomedical and Health Informatics*, v. 24, n. 12, p. 3529–3538, 2020. ISSN 21682208.

WU, Y. H. et al. JCS: An Explainable COVID-19 Diagnosis System by Joint Classification and Segmentation. *IEEE Transactions on Image Processing*, v. 30, p. 3113–3126, 2021. ISSN 19410042.

YANG, F. et al. Deep learning for smartphone-based malaria parasite detection in thick blood smears. *IEEE Journal of Biomedical and Health Informatics*, v. 24, n. 5, p. 1427–1438, 2020. Disponível em: <<https://doi.org/10.1109/JBHI.2019.2939121>>.

YANG, F. et al. Parasite detection in thick blood smears based on customized faster-rcnn on smartphones. In: *2019 IEEE Applied Imagery Pattern Recognition Workshop (AIPR)*. [S.l.]: IEEE, 2019.

ZHANG, Q.; REN, X.; WEI, B. Segmentation of infected region in CT images of COVID-19 patients based on QC-HC U-net. *Scientific Reports*, Nature Publishing Group UK, v. 11, n. 1, p. 1–11, 2021. ISSN 20452322. Disponível em: <<https://doi.org/10.1038/s41598-021-01502-0>>.

ZHOU, L. et al. A Rapid, Accurate and Machine-Agnostic Segmentation and Quantification Method for CT-Based COVID-19 Diagnosis. *IEEE Transactions on Medical Imaging*, v. 39, n. 8, p. 2638–2652, 2020. ISSN 1558254X.

APÊNDICE A – ARTIGO PUBLICADO NA REVISTA *BIOMEDICAL SIGNAL PROCESSING AND CONTROL* EM 2022

Neste apêndice, é apresentado o artigo *A new approach for malaria diagnosis in thick blood smear images*, publicado na revista *Biomedical Signal Processing and Control* em 2022. Este é o primeiro artigo resultante deste trabalho.



Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

journal homepage: www.elsevier.com/locate/bspc

A new approach for malaria diagnosis in thick blood smear images

Anne de Souza Oliveira^a, Marly Guimarães Fernandes Costa^a, Maria das Graças Vale Barbosa^b, Cicero Ferreira Fernandes Costa Filho^{a,*}^a Center for Research and Development of Electronic Technology and Information (CETELI), Federal University of Amazonas (UFAM), Manaus, Amazonas, Brazil^b Amazonas State University, Manaus, Amazonas, Brazil

ARTICLE INFO

Keywords:

Malaria

Convolutional neural networks

Thick blood smear image

ABSTRACT

This paper presents a new approach for detecting malaria parasites in full images of thick blood smear using pixel classifiers for obtaining foreground objects and delimiting parasite-stained objects. For both processes, the HSV components were used as input variables of the following pixel classifiers: multilayer perceptron and a decision tree. The obtained patches were classified using a deep neural network with 34 layers, trained from scratch. The image dataset used was divided into sets with different parasite sizes. This enables characterizing performance metrics (accuracy, sensitivity, specificity, precision, and F1-score) for parasite detection with varying parasite sizes. The best metric values were obtained in images with large parasite sizes. For image sets 1 and 2, with large parasite sizes, precision rates of 91.71% and 93.14% were obtained. For image sets 3 and 4, with small parasite sizes, precision rates of 76.58% and 71.58% were obtained. As shown by the literature review, these results are comparable to others previously published. Nevertheless, a rigorous comparison could not be done, as different works use different datasets.

1. Introduction

The incidence and prevalence of malaria has decreased worldwide between 2000 and 2019. However, malaria continues to cause unacceptably high levels of disease and death. According to the World Health Organization (WHO) [1], in 2000 the prevalence was 238 million cases, while in 2019 it was 229 million cases. The incidence (i.e., cases per 1000 population at risk), decreased from 80 in 2000 to 57 in 2019.

Globally, between 2000 and 2019, the number of deaths decreased from 736,000 in 2000 to 409,000 in 2019. Although malaria affects all age groups, in 2000, 84% of cases occurred in children under 5 years old, and 67% in 2019 [1].

Around 95% of malaria cases occurred in just 31 countries. The following countries accounted for 51% of malaria cases in the world: Nigeria (23%), the Democratic Republic of the Congo (11%), the United Republic of Tanzania (5%), Mozambique (4%), Niger (4%) and Burkina Faso (4%) [1].

In the Americas, between 2000 and 2019 the incidence of malaria decreased by 40%. Progress in the region in recent years has been compromised by the increase in cases in Venezuela, which in 2000 had 35,500 cases, while in 2019, it had 467,000 cases. Brazil, Colombia, and Venezuela account for 85% of cases in the region. In Brazil, 99.7% of

malaria is registered in the Amazon region [2], where environmental and sociocultural conditions favor the expansion of its transmission.

Although, in low-income countries, the incidence, prevalence, and number of deaths in the world due to malaria has decreased between 2000 and 2019, malaria is still the 6th most frequent cause of death [3]. It also represents a high risk for travelers and migrants, with imported cases in nonendemic areas. Because of this, the WHO recommends that its early diagnosis and rapid treatment should be the first building blocks established in any control program.

Malaria is a potentially fatal disease, caused by parasites transmitted to people through the bite of infective female *Anopheles* mosquitoes. Malaria symptoms and signs are not specific. Clinical suspicion of malaria is based on the occurrence of fever or a history of fever. Diagnosis based on clinical assessment alone is not very specific and results in overtreatment. The goal in malaria diagnosis has to be to identify patients who actually have malaria in order to rationally guide the use of antimalarial medication. At present, the three most commonly used methods for diagnosing malaria are polymerase chain reaction (PCR), rapid diagnostic test (RDT) and light microscopy [4]. Currently, molecular diagnostic tools based on nucleic acid amplification techniques (e.g. loop-mediated isothermal amplification or PCR) are not used in the clinical management of malaria. The two methods most frequently used

* Corresponding author.

E-mail address: ccosta@ufam.edu.br (C. Ferreira Fernandes Costa Filho).<https://doi.org/10.1016/j.bspc.2022.103931>

Received 17 January 2022; Received in revised form 12 April 2022; Accepted 18 June 2022

Available online 1 July 2022

1746-8094/© 2022 Elsevier Ltd. All rights reserved.

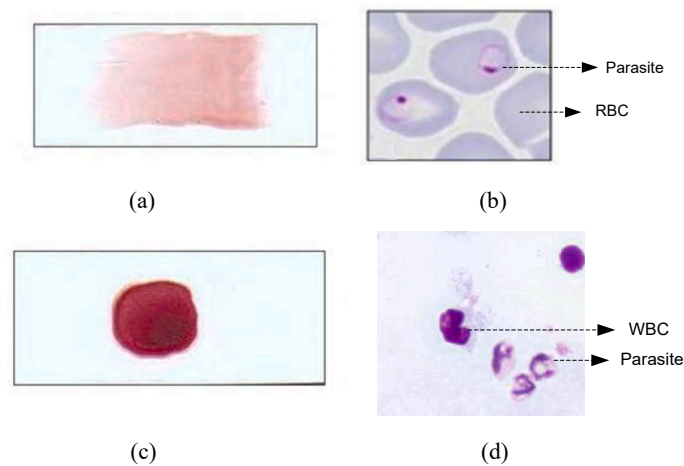


Fig. 1. (a) Thin blood smear slide (b) microscopic image of thin blood smear; (c) thick blood smear slide; (d) microscopic image of thick blood smear.

for the parasitological diagnosis of malaria are RDT and light microscopy. The first detects parasite-specific antigens or enzymes that are either genus or species specific. Light microscopy is still considered the “field standard” against which the sensitivity and specificity of other methods must be assessed. Light microscopy has the following advantages over other diagnostic methods: high sensitivity; low direct costs, if laboratory infrastructure to maintain the service is available; differentiation of *Plasmodia* species; determination of parasite densities and allows monitoring of responses to therapy [4]. However, training and managing specialized microscopists using light microscopy is costly, and diagnostic accuracy is heavily dependent on the expertise of the microscopist. This is even more true in low-income countries. In a study to assess the accuracy of malaria microscopy in private health facilities in Entebbe, Uganda [5], the authors concluded that the accuracy in diagnosing malaria using light microscopy was high. However, one-third of the patients diagnosed with malaria did not have the disease. Most mistakes were made by microscopists with little experience.

The microscopic diagnosis of malaria is based on counting the number of parasites (parasitemia) in stained blood slide. There are two types of microscopic exams: thin film and thick film. In the first one, a drop of blood is spread over the glass slide, forming a thin film, as shown in Fig. 1(a). In the second, a drop of blood is spread over a smaller surface, between 1.2 cm^2 and 1.5 cm^2 , as shown in the Fig. 1(c) [6,7].

Before the films are evaluated by microscopy, they are stained to highlight the malaria parasite. In the thin film preparation, after blood deposition, the slide is fixed with methyl alcohol for one minute and

then stained with Giemsa dye, in the proportion of one drop of dye to 1 ml of buffered water. Fig. 1(b) shows the appearance of the thin film under the microscope at $1000\times$ magnification. In this figure, a red blood cell (RBC) and parasites inside RBCs are shown. Other blood elements, such as white blood cells and platelets can also be found in thin blood images. In thick film preparation, the slides are initially subjected to a hypotonic methylene blue solution for deglobalization. Then, the slide is stained for 10 min in a solution of Giemsa stain. Fig. 1(d) shows the appearance of the thick film under the microscope at $1000\times$ magnification. In this figure, a white blood cell (WBC) and malaria parasites are shown. Differing from thick film, the parasites are not inside RBC cells [7].

Due to the fixation of the red blood cells and, therefore, keeping the shape of the parasites in their interior unaltered, the thin film method allows better examination of their morphology. In thick blood film, the blood elements (including parasites, if any) are more concentrated ($20\sim 30\times$) than in an equal area of a thin blood film, allowing a greater volume of blood to be examined. Because a larger volume ($6\text{ }\mu\text{l}$) of blood is examined in the thick film, it is more sensible to use than thin blood film, and is much better than the thin blood film for detecting low levels of parasitemia [6,7]. Thus, the thick blood film exam is performed to detect and quantify (parasite density) malaria parasites in routine malaria microscopic diagnosis.

There are five types of plasmodium protozoa that infect humans: *Plasmodium falciparum*, *Plasmodium vivax*, *Plasmodium ovale*, *Plasmodium malariae* and *Plasmodium knowlesi* [6]. Within RBCs, depending on their

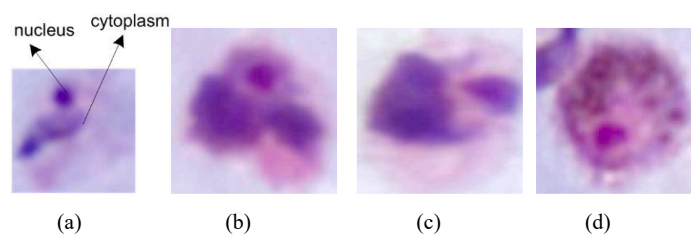


Fig. 2. Examples of parasites stained with Giemsa dye in thick blood smear. (a): ring or immature trophozoite; (b) ameboid or mature trophozoite; (c) schizont; (d) gametocyte.

stage of development and specific characteristics, the parasites can take several forms and sizes (length between 1 μm and 5 μm or between 40 μm and 80 μm long): ring or immature trophozoite; amoeboid or mature trophozoite; schizont and gametocyte. Fig. 2 shows examples of parasites stained with Giemsa dye in thick blood smear of *P. vivax* parasites. From Fig. 2, we drew some conclusions about the problem of malaria parasite detection in thick blood smear:

- Depending on the reproductive stage, the parasite has different sizes. The smaller ones are the immature trophozoites. The larger sizes are the gametocytes.
- A parasite can be made up of several stained objects (e.g. nuclei + cytoplasm).
- A parasite can take many forms.

These different characteristics represent a challenge in the automatic detection of malaria parasites.

2. Literature review

In the last decade, to mitigate the problems arising from the lack of skilled labor for diagnosing malaria with bright microscopy, a huge effort has been made in the research and development of automatic or semiautomatic solutions for diagnosing malaria using thick blood film: [8–16].

Below follows a review of trends observed in the works published in the literature in the last decade, on automatic diagnosis of malaria using thick blood films.

Regarding the techniques used to identify the parasites, three different groups are identified. In the first group, only digital image processing techniques are used to detect the parasites [17–20]. The second group is characterized by two moments. At first, characteristics are extracted from objects. In a second moment, these characteristics are used in classifiers such as support vector machines, perceptron multilayer, k-means, which differentiate parasites from other structures, such as white blood cells, platelets, and noise [21,22,16]. Finally, the third group is characterized by using deep learning techniques [9,13,23,10,8,11,12,14,15,24,25].

The appearance of microscopic images of thick blood films may present different characteristics in terms of color, based on the pH of the Giemsa stain [22], such as lighting, due to different illumination conditions used [26]. To preserve uniformity in the images of thick blood films, it is common, in the works of the first and second groups, to use image processing techniques.

In [26], the captured images are converted in HSV color space to reduce variations due to brightness and concentration of stain solution, which, according to the authors, highly affects images in RGB color space. In [16], to contrast enhancement of images, the Modified Global Contrast Stretching (MGCS) algorithm is used, which highlights and sharpens the thick blood smear image quality. In the sequence, for color normalization, a gray world algorithm was used. In [17], the authors improve the image applying contrast stretching [27] and dark stretching [28].

In the papers of group 1, the malaria parasites are detected using only segmentation techniques borrowed from digital image processing, followed by filtering operations to remove noise. In [17], the parasites are segmented using the technique of Active Contour Without Edge (ACWE), a modified version of the active contours technique, also known as snakes. In the sequence, the authors apply morphological filters to remove noise and, finally, a thresholding operation. In a dataset of 50 images with 140×140 pixels, the authors reported an accuracy of 97.57% and a false positive rate of 12.04%. In [18], the authors used 87 microscopic fields with 3136×2352 pixels. Each microscopic field is captured at 10 different focus planes. The ground truth is obtained through a manual annotation of the x-y coordinates of parasites. As a pre-processing stage, noise filtering is done using maximum Laplacian

filter values obtained from 10 different focus planes. For detecting parasites, the authors first separate the foreground (stained particles i.e. parasites, white blood cells (WBCs), platelets and noise) from the background (nonstained particles) through adaptive thresholding of the saturation channel histogram. After, the nonparasite objects of the foreground are eliminated using the area feature of connected components. According to the authors, the maximum area of parasites is 700 pixels. Using this area threshold, objects other than parasites are eliminated. The authors reported an accuracy of 90.81% and a false positive rate of 10.12%.

In the papers of group 2, the following operations are employed to detect malaria parasites: segmentation + characteristic extraction + classifier. In [26], stained particles, i.e., parasites of all life stages, WBCs, platelets and many other artifacts are segmented from the background, using adaptive histogram thresholding technique of the saturation channel of the HSV space. In the sequence, for each ROI, 123 features are extracted (central moments, Hu's invariant moments, Zernike's moments, area, etc.) and submitted to a support vector machine. In a dataset of 87 images of microscopic fields, with 3136×2352 pixels, collected from distinct patients, with variation in staining solution concentration and parasite density, the authors reported a sensitivity of 86.34% and a specificity of 96.60%.

In [16], the thick blood smear images were segmented and classified using fast k-means (FKM) clustering, using 15 color components of the following spaces: RGB, CMY, XYZ, HSV and YUV. In FKM algorithm, a combination of k-means algorithm with level histogram discrete function value is employed for determining cluster centers of the database training. The extracted color components are clustered into three categories, which are WBCs and malaria parasites (black), staining precipitation (gray) and background (white). The watershed algorithm is used to separate connected parasites, and a noise filter discards objects smaller than 7 pixels and larger than 200 pixels (WBCs). Using a dataset of 100 images with 1294×980 pixels, the authors reported an accuracy of 99.74%, a sensitivity of 77.30% and a specificity of 99.76%.

In the papers of group 3, deep learning techniques are used for patch classification [9,13,23], image classification [10], or for parasite detection and classification in full images [8,11,12,14,15,24,25].

One approach for parasite detection and classification in full images involves obtaining patches and classifying them with deep neural networks. For example, in [13] the authors use Otsu's threshold method for WBC filtering in grayscale images, iterative global minimum screening (IGMS) for obtaining patches (generate parasite candidates by localizing minimum intensity values in a grayscale image) and CNN for patch classification. In a dataset with 1,819 thick blood smear images and 84,961 annotated parasites, the authors reported an accuracy of 97.26%, a sensitivity of 82.73%, a specificity of 98.39% and a precision of 78.98%.

Another approach for parasite detection and classification in full image, is using the state-of-the-art generic object detection algorithms, such as you only look once (YOLO) [11,12], single short detection (SSD) [8,11,14], faster regional convolutional neural network (Faster RCNN) [8,11,14,24], Retina Net [8], and Refine Det [8].

In [11], the authors stated that YOLOv3 and YOLOv4 models are not optimized for detecting small objects such as malaria parasites in microscopic images, so they modified these models by increasing feature scale and adding more detection layers to enhance their capability of detecting small objects without notably decreasing detection speed. In this work the authors used a database of 1182 RGB images with 750×750 pixels, captured with a smartphone attached on the microscope eyepiece. The ground truth is obtained with bounding boxes around the malaria parasites. The main results obtained by with YOLOv4 CNN are the following: Mean Average Precision – 96.32, Precision – 95, Recall – 94, F1-Score: 94, Average IOU – 62.12 and FPS – 29.60.

In [8], with RetinaNet, and using 643 images with 750×750 pixels, the authors reported an accuracy of 86.97% and a recall of 60.86%. In [12], with YOLOv3, in a dataset of 2703 images with 1024×768 pixels,

Table 1
Summary of works published in the last decade concerning malaria parasite detection in thick blood smear.

Reference	Database	Pre-Processing	Ground Truth	Segmentation/Classifier	Performance evaluation	Results (%)
[8]	643 images with 750×750 pixels, captured with a smartphone attached on the microscope eyepiece	Data augmentation through random horizontal and vertical flipping.	Ground truth: bounding boxes around the malaria parasites.	Pre-trained deep learning models: faster-R_CNN (ResNet50, ResNet101, Inception V2), RetinaNet (SSD ResNetFPN), SSD-MobileNetV2, SSD-InceptionV2.	Parasite detection in full image	faster-R-CNN(ResNet50): Precision:77.91, Recall: 89.81; RetinaNet: Precision: 86.97, Recall:60.86; SSD-InceptionV2: Precision:91.50, Recall: 37.53
[9]	27,558 cell images: 13,779 images being normal cell images and 13,779 images being infected with malaria parasites. This dataset has been collected from the NIH.	Reshape images to a resolution of 128×128 pixels	Ground truth: positive erythrocytes and Negative erythrocytes manually annotated.	Capsule Networks	Patch classification	Accuracy: 98.7 ± 0.108 , AUC-ROC: 98.73 ± 0.105 , Sensitivity: 98.50 ± 0.109 Specificity: 98.97 ± 0.189
3	1182 RGB images with 750×750 pixels, captured with a smartphone attached on the microscope eyepiece.	Not applied	Ground Truth: images classified as malaria-infected or non-infected.	Functional Link Artificial Neural Network (FLANN) and Stacked Sparse Autoencoder (SSAE).	Full image classification	Accuracy: 89.10
[11]	1182 images with 750×750 pixels, captured with a smartphone attached on the microscope eyepiece.	Not applied	Ground Truth: bounding boxes around the malaria parasites.	Faster RCNN, SSD, YOLOV3 CNN and YOLOV4 CNN	Parasite detection in full images	YOLOV4 CNN: Mean Average Precision: 96.32, Precision: 95, Recall: 94, F1-Score: 94, Average IOU: 62.12, FPS: 29.60. Faster RCNN: Precision: 92.7; Recall: 86.9F1-score: 89.71
[12]	Dataset A: 2703 images with 1024×768 pixels, obtained from a camera attached on the microscope.Dataset B: 1182 images with 750×750 pixels, captured with a smartphone attached on the microscope eyepiece.	Not applied	Ground truth: bounding boxes around the malaria parasites.	YOLOV3 CNN	Parasite detection in full images	Dataset A: Mean Average Precision: 88.70. Dataset B: Mean Average Precision: 90.02
[13]	Image set acquired from 150 patients, including 1819 thick smear images and 84,961 annotated parasites	Not applied	Ground truth: manual annotation of the x-y coordinates of parasites.	Otsu's method for WBC filtering, Iterative Global Minimum Screening (IGMS) for parasite segmentation and CNN for parasite classification	Parasite detection in full image	Accuracy: 97.26, Sensitivity: 82.73, Specificity: 98.39, Precision: 78.98%, AUC: 97.34
[14]	903 images with 3264×2448 pixels.	Data augmentation through random vertical and horizontal flipping	Ground truth: bounding boxes drawn around parasites and around WBCs.	Faster R-CNN and SSD MobileNet	Parasite detection in full image	Mean Average Precision: 55.06, Precision: 67.20, Recall: 80.20, F1-Score:73.12
[15]	1182 RGB images of 7245 objects, captured with a smartphone attached on the microscope eyepiece.	Conversion of malaria trophozoite and WBC labels to TF-Record format	Ground truth: bounding boxes around the malaria parasites.	CNN with a sliding window with the non-maximum-supression algorithm.	Parasite detection in full image	AUC = 100
[16]	100 images with 1294×980 pixels.	Modified Global Contrast Stretching (MGCS) algorithm. Color normalization using Gray world algorithm.	Ground truth: manual annotation of the x-y coordinates of parasites.	fast k-means algorithm applied in RGB, HSV, CMY and YUV space color components.	Parasite detection in full image	Accuracy: 99.74, Specificity: 99.76, Sensitivity: 77.30
[17]	2967 images with 4032×3024 pixels, from 200 patients, captured with a smartphone attached on the microscope eyepiece.	Not applied	Ground truth: manual annotation of the x-y coordinates of parasites.	Faster-RCNN.	Parasite detection in full image	Sensitivity: 96.84
[18]	50 images with 140×140 pixels (at $1000 \times$ optical magnification).	noise filtering (median filtering, gaussian filtering and lowpass) and image enhancement (contrast stretching and dark stretching).	Ground truth: manual annotation of the x-y coordinates of parasites.	Digital image processing tools for parasit detection: active contour without edge, then erosion, dilation, masking, contrast stretching, and thresholding.	Parasite detection in full image	Accuracy: 97.57, False Negative Rate: 12.04
[19]	87 microscopic fields with 3136×2352 pixels. Each microscopic field is captured at 10 different focus planes.	noise filtering using maximum Laplacian filter values obtained from 10 different focus planes.	Ground truth: manual annotation of the x-y coordinates of parasites.	Segmentation: threshold of the saturation histogram. Classifier: area threshold to separate parasites from white blood cells and platelets.	Parasite detection in full image	Accuracy thick smear: 90.81, Accuracy thin smear: 97.87
[20]		Colour processing not detailed	Ground truth: manual	Support vector machines, neural networks and decision		Sensitivity: 100, Specificity: 95%

(continued on next page)

Table 1 (continued)

Reference	Database	Pre-Processing	Ground Truth	Segmentation/Classifier	Performance evaluation	Results (%)
[21]	183 positive slides and 98 negative slides of 7 locations worldwide. A total of 27,558 images are available, with 13,779 images being normal cell images and the rest of 13,779 images being infected with malaria parasites. This dataset has been collected from the NIH.	All the images have been rescaled to 224 X 224.	annotation of the x-y coordinates of parasites. Ground truth: positive erythrocytes and negative erythrocytes manually annotated.	trees with morphological, colour and texture features of candidates. VGG16 with adversarial training for classification of positive erythrocytes and negative erythrocytes manually annotated	Parasite detection in full image Patch classification	Accuracy: 95.9; Precision: 96.78%; Recall: 95.10%; F1-Score: 95.59%.
[22]	643 images with 750 × 750 pixels captured with a smartphone attached on the microscope eyepiece..	Data augmentation: random horizontal and vertical flipping to randomly selected images.	Ground truth: bounding boxes around the malaria parasites.	Faster RCNN	Parasite detection in full image	Accuracy: 93.06; Precision: 74.61; Recall: 90.62.
[23]	87 images of microscopic fields, with 3136 × 2352 pixels, collected from distinct patients, with variation in staining solution concentration and parasite density.	HSV conversion	Ground truth: manual annotation of the x-y coordinates of parasites.	Segmentation: threshold of the saturation histogram. Classifier: Support vector machine using 123 features: central moments, Hu's invariant moments, Zernike's moments, area, etc.	Parasite detection in full image	Sensitivity: 86.34, Specificity: 96.60
[24]	36 images of microscopic fields, with 1600x1200 pixels. The average number of Plasmodium parasites in each image was 15 parasites	Contrast stretching and median filtering.	Ground truth: manual annotation of the x-y coordinates of parasites.	Parasites are detected using a bottom-hat morphological operation, adaptive entropy thresholding and active contour without edge methods.	Parasite detection in full image	Sensitivity: 98.04, Accuracy: 61.53
[25]	2703 images from 133 individuals, with 50,255 parasites, obtained with a Motic MC1000 microscope camera.	Not applied	Ground truth: bounding boxes around the malaria parasites.	Segmentation: threshold of gray level images. Classifier: Extremely Randomized Trees classifier using morphological features and moments.	Parasite detection in full image	Precision: 90%, Recall: 20%

the authors reported an average mean accuracy of 88.70%.

In recent years, with the argument for accelerating diagnosis in remote locations with few resources, some works have published articles with applications aimed at mobile devices. The mobile device is attached to the microscope objective using a plastic fixture. The thick film images are then acquired and processed in the mobile devices [8,12] In another work, the authors developed an application for diagnosing malaria based on a WEB application. In this case, images are acquired locally and processed remotely. The result is then sent back to the image collection point [25].

Table 1 shows a summary of some papers published in the last decade concerning malaria parasite detection in thick blood smear.

In the papers previously presented that do not use generic object detection algorithms, we observe the following trends: a) the segmentation of foreground objects (platelets, WBCs, parasites and some artifacts) is usually affected by applying thresholds to saturation channel histograms [18–26] and thresholds to gray-level values [13,20] b) as shown in the introduction, in thick blood smear images, the parasites are usually made up of one or more stained objects. In order not to have multiple object detections (false positives), the authors usually adopted procedures such as morphological closing operations [19], evaluation of a neighborhood of a detected object [13], etc.

In this work we propose a new approach to malaria parasite detection classified as group 3, using patch classification, with the following contributions:

- 1) Segment foreground objects using pixel classifiers, as MLP, decision trees of logistic regression, and HSV components.
- 2) Delimit a parasite patch (objects that made up a parasite) using pixel classifiers, as MLP, decision trees of logistic regression, and HSV components.

Table 2

Number of parasites in each set.

Number of parasites				
Image set				
Statistic	1	2	3	4
mean	18.85	14.18	12.76	5.21
std	4.99	4.98	4.27	3.47

- 3) Separate malaria parasite from other foreground objects, form WBCs, platelets and artifacts, using a convolution neural network (CNN) trained from scratch.
- 4) Evaluate the malaria parasite detection performance with image sets containing different parasite sizes.

3. Materials and methods

3.1. Materials

The database consists of 4 sets of 169 images, totaling 676 images of size 1388 × 1040 pixels of thick blood smear films. Each slide corresponds to a different patient. The image acquisition was performed at the Mycobacteriology Lab from the Instituto Nacional de Pesquisa da Amazônia–INPA, Brazil.

The images were captured on a ZEISS Axio Imager M2 microscope, with the software ZEN 2 Pro at 1000 × magnification, from four slides containing *Plasmodium vivax* provided by Fundação de Medicina Tropical, Dr Heitor Vieira Dourado. Each set of images was taken from a different patient. Table 2 shows the mean number of parasites per image in each set. As shown in Table 2, the parasite density in sets 1 and 2 is higher than in sets 3 and 4. The patient corresponding to set 1, with an

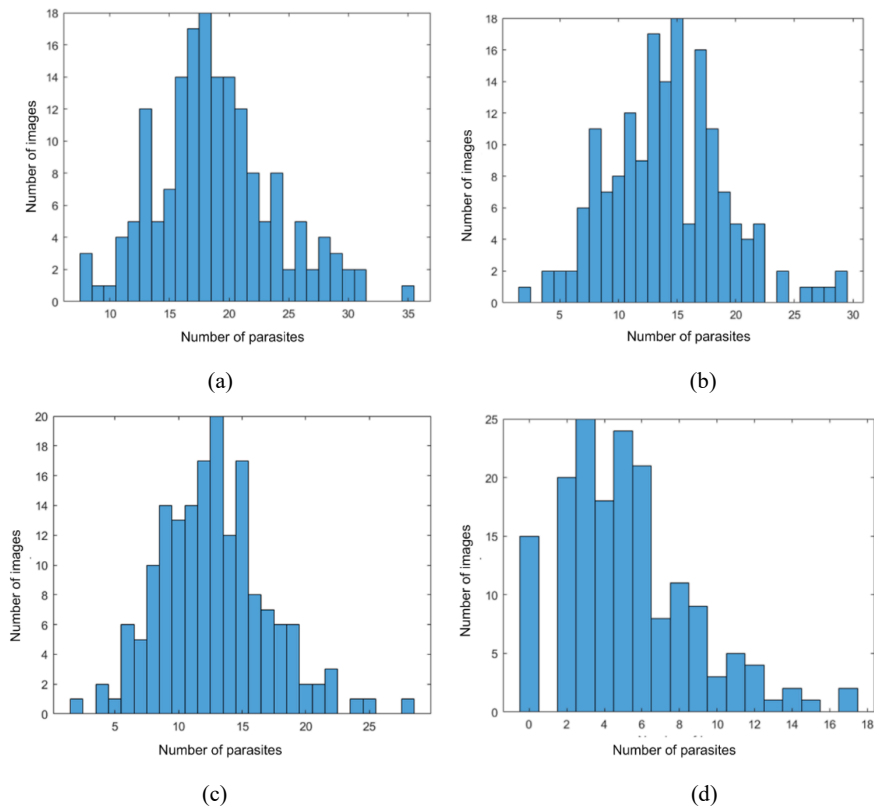


Fig. 3. Histograms of number of parasites per image in each set. (a) set 1; (b) set 2; (c) set 3 and (d) set 4.

average of 18.85 parasites per slide, corresponds to the highest level of infection, while the patient corresponding to set 4, with an average of 5.21 parasites per slide, corresponds to the lowest level of infection. Fig. 3 shows histograms of the number of parasites per image in each set. As shown, the set with a higher parasite density is set 1, while the one with a lowest parasite density is set 4.

3.2. Methods

Fig. 4 shows a block diagram of the new approach adopted in this work to detect malaria parasites in thick blood film slides.

The proposed method uses deep learning for classification of selected patches, containing objects from the foreground. These objects are stained similarly when applying Giemsa dye. Originally, we proposed that the segmentation of these objects be done through pixel classifiers, using color components from the HSV space.

As explained in the previous section, a problem that arises in detecting parasites, is that, not infrequently, they are composed of more than one stained object. Thus, it is necessary to group these objects in a single patch, so as not to generate false positives. In the current work, a second pixel classifier also uses the HSV components to detect the region near foreground objects. Although the deglobalization process breaks down the RBCs, a region with a different hue from the background still involves the malaria parasites. Fig. 5 shows an HSV image of a stained thick blood film, showing some malaria parasites and WBCs. As seen, a

region, with HSV characteristics different from the background, involves the parasites and WBCs. Also, in the image, there are other regions with the same characteristics, formed by rest of the RBCs.

In the proposed method, shown in Fig. 4, the patches are formed after a fusion process from the results of the two-pixel classifiers: classifier of foreground objects and classifier of regions near foreground objects. In this fusion process, an intersection is made between the foreground objects and the regions detected by the second classifier.

In the development of this work, it was observed that the size of patches containing parasites vary greatly in size. Normalizing all patches to the same size causes a lot of distortion. Thus, two distinct patch classification CNNs were trained. The first one was trained with patches smaller than 1600 pixels, which corresponds to immature trophozoites. The second, with patches larger than 1600 pixels, which corresponds to ameboid or mature trophozoites, schizonts and gametocytes, as well as to WBCs and platelets, which are usually larger than parasites [18].

The following presents each block shown in Fig. 4.

3.2.1. Detecting foreground objects

Fig. 6 shows HSV component values of the three types of pixels found in thick blood films: foreground object pixels, pixels of regions near foreground objects and the rest of the RBCs, and background pixels. We extracted information from 10,000 pixels of foreground objects (5,000 from parasites and 5,000 from WBCs and platelets), 5,000 pixels of regions near foreground objects and of 5,000 pixels from background.

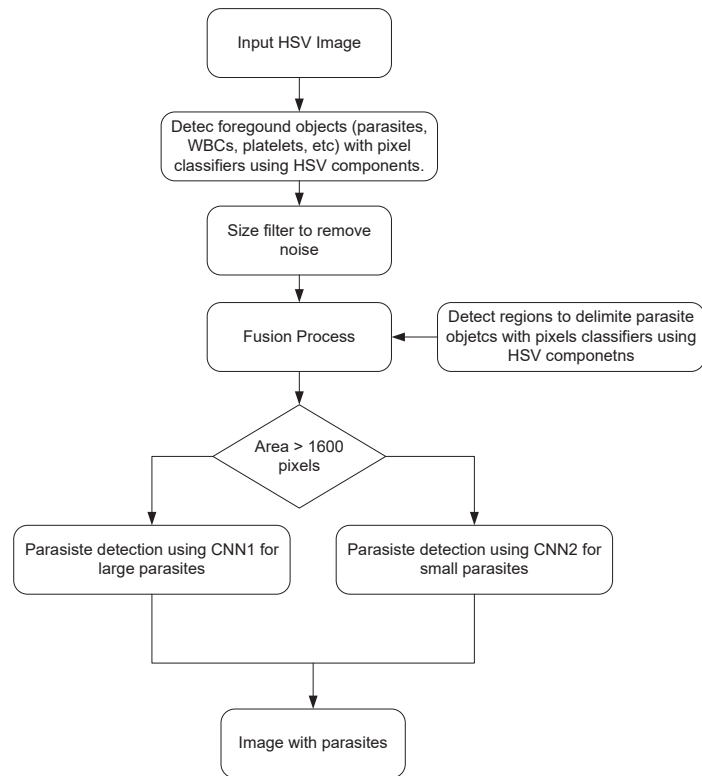


Fig. 4. Block diagram of the new approach proposed for detecting malaria parasites in thick blood film slides.

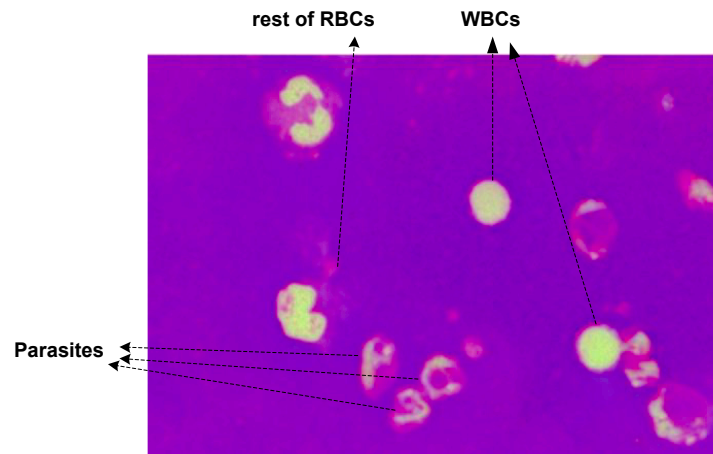


Fig. 5. HSV image of a thick blood film showing foreground objects involved in regions with a hue different from the background. Also, the rest of RBC cells are shown.

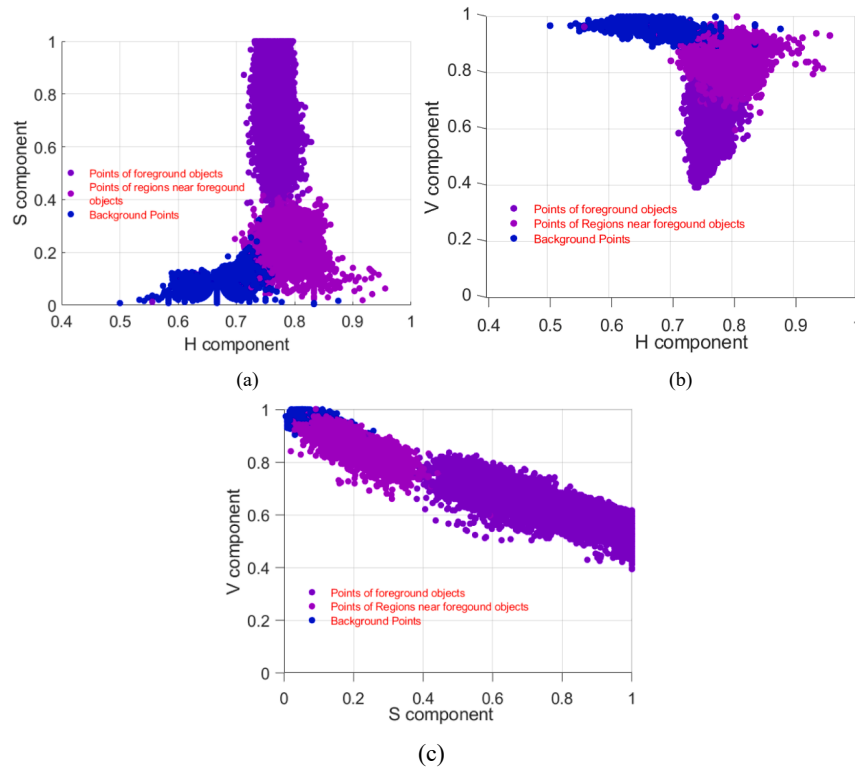


Fig. 6. Graphs of HSV pixels of foreground objects, regions near foreground objects and background: (a) S $\times$ H component; (b) V $\times$ H component and (c) V $\times$ S component.

From each image set, a subset with 5 images was used for extracting these pixels. As shown in Fig. 6, the foreground pixels have high saturation and low intensity. The background pixels have high intensity, low saturation, and low hue. The pixels of regions near the foreground objects have high hue and low saturation.

To segment the foreground objects, we evaluated two binary classifiers, a multilayer perceptron (MLP) and a decision tree (DT). The performance of both classifiers is shown in the results section. The input of both classifiers are the H, S and V components of HSV color space.

3.2.2. Detecting regions to delimit parasite objects

Fig. 7 shows the profiles of components H, S, and V of the pixels of a region near a parasite + parasite (R + P), versus the background (B). As shown, the first group of pixels has low intensity and high saturation and hue, when compared to background pixels.

To segment the region to delimit parasite objects, we also evaluated two binary classifiers, a multilayer perceptron (MLP) and a decision tree (DT). The input of both classifiers are the H, S and V components of HSV color space. The results will be presented in the results section.

3.2.3. Fusion process

The fusion process consists of creating the intersection of the foreground objects with the regions detected near them. Only the regions that intersect a foreground object are maintained. Fig. 8 illustrates this procedure. Fig. 8(a) shows an original thick blood image. Fig. 8(b) shows the foreground objects detected using the classifier described in

section 3.2.1. Fig. 8(c) shows the regions detected near the foreground objects using the classifier described in section 3.2.2. Fig. 8(d) shows the fusion image. Fig. 8(e) shows the patches in the original image corresponding to the regions of the fusion image. These patches are the ones submitted to a CNN that separates malaria parasites from other foreground objects. Fig. 7(f) shows some patches magnified. As shown, although some parasites have more than one stained region, they are all included in the same patch, minimizing false positives. The patch in the lower line and right column is a WBC.

3.2.4. Parasite detection

The patches shown in Fig. 8(e), corresponding to foreground objects, are submitted to a CNN that discriminates malaria parasites from WBCs, platelets, and noise. The patches for training the CNNs were extracted from 564 images, while 92 images (23 images of each image set) were used for evaluating the method of parasite detection in full images.

The size of patches varies widely. There are patches with areas smaller than 700 pixels and patches with areas larger than 3500 pixels. Fig. 9 shows graphs of the number of patches with areas smaller than 1600 pixels (40×40) and with areas larger than 1600 pixels in each image set. Table 3 shows the mean value of patch size in all three sets. As noted, in sets 1 and 2, there are larger patches, while in sets 3 and 4 there are smaller patches. Sets 1 and 2 correspond to patients containing more parasites in advanced developmental stages (schizonts and gametocytes), with larger sizes, while sets 3 and 4 correspond to patients containing more parasites in less advanced developmental stages (ring or

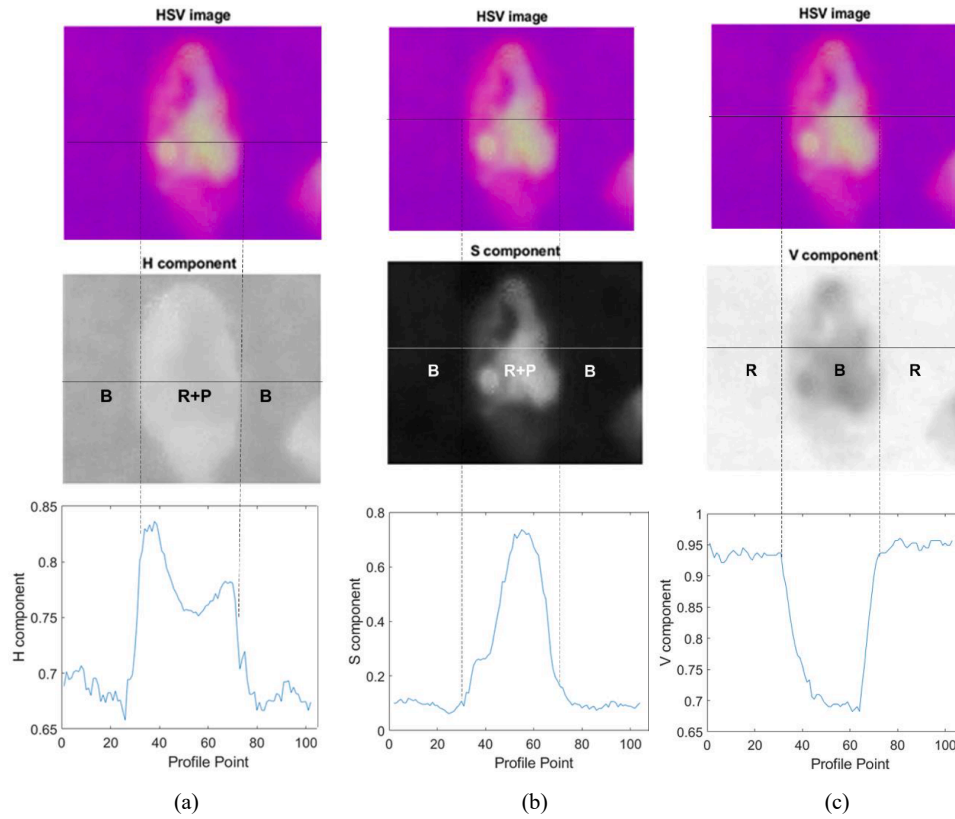


Fig. 7. Profiles of components H, S, and V of the pixels of a region near a parasite + parasite (R + P), versus the background (B). (a) H profile; (b) S profile and (c) V profile.

immature trophozoites). The choice of these sets of images allows comparing the CNN performance with larger and smaller patches.

For CNN's input, it was necessary to standardize the patch sizes. We adopted two standardization approaches. In the first one, a patch with area < 1600 pixels is resized to 40×40 pixels and a patch with area greater than 1600 pixels is resized to 100×100 pixels. To classify the smaller patches, a CNN1 network was used. To classify larger squares, a CNN2 network was used. The difference between CNN1 and CNN2 is only in the first layer, which receives images of different sizes. In the second approach, all patches are resized to 100×100 pixels, and CNN2 network is used. The reason for evaluating two different patch sizes is that we noticed that when smaller patches are resized to 100×100 pixels, it gets very distorted, so, we want to compare the performance with different patch sizes.

The number of positive and negative patches used for training and validating CNN1 (small patches) and CNN2 (large patches) is shown in Table 4. Eighty-five percent of the patches were used for training, while 15%, for validation.

The architecture of both CNN1 and CNN2 are shown in Fig. 10. There are 34 layers and 4 blocks of characteristic extraction, composed of CONV-BATCH-RELU-CONV-BATCH-RELU layers. In all blocks the kernel size was adjusted to [3, 3]. The depth of each CONV layer in block 1, block 2, block 3 and block 4 are 64, 128, 256 and 512, respectively.

According to [29], the batch normalization layer is a recent method to address the vanishing and exploding gradient problems, which cause activation gradients in successive layers to either reduce or increase in magnitude. Also, this layer can reduce the effect of the internal covariate shift, which affects the training of deep neural networks.

To reduce the map representation three maxpooling layers are inserted between the blocks of characteristic extraction. The maxpooling layers extract maximum values from an area of 2×2 pixels, with a stride of [2, 2]. After each maxpooling operation the map representation is divided by 2.

At the end of the fourth extraction characteristic block, to improve generalization, there is a dropout layer, with a dropout factor of 0.5.

Several optimization methods were evaluated: root mean square propagation (RMSProp), stochastic gradient descent with momentum (SGDM) and ADAM. The best results were obtained with the ADAM method. The number of epochs was fixed in 100. With the ADAM method, an initial learn rate of 3×10^{-4} and a squared gradient decay factor of 0.99 were employed. The minibatch size was 128.

3.2.5. Experimental Setup

The simulations were done in the MATLAB environment, on a computer with a processor intel Core i5-9300H CPU 2.40 GHz, 8 GB RAM, graphical unit processor NVIDIA GeForce GTX 1650 (4 GB RAM).

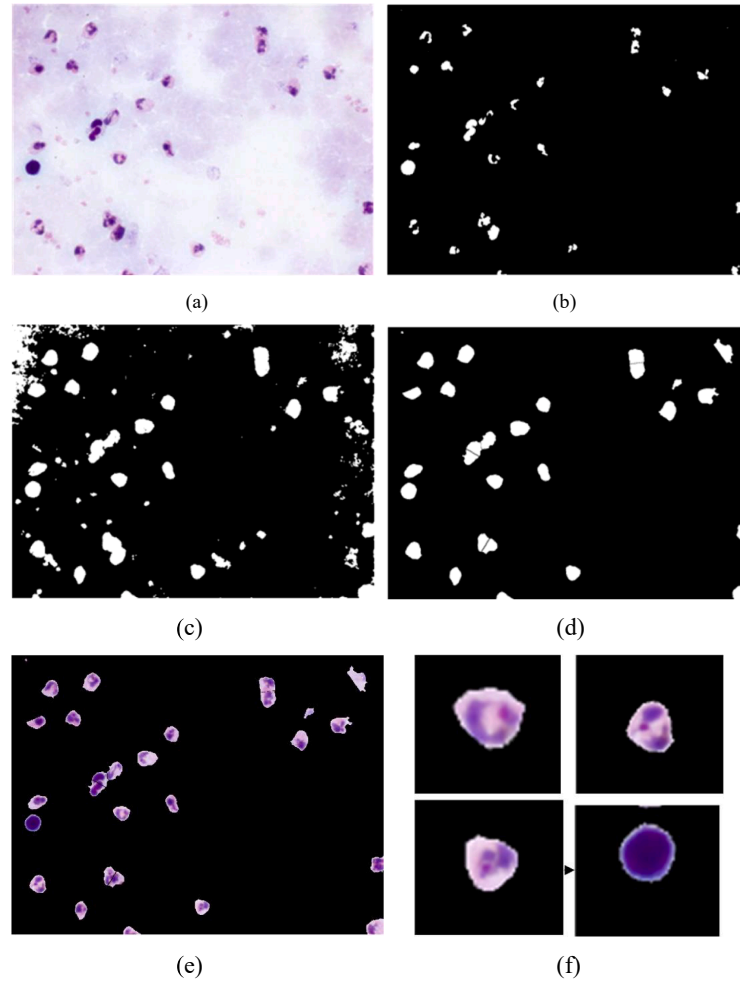


Fig. 8. (a) A thick blood smear image; (b) result of the foreground detection; (c) result of extraction of regions near foreground objects; (d) fusion image; (e) corresponding patches in original image; (f) some patches magnified.

4. Results

In this section, four types of results are presented: results for detecting foreground objects, results for detecting regions near foreground objects, results for classifying parasite patches and, finally, results for detecting malaria parasites in full image.

4.1. Results for detecting foreground objects

To detect foreground objects, the architecture of the MLP was $3 \times 4 \times 1$. The data was divided into 70%, 15%, 15% for training, validation, and testing, respectively. The stop criterium 16 epochs. Fig. 11 shows the mean squared error performance during the training. A 10-fold cross-validation loss of 0.6% was obtained. Concerning the tree classifier, we simulate trees with different values of minimum leaf size. Fig. 12 (a) shows the classification error as a function of minimum leaf size. A

minimum leaf size of 45 was used in the optimal tree. The optimal tree is shown in Fig. 12(b) (variable x_2 corresponds to S component of HSV space). A 10-fold cross-validation loss of 0.78% was obtained.

The MLP classifier was selected for segmenting the background objects, since it has the lowest loss value.

4.2. Results for detecting region near foreground objects

To detect the region near the foreground objects, the MLP architecture was also $3 \times 4 \times 1$. The data was also divided into 70%, 15%, 15% for training, validation, and testing, respectively. The stop criterium was 15 epochs. A 10-fold cross-validation loss of 0.95% was obtained in the testing set. Concerning the tree classifier, we simulate trees with different values of minimum leaf size. A minimum leaf size of 45 was used in the optimal tree. A 10-fold cross-validation loss of 1.33% was obtained.

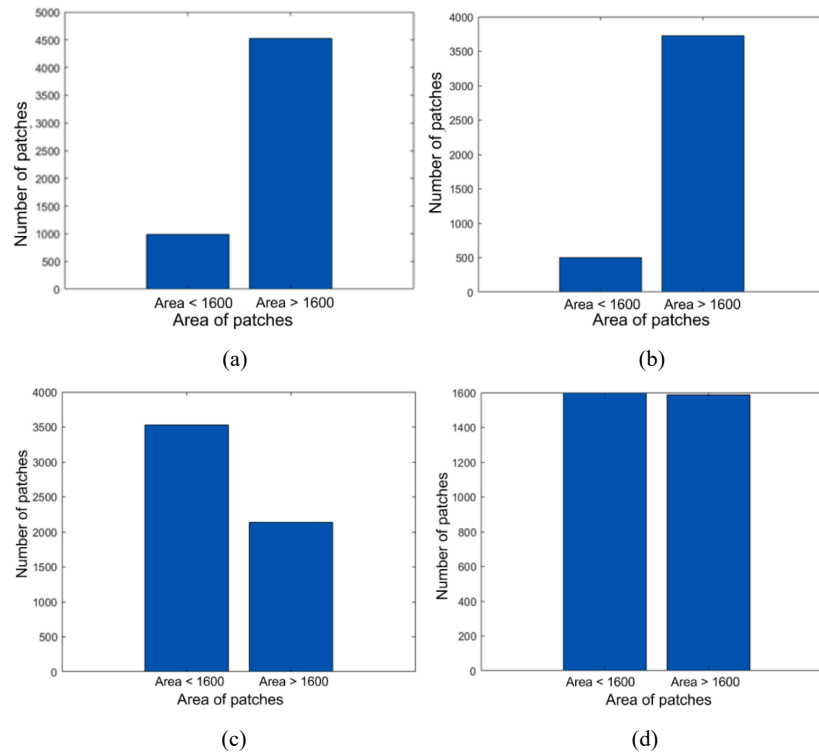


Fig. 9. Graphs of the number of patches with area smaller than 1600 pixels (40×40) and with area larger than 1600 pixels in each image set: (a) set 1; (b) set 2; (c) set 3 and (d) set 4.

Table 3
Patch size per image set.

Patch Size				
Image set				
Statistic	1	2	3	4
Mean	2481.7	2801.9	1487.2	1920.9
Std	1691.4	1478.7	3191.0	2351.0

Table 4
Number of patches for training and validation.

Type		
Size	Positive	Negative
Large	5345	5702
Small	4778	4425

The MLP classifier was selected for segmenting the regions near the foreground objects, since it has the lowest loss value.

Sometimes, when there are two parasites near each other, as shown in Fig. 13, the regions near each parasite are united. Two procedures are applied to solve this problem. The first one, is a watershed operation. If, after the watershed, the region size is greater than 4000 pixels, we divide it in two equal parts. In Fig. 13(a) there are two parasites, one in

the upper side and other in the lower side. After the classifier, only one region is detected, as shown in Fig. 13(b). The separation of one region in two regions is done with the watershed procedure. The result is shown in Fig. 13(c).

4.3. Results for classifying parasite patches

Fig. 14 shows some examples of detected positive and negative parasite patches. As seen, a parasite can have more than one stained object. Fortunately, by the method proposed in this work, all of them are included in only one patch.

Table 5 shows the results for classifying parasite patches. Three results are shown: CNN1 + small patches with size 40×40 pixels, CNN2 + large patches with size 100×100 pixels and CNN2 + all patches with size 100×100 pixels. A 10-fold cross-validation process was used. The following metrics are calculated: Accuracy (Acc), Sensitivity (Sens), Specificity (Spec), Precision (Prec), F1-score and Area Under ROC Curve (AUC), together with the standard deviation of each one. As seen in Table 5, the results of the first two lines, with CNNs dedicated to small patches and large patches, is better than the results when all the patches are converted to 100×100 pixels and classified with CNN2. Therefore, the proposed method uses the two CNNs, CNN1 and CNN2 for patch classification.

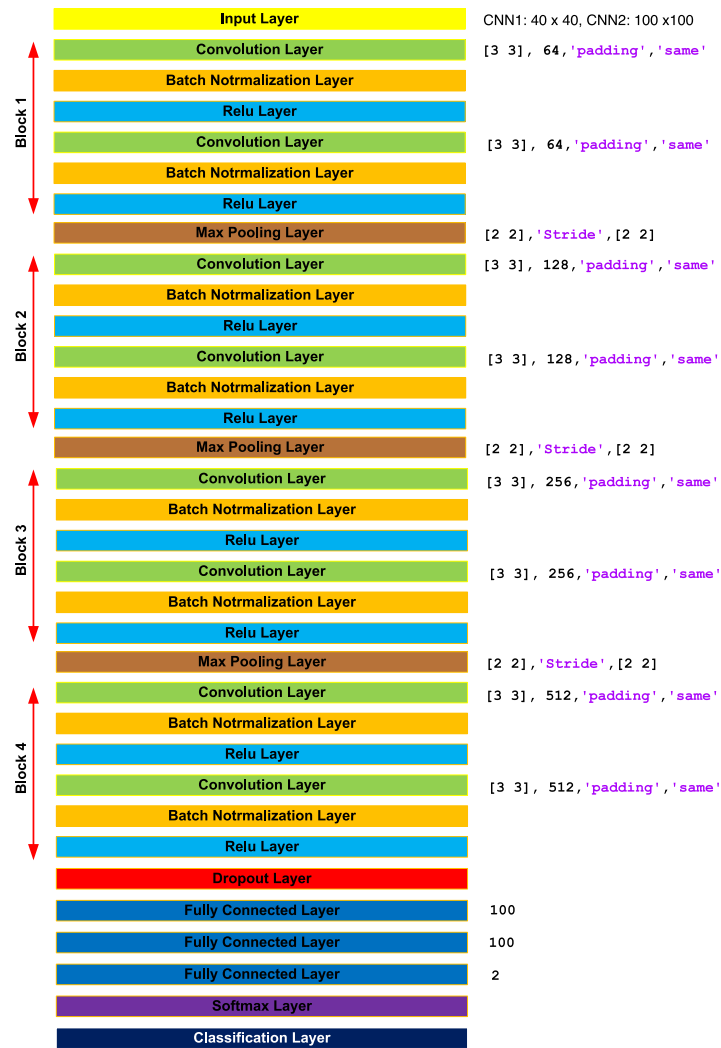


Fig. 10. Architecture of CNN1 and CNN2, showing with 34 layers and 4 blocks of characteristic extraction.

4.4. Results for classifying parasite in full images

Eighty images from the 4 image sets, 20 images per set, were used to evaluate the parasite detection in full images. A 10-fold cross-validation process was used. Each fold consists of eight images, two images from each set (2 images of set 1, 2 images of set 2, 2 images of set 3 and 2 images of set 4). The following metrics were evaluated: accuracy (acc), sensitivity (sens), specificity (spec), precision (prec) and F1-score, together with the standard deviation of each one. Table 5 shows these metric values.

The time required to evaluate an image of 1044×1388 pixels is approximately 6 s. As a malaria diagnosis require evaluation of 160 images, this implies a time of 16 min.

5. Discussion

This work presents an original method for detecting malaria parasites using CNNs for patch classification. We can divide it into four parts: detecting foreground objects, delimiting patches, classifying patches, and detecting parasites in a full image.

Initially, we will comment about detecting foreground objects. In the proposed method, foreground objects are detected using pixel classifiers. As previously mentioned, in other works in the literature, the foreground detection is usually effected by applying thresholds to saturation channel histograms [18–26] and thresholds to gray level values [13,20]. This work evaluates two classifiers: the MLP and a decision tree, using the components of the HSV system as input data. Although the decision tree had an optimal performance with a simple

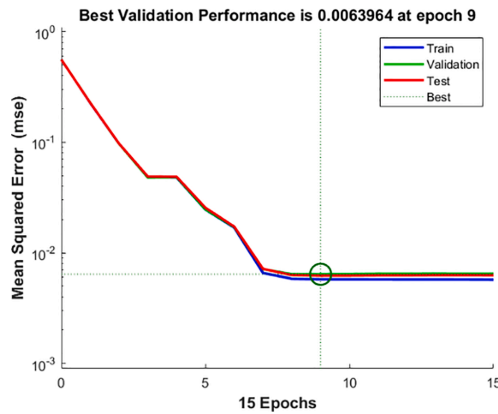


Fig. 11. Mean squared error performance during the training of the MLP for foreground object classification.

threshold applied in the S component, coinciding with the proposed in [18,26], the best performance was obtained with the MLP, using the three components of the HSV space.

About delimiting patches, we also propose pixel classifiers using HSV components to obtain a region near the foreground objects. In other works the authors usually adopt procedures such as morphological closing operations [19] or evaluation of a neighborhood of a detected object [13]. Probably the method proposed in this work contributed to obtain a low false positive rate, which reflects in the high specificity values obtained in Table 7.

Concerning the patch classification, the approach adopted of classifying small patches and large patches with different CNNs showed interesting results. Table 5 shows that CNN1, trained with small patches (area < 1600 pixels), and CNN2, trained with large patches (area > 1600 pixels), obtained AUC of 0.966 and 0.958, respectively. These AUC values are better than the value of 0.941, obtained with CNN2 trained with all patches resized to 100×100 pixels. We believe that resizing small patches to 100×100 pixels generates too much distortion, making CNN recognition more difficult.

Table 7 shows that the sensitivity in sets 3 and 4 is smaller than sensitivity in sets 1 and 2. The following reasons explain this fact: 1) The sensitivity of CNN1, trained with small patches, is lower than sensitivity of CNN2, trained with large patches; 2) sets 1 and 2, as shown in Fig. 9, have more large patches than small patches, while sets 3 and 4 have more small patches than larger patches. Small patches are more difficult

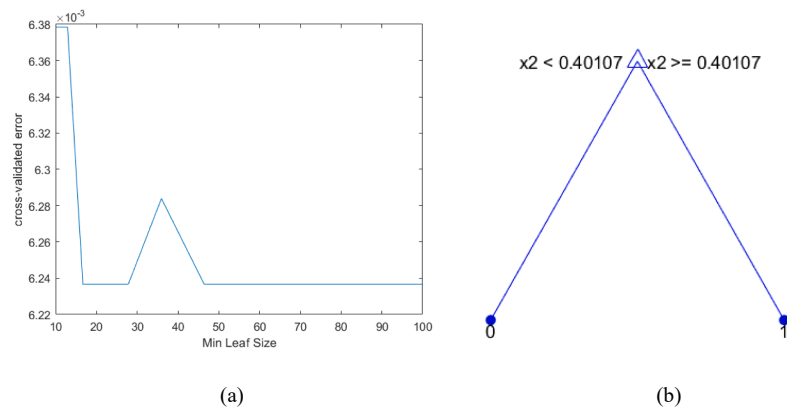


Fig. 12. Decision tree for foreground object classification. (a) 10-fold cross-validation error as a function of minimum leaf size; (b) tree architecture (variable x_2 corresponds to S component).

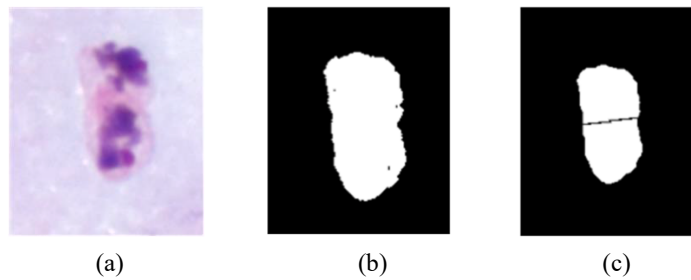


Fig. 13. (a) Image of two parasites near each other; (b) segmented regions near the two parasites united; (c) regions separated using watershed operation.

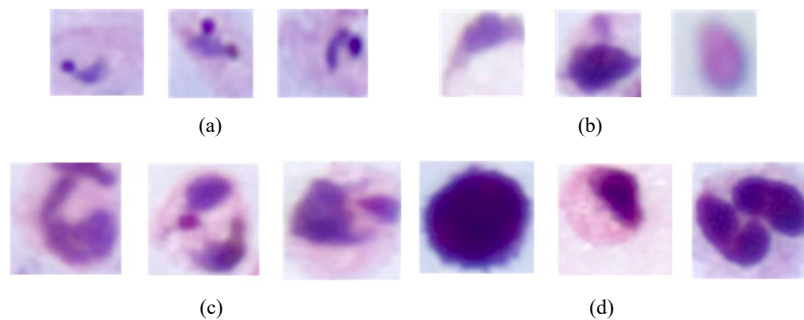


Fig. 14. Images of extracted patches (a) small positive patches; (b) small negative patches; (c) large positive patches; (d) large negative patches.

Table 5
Results for classifying parasite patches.

CNN/ patch	Measure					
	Acc $\pm$ std	Sens $\pm$ std	Spec $\pm$ std	Prec $\pm$ std	F1-score $\pm$ std	AUC $\pm$ std
CNN1/ small patches	90.10 $\pm$ 0.20	85.78 $\pm$ 0.17	96.01 $\pm$ 0.13	85.88 $\pm$ 0.19	87.89 $\pm$ 0.14	0.961 $\pm$ 0.12
CNN2/ large patches	88.91 $\pm$ 0.26	87.75 $\pm$ 0.16	89.90 $\pm$ 0.22	87.82 $\pm$ 0.11	87.90 $\pm$ 0.2	0.952 $\pm$ 0.16
CNN2/all patches	87.80 $\pm$ 0.15	76.79 $\pm$ 0.21	95.91 $\pm$ 0.25	76.84 $\pm$ 0.15	87.91 $\pm$ 0.17	0.939 $\pm$ 0.11
Small Cell	90.36	85.90	96.08	85.90	88.00	0.966

to classify because they can be confounded with noise.

This work presented two types of results: results concerning patch classifications and results concerning parasite classification in full images. The results previously published in the literature present different metric values, probably, because they use different datasets. For example, concerning parasite classification in full images, in [11,14], two works developed using the state-of-the-art generic object detection algorithms, faster regional convolutional neural network (Faster RCNN), precisions of 92.7% and 67.2%, respectively, were obtained. In the first one, a dataset with 1182 images of 750×750 pixels was used, while in the second a database with 903 images of 3264×2448 pixels was used. Another example of parasite detection in full image, adopting the approach used in this paper, patch determination and classification with deep learning, is found in [13]. In this paper, the authors obtained a precision of 78.98%. In [12], also using state-of-the-art generic object detection algorithms, a only look once (YOLO) network, the authors obtained multi average precisions between 64.5% and 83.2% [12]. In this work, precision values varying between 71.55% and 91.71% were obtained. The smaller values are obtained with images containing smaller parasites.

Concerning the sensitivities and specificities values, in [16] the authors, using fast k-means algorithm, obtained results similar to those obtained in this work: high specificity values, and median sensitivity values, in the range between 55.28% and 79.71%. In this work the sensitivities varied between 55.46% and 80.23%. The smaller values are obtained with images containing smaller parasites. In [13], using deep learning techniques, as Faster RCNN and YOLO, the authors conclude that these object detection networks do not work well for very small objects like parasites, with an average size of 44×44 pixels, resulting in many false negatives.

Thus, it can be stated that the method presented in this paper for malaria parasite detection has precision, sensitivity, and specificity

metrics that are comparable to state-of-the-art methods for object detection. A clinical evaluation of the proposed method, through a comparison with results from experienced microscopy technicians, is the next step in this research.

6. Conclusion

This paper presents a new approach for malaria parasite detection in full images using pixel classifiers for obtaining foreground objects and delimiting parasite-stained objects. Both processes were accomplished using HSV components. For patch classification, a CNN network with 34 layers, trained from scratch, was used. The results showed that different metric values are obtained with images containing different parasite sizes, or different stages of development. The better metric values were obtained with images with large parasite sizes, probably because small parasite sizes are easily confounded with noise. The obtained results are comparable with other results previously obtained in the literature. However, a rigorous comparison could not be done, because different datasets are used in each published work.

CRediT authorship contribution statement

Anne de Souza Oliveira: Software, Methodology. **Maria das Graças Vale Barbosa:** Resources, Data curation. **Marly Guimarães Fernandes Costa:** Conceptualization, Methodology, Visualization. **Cicero Ferreira Fernandes Costa Filho:** Writing – original draft, Writing – review & editing, Supervision, Project administration, Methodology, Software.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

Academic English Solutions (<https://www.academicenglishsolutions.com>) revised this paper.

References

- [1] World Health Organization, *World Malaria Report: 20 years of global progress and challenges*, vol. WHO/HTM/GM, no. December, 2020.
- [2] E. C. de S. Gomes et al., "Outbreak of autochthonous cases of malaria in coastal regions of Northeast Brazil: the diversity and spatial distribution of species of Anopheles," *Parasites and Vectors*, 13(1), pp. 1–11, 2020, doi: 10.1186/s13071-020-04502-7.
- [3] World Health Organization, *World health statistics 2021: monitoring health for the SDGs, sustainable development goals*. 2021.

- [4] WHO, "Guideline WHO Guidelines for malaria," 16 Febr. 2021, p. 225, 2021, [Online]. Available: <http://apps.who.int/bookorders>.
- [5] T. Mutabazi, et al., Assessment of the accuracy of malaria microscopy in private health facilities in Entebbe Municipality, Uganda: a cross-sectional study, *Malar. J.* 20 (1) (2021) 1–9, <https://doi.org/10.1186/s12936-021-03787-y>.
- [6] Ethiopian Public Health Institute - Federal Ministry of Health, *Manual for the Laboratory Diagnoses of Malaria*, 2nd ed. 2020.
- [7] B. Health Surveillance Secretariat - Ministry of Health, *Manual for the Laboratory Diagnoses of Malaria*. 2009.
- [8] R. Nakasi, E. Mwebaze, A. Zawedde, J. Tusubira, B. Akera, G. Maiga, A new approach for microscopic diagnosis of malaria parasites in thick blood smears using pre-trained deep learning models, *SN Appl. Sci.* 2 (7) (2020) 1–7, <https://doi.org/10.1007/s42452-020-3000-0>.
- [9] G. Madhu et al., "Imperative Dynamic Routing between Capsules Network for Malaria Classification," *Comput. Mater. Contin.*, 68(1), pp. 903–919, 2021, doi: 10.32604/cmc.2021.016114.
- [10] P.A. Pattanaik, M. Mittal, M.Z. Khan, Unsupervised Deep Learning CAD Scheme for the Detection of Malaria in Blood Smear Microscopic Images, *IEEE Access* 8 (2020) 94936–94946, <https://doi.org/10.1109/ACCESS.2020.2996022>.
- [11] F. Abdurahman, K. A. Fante, and M. Aliy, "Malaria parasite detection in thick blood smear microscopic images using modified YOLOV3 and YOLOV4 models," *BMC Bioinformatics*, 22(1), 2021, doi:10.1186/s12859-021-04036-4.
- [12] S. Chibuta and A. C. Acar, "Real-time Malaria Parasite Screening in Thick Blood Smears for Low-Resource Setting," *J. Digit. Imaging*, 33(763–776), 2020, doi: 10.1007/s10278-019-00284-2.
- [13] F. Yang et al., "Deep Learning for Smartphone-Based Malaria Parasite Detection in Thick Blood Smears," *IEEE J. Biomed. Heal. Informatics*, 24(5), pp. 1427–1438, 2020, doi:10.1109/JBHI.2019.2939121.
- [14] R. Nakasi, E. Mwebaze, and A. Zawedde, "Mobile-aware deep learning algorithms for malaria parasites and white blood cells localization in thick blood smears," *Algorithms*, 14(1), 2021, doi:10.3390/a14010017.
- [15] J. A. Quinn, R. Nakasi, P. K. B. Mugagga, P. Byanyima, W. Lubega, and A. Andama, "Deep Convolutional Neural Networks for Microscopy-Based Point of Care Diagnostics," in: *Machine Learning for Healthcare Conference*, 2016, pp. 271–281, [Online]. Available: <http://arxiv.org/abs/1608.02989>.
- [16] T. A. Aris, A. S. A. Nasir, L. C. Chin, H. Jaafar, and Z. Mohamed, "Fast k-means clustering algorithm for malaria detection in thick blood smear," in: *2020 IEEE 10th International Conference on System Engineering and Technology, ICSET 2020 - Proceedings*, 2020, pp. 267–272, doi:10.1109/ICSET51301.2020.9265380.
- [17] S. R. Abidin, U. Salamah, and A. S. Nugroho, "Segmentation of malaria parasite candidates from thick blood smear microphotographs image using active contour without edge," 2016, doi:10.1109/IBIOMED.2016.7869824.
- [18] I. R. Dave and K. P. Upla, "Computer aided diagnosis of Malaria disease for thin and thick blood smear microscopic images," in: *2017 4th International Conference on Signal Processing and Integrated Networks, SPIN 2017*, 2017, pp. 561–565, doi: 10.1109/SPIN.2017.8050013.
- [19] F. M. Azif, H. A. Nugroho, and S. Wibirama, "Adaptive Threshold Determination Based on Entropy in Active Contour without Edge Method for Malaria Parasite Candidate Detection," in: *Proc. - 2018 4th Int. Conf. Sci. Technol. ICST 2018*, pp. 0–5, 2018, doi: 10.1109/ICSTC.2018.8528585.
- [20] J.A. Quinn, A. Andama, I. Munabi, F.N. Kiwanuka, Automated blood smear analysis for mobile malaria diagnosis, *Mob. Point-of-Care Monit. Diagnostic Device Des.* (September) (2014) 115–132, <https://doi.org/10.1201/9781315215723-10>.
- [21] Z. Zhang et al., "Image classification of unlabeled malaria parasites in red blood cells," in: *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, 2016, vol. 2016-Octob, pp. 3981–3984, doi:10.1109/EMBC.2016.7591599.
- [22] C. B. Delahunt et al., "Automated microscopy and machine learning for expert-level malaria field diagnosis," in: *Proc. 5th IEEE Glob. Humanit. Technol. Conf. GHTC 2015*, pp. 393–399, 2015, doi: 10.1109/GHTC.2015.7344002.
- [23] A. Huq and M. T. Pervin, "Robust Deep Neural Network Model for Identification of Malaria Parasites in Cell Images," in: *2020 IEEE Region 10 Symposium, TENSYP 2020*, 2020, pp. 1456–1459, doi:10.1109/TENSYP50017.2020.9230832.
- [24] F. Yang, H. Yu, K. Silamut, R. J. Maude, S. Jaeger, and S. Antani, "Parasite Detection in Thick Blood Smears Based on Customized Faster-RCNN on Smartphones," in: *Proceedings - Applied Imagery Pattern Recognition Workshop*, 2019, vol. 2019-Octob, doi:10.1109/AIPR47015.2019.9174565.
- [25] R. Nakasi, J. F. Tusubira, A. Zawedde, A. Mansourian, and E. Mwebaze, "A web-based intelligence platform for diagnosis of malaria in thick blood smear images: A case for a developing country," in: *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, 2020, vol. 2020-June, pp. 4238–4244, doi:10.1109/CVPRW50498.2020.00500.
- [26] I. R. Dave, "Image analysis for malaria parasite detection from microscopic images of thick blood smear," in: *Proceedings of the 2017 International Conference on Wireless Communications, Signal Processing and Networking, WISPNET 2017*, 2017, vol. 2018-Janua, pp. 1303–1307, doi:10.1109/WISPNET.2017.8299974.
- [27] J. Kaur, A. Choudhary, *Comparison of Several Contrast Stretching Techniques on Acute Leukemia Images*, *Int. J. Eng. Innov. Technol.* 2 (1) (2012) 332–335.
- [28] N. S. M. M. Hanif, M. Y. Mashor, and Z. Mohamed, "Image enhancement and segmentation using dark stretching technique for Plasmodium Falciparum for thick blood smear," in: *Proc. - 2011 IEEE 7th Int. Colloq. Signal Process. Its Appl. CSPA 2011*, pp. 257–260, 2011, doi: 10.1109/CSPA.2011.5759883.
- [29] C. C. Aggarwal, *Neural Networks and Deep Learning*. 2021.

APÊNDICE B – ARTIGO PUBLICADO NA REVISTA *DIAGNOSTICS* EM 2024

Neste apêndice, é apresentado o artigo *Comparing different data partitioning strategies for segmenting areas affected by COVID-19 in CT scans*, publicado na revista *Diagnostics* em 2024. Este é o segundo artigo resultante deste trabalho.

Article

Comparing Different Data Partitioning Strategies for Segmenting Areas Affected by COVID-19 in CT Scans

Anne de Souza Oliveira ¹ , Marly Guimarães Fernandes Costa ¹ , João Pedro Guimarães Fernandes Costa ² and Cícero Ferreira Fernandes Costa Filho ^{1,*} 

¹ R&D Center in Electronic and Information Technology, Federal University of Amazonas, Manaus 69077-000, Brazil; anneollivveira@gmail.com (A.d.S.O.); mcosta@ufam.edu.br (M.G.F.C.)
² Cancer Institute of São Paulo State, São Paulo 01246-000, Brazil; joaopedrogfc@gmail.com
 * Correspondence: ccosta@ufam.edu.br; Tel.: +55-(92)-991464954

Abstract: Background/Objectives: According to the World Health Organization, the gold standard for diagnosing COVID-19 is the Reverse Transcription Polymerase Chain Reaction (RT-PCR) test. However, to confirm the diagnosis in patients who have negative results but still show symptoms, imaging tests, especially computed tomography (CT), are used. In this study, using convolutional neural networks, we compared the following topics using manual and automatic lung segmentation methods: (1) the performance of an automatic segmentation of COVID-19 areas using two strategies for data partitioning, CT scans, and slice strategies; (2) the performance of an automatic segmentation method of COVID-19 when there was interobserver agreement between two groups of radiologists; and (3) the performance of the area affected by COVID-19. Methods: Two datasets and two deep neural network architectures are used to evaluate the automatic segmentation of lungs and COVID-19 areas. The performance of the U-Net architecture is compared with the performance of a new architecture proposed by the research group. Results: With automatic lung segmentation, the Dice metrics for the segmentation of the COVID-19 area were $73.01 \pm 9.47\%$ and $84.66 \pm 5.41\%$ for the CT-scan strategy and slice strategy, respectively. With manual lung segmentation, the Dice metrics for the automatic segmentation of COVID-19 were $74.47 \pm 9.94\%$ and $85.35 \pm 5.41\%$ for the CT-scan and the slice strategy, respectively. Conclusions: The main conclusions were as follows: COVID-19 segmentation was slightly better for the slice strategy than for the CT-scan strategy; a comparison of the performance of the automatic COVID-19 segmentation and the interobserver agreement, in a group of 7 CT scans, revealed that there was no statistically significant difference between any metric.

Keywords: computed tomography; convolutional neural networks; coronavirus; COVID-19; deep learning; image segmentation; lung; radiology



Citation: Oliveira, A.d.S.; Costa, M.G.F.; Costa, J.P.G.F.; Costa Filho, C.F.F. Comparing Different Data Partitioning Strategies for Segmenting Areas Affected by COVID-19 in CT Scans. *Diagnostics* **2024**, *14*, 2791. <https://doi.org/10.3390/diagnostics14242791>

Academic Editor: Thomas Frauenfelder

Received: 7 November 2024
 Revised: 9 December 2024
 Accepted: 10 December 2024
 Published: 12 December 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In the previous decades, none of the coronaviruses had caused serious illness in humans, but on 31 December 2019, the World Health Organization was alerted to cases of pneumonia in the city of Wuhan, China. These were confirmed on 7 January 2020 to belonging to a new coronavirus variant, SARS-CoV-2. The infectious disease caused by the SARS-CoV-2 coronavirus has been named COVID-19 [1].

On 11 March 2020, COVID-19 was characterized as a pandemic due to the geographical distribution of the disease in several countries around the world. The severity of COVID-19 was defined by the World Health Organization as critical, severe, and not severe [2] and the symptoms listed were a fever, tiredness, and a dry cough, as well as less common symptoms such as a loss of taste or smell, nasal congestion, conjunctivitis, sore throat, headache, muscle or joint pain, skin rashes, nausea or vomiting, diarrhea, chills, and dizziness.

The diagnostic confirmation of COVID-19 is realized with the identification of viral RNA via a Reverse Transcription Polymerase Chain Reaction (RT-PCR) test, which is considered the gold-standard test. However, confirmation of the diagnosis can be performed

via chest imaging in situations where there are symptoms of the disease, but RT-PCR is not available; there is a delay in obtaining the results; or the RT-PCR results are initially negative, but there are symptoms that suggest the disease.

Chest imaging tests include chest radiography, chest computed tomography (CT), and lung ultrasound analysis. In a systematic review carried out in [3], for the first months of the COVID-19 pandemic, 23 studies were found that used these three types of imaging to assess the accuracy of the diagnosis in relation to a reference standard. Overall, 3 of these used chest X-ray exams, 19 used chest CT scans, and 1 used a lung ultrasound. Among the 19 CT studies, 5 used artificial intelligence (AI) to interpret the images. Among the studies that used AI, two used deep learning models, with one applying convolutional neural networks (CNNs). The tasks performed in these studies largely involved classification, where the model determines whether the CT classified the patient as with or without COVID-19.

In the literature review carried out in this work, studies were also found that used deep learning models trained with convolutional neural networks to delimit the regions affected by COVID-19 in CT scans. Most of the articles used to segment areas affected by COVID-19 only use the features extracted by the deep network itself. It should be noted, however, that some approaches use the fusion of these features with handcrafted features extracted using the Weber Local Descriptor (WLD). The authors in [4] claim that, regarding fusion characteristics, the results are better than those obtained with only ResNet101 or DensNet201.

To undertake this review, the databases selected were PubMed, IEEE, ScienceDirect, and ACM and the general expression used in the query was as follows: ((machine learning) AND (deep learning)) AND ((image segmentation) AND (computerized tomography)) AND ((lung infection) AND (coronavirus OR COVID-19)). We made appropriate adaptations for each database. The total number of articles returned without repetition was 56.

In the screening phase, 31 articles were retained and 25 were excluded according to the following established criteria: 1. The article was included if it trained convolutional neural networks to segment COVID-19. 2. The article was included if it cited the databases used. 3. An article was included if it cited the data partitioning used for training and testing (CT-scan or slice strategy). 4. The article was included if it cited the networks used to perform the segmentation task(s). 5. The article was included if it cited the Dice metric.

In the analysis phase, 22 articles were retained and 9 were excluded. The criteria established were as follows: 1. the article was included if it cited the network architecture used to segment the lungs, COVID-19, or both; 2. the article was included if it cited the type of method adopted, whether automatic or iterative; and 3. the article was included if it aimed to train a model to segment COVID-19.

After analyzing the articles, we observed that segmenting the affected areas is of fundamental importance for the clinical follow-up of patients, as it makes it possible to monitor the proportion of COVID-19 in the lung region on CT scans, and that studies dealing with COVID-19 segmentation can be classified according to the number of steps used in the segmentation process and according to the methodology used for data partitioning.

With respect to the number of steps used in the segmentation process, we identified two approaches in the literature. First, as shown in Table 1, COVID-19 segmentation is carried out in a single step [5–19]. This approach is referred to in this article as one-step segmentation. In the second approach, as shown in Table 2, the segmentation of the area affected by COVID-19 is preceded by a lung segmentation step [20–26]. This approach is referred to in this article as two-step segmentation.

With respect to the methodology used for data partitioning (DP), we also identified two strategies. In the first strategy, the data are divided by CT scans, and the slices of a patient's CT scan are only used for training or testing the classifier, as shown in [5–11,20–22]. In the second strategy, the data are divided into slices. In this strategy, slices from the same patient's CT scan can be used for both training and testing the classifier, as shown in [12–19,23–26]. This first strategy is referred to in this article as the CT-scan strategy, and the second strategy is referred to as the slice strategy.

Table 1. Summary of published papers on COVID-19 segmentation in CT images via one-step segmentation methods.

Strategy	Ref.	Classifier	Dataset	Train	Validation	Test	Dice (%)
CT scan	[5]	Attention-RefNet	Proprietary	149 CT scans	29 CT scans	32 CT scans	92.75
			COVID-19 CT Seg	3-fold CV	1-fold CV	1-fold CV	91.37
			MICCAI	3-fold CV	1-fold CV	1-fold CV	81.22
	[6]	VB-Net	Proprietary	249 CT scans	-	300 CT scans	91.60 ± 10.00
	[7]	COPLE-Net	Proprietary	378 CT scans	50 CT scans	130 CT scans	80.29 ± 11.14
	[8]	U-Net	Harbin	4-fold CV		1-fold CV	78.30
	[9]	U-Net with backbone ResNet 34	Proprietary	531 CT scans	-	30 CT scans	Radiologist with 20 years' experience 74.00 ± 0.28.
							Radiologist with 15 years' experience 76.00 ± 0.29.
	[10]	nCoVSegNet	LIDC-IDRI	875 CT scans	$\frac{1}{4}$	-	-
			MosMedData Coronacases	40 CT scans -	- -	10 CT scans 10 CT scans	68.43 68.94
Slice	[11]	U-Net	Proprietary	50 CT scans		130 CT scans	67.00
	[12]	CNN	COVID-CT	70%	10%	20%	92.00
	[13]	Based on U-Net	COVID-CT COVID-19 CT Henri Becquerel Cancer Center	1069 slices	150 slices	150 slices	88.00
	[14]	O-Net	COVID-19 CT Seg MosMedData	70%	-	30%	86.00
	[15]	QC-HC U-Net	MSD	80%	-	20%	85.31
			COVID-19 CT Seg	80%	-	20%	
	[16]	STCNet	CC-CCII segmentation	70%	-	30%	82.78
			COVID-19 CT Seg	70%	-	30%	79.92
	[17]	JSC	COVID-CS	2794 slices	-	1061 slices	78.50
	[18]	Semi-Inf-Net	COVID-19 CT	45 slices	5 slices	50 slices	73.90
			COVID-19 CT collection	1600 slices	-	-	
	[19]	SSInf-Net	COVID-19 ICTCF	698 slices 6654 slices	114 slices -	117 slices -	63.00

Table 2. Summary of published papers on COVID-19 segmentation in CT images via two-step segmentation methods.

Strategy	Ref.	Classifier	Dataset	Train	Valid.	Test	Dice (%)
CT scan	[20]	Lungs Model by another author COVID-19 3D dense CNN	Proprietary	64 CT scans	23 CT scans	77 CT scans	COVID-19 Institute 1 71.30
				58 CT scans	14 CT scans	87 CT scans	Institute 2 65.70 Institute 3 64.10
	[21]	Lungs nnU-Net COVID-19 nnU-Net	COVID-19 CT Seg	4-fold CV with 16 CT scans	-	1-fold CV with 4 CT scans	Right lung 87.90 ± 9.30 Left lung 85.80 ± 10.50 COVID-19 67.30 ± 22.30
			MosMed	-	-	50 CT scans	COVID-19 58.80 ± 20.60
	[22]	Lungs Model by another author COVID-19 U-Net	Proprietary	COVID-19 Infection 80 CT scans	COVID-19 Infection 7 CT scans	COVID-19 Infection 62 CT scans	COVID-19 Infection 63.20
			Proprietary	COVID-19 Consolidation 19 CT scans	-	-	COVID-19 Consolidation 62.80
			MedSeg	-	-	COVID-19 Consolidation 9 CT scans	
	[23]	Lungs Model by another author COVID-19 U-Net	COVID-19 CT	4-fold CV with 80 slices	-	1-fold CV with 20 slices	COVID-19 85.00
	[24]	Lungs Model by another author COVID-19 Anam-Net	COVID-19 CT Seg (exp.2)	3-fold CV	-	1-fold CV with 545 slices	COVID-19 79.80
			COVID-19 CT	270 slices	-	-	COVID-19 75.50
			MedSeg	-	-	704 slices	
			-	Model trained in exp. 2	-	-	COVID-19 66.40
			MedSeg	-	-	704 slices	
Slice	[25]	Lungs RCNN-ResNetSt-200 COVID-19 RCNN-ResNetSt-200	NSCL COVID-19 CT Seg	50,756 slices	-	1222 slices	Lungs 97.18
			MosMed MSD COVID-19 CT Seg	5854 slices	-	1117 slices	COVID-19 78.06
	[26]	Lungs Model by another author COVID-19 SegNet	COVID-19 CT	72 slices	10 slices	18 slices	COVID-19 74.90

Table 1 (one-step segmentation methods) shows the papers with the highest values in terms of the Dice metric. Importantly, among these studies, refs. [5,6] adopted the CT-scan strategy for data partitioning, which presented Dice metric values of 92.75% and 91.60%, respectively, and used interactive methods to obtain COVID-19 segmentation models. The other studies used automatic methods to segment COVID-19, with [7] achieving a Dice metric of 80.29%. Among the studies that adopted the slice strategy for data partitioning, all used automatic methods, with the highest value for the Dice metric achieved in the work of [11], at 92%, followed by the work of [13], at 88%. The studies in [5,14–16] employed one of the datasets used in this work: the COVID-19 CT Lung and Infection Segmentation Dataset [27]. This dataset is referred to in this work as COVID-19 CT Seg.

All the papers in Table 2 (two-step segmentation methods) implemented automatic segmentation methods. Among the studies that adopted the CT-scan strategy for data partitioning, we highlight those of [20–22], which obtained Dice metric values of 71.30%, 67.30%, and 63.20%, respectively. Among the studies that adopted a slice strategy for data partitioning, we highlight those of [23–26], with Dice metrics of 85.00%, 79.80%, 78.06%, and 74.90%, respectively.

Some studies [21,22,24,25] employed the datasets used in this work: COVID-19 CT Seg [27] and/or MedSeg COVID Dataset 2 [28]. The latter is referred to in this work as MedSeg. Ref. [21] used nnU-Net to segment the left and right lungs and detect COVID-19 infection. In [22], the U-Net architecture was used to segment regions infected with COVID-19. Ref. [24] proposed the Anam-Net architecture and provided the percentages of normal and COVID-19-infected regions. In [25], the RCNN-ResNeSt-200 mask-cascade architecture was used in the first module of the three proposed by the authors to train and evaluate lung and COVID-19 segmentation models.

Among the studies listed in Tables 1 and 2, we observed that no study compared the automatic segmentation of COVID-19 using two DP strategies, i.e., CT-scan and slice methods; that one study presented a comparison of the performance of the automatic segmentation of COVID-19 with the performance of the segmentation conducted by two groups of radiologists on the same set of images and the interobserver agreement between these two groups [8]; and that two studies used interactive methods to segment COVID-19, i.e., they relied on human intervention in the segmentation process [5,6].

In view of what has been described above, the contributions of this work are as follows:

- Using CT images, we compare the performance of deep learning tools by employing the two-step segmentation method for COVID-19; we use both the data partitioning strategies, i.e., CT-scan and slice methods.
- Using CT images, we compare the performance of deep learning tools by employing a two-step segmentation method for COVID-19, with the agreement level of segmentation assessed by two distinct groups of radiologists for the same dataset.
- Using CT images, we compare the performance of deep learning tools in a two-step segmentation process for COVID-19 employing manual and automatic lung segmentation.

2. Materials and Methods

2.1. Materials

Dataset 1, known as COVID-19 CT Seg, consists of 20 CT scans [27]. This dataset contains 3520 slices that were assessed in 2 stages, an initial assessment and a review, by 2 radiologists. A total of 1644 slices were positive. Figure 1a shows the number of slices per CT scan in this dataset. CT scans, lung masks, and COVID-19 masks were available. The number of slices with COVID-19 in each CT scan ranged from 2 to 216. In the CT scans, the patient's chests were turned to the right and left. The original sizes of the images were 630×630 pixels and 512×512 pixels. All image regions associated with COVID-19 were labeled with ground glass opacities.

Dataset 2, known as MedSeg, consists of 9 CT scans [28]. This dataset has 829 slices, of which 373 slices are positive, that were evaluated by a radiologist. Figure 1b shows the number of slices per CT scan in this dataset. It contains 7 CT scans in common with Dataset 1. As the COVID-19 areas of the exams in Dataset 2 were evaluated by a different group of radiologists, it was possible to compare the level of agreement of the method proposed in this study with the results of the two groups of radiologists when segmenting the COVID-19 area in the 7 CT scans in common.

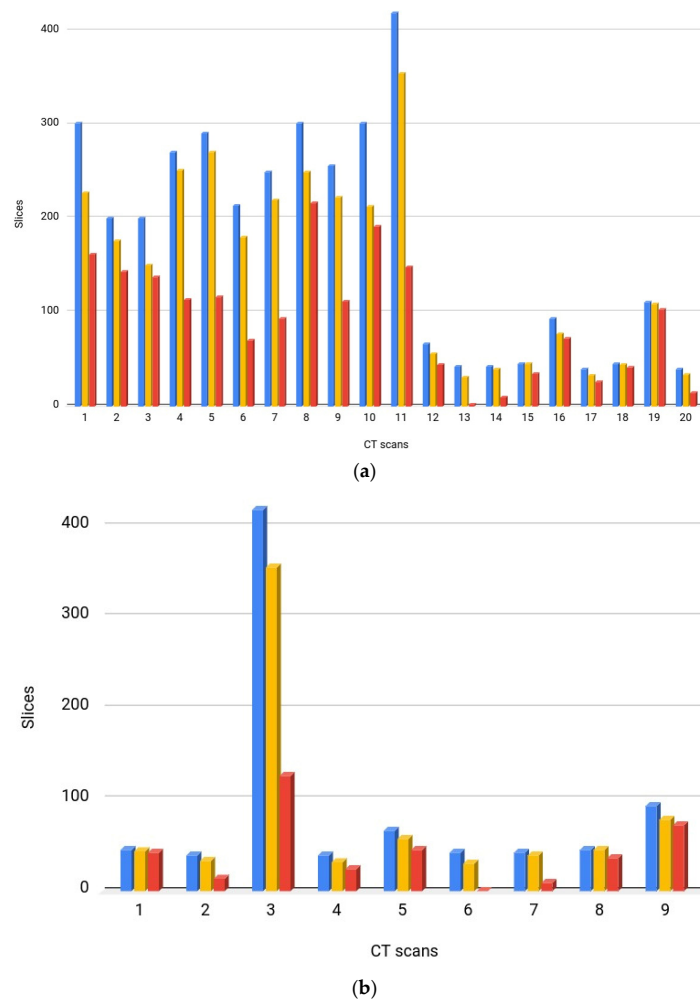


Figure 1. Graphs showing, for each CT scan of the dataset, the total number of slices, the number of slices containing the lung, and the number of slices containing regions with COVID-19. (a) Dataset 1; (b) Dataset 2. Blue bars: number of slices; yellow bars: number of slices with lung; red bars: number of slices with COVID-19.

2.2. Methods

This work uses two-step segmentation. Figure 2 illustrates the methodology adopted to evaluate COVID-19 area segmentation in CT scans. Two ways of dividing the dataset are evaluated, namely, the slice and the CT-scan strategies.

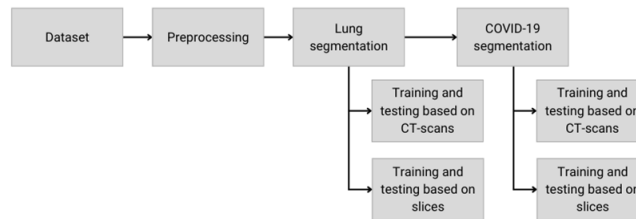


Figure 2. The steps of the methodology adopted in this work.

2.2.1. Preprocessing

The preprocessing steps are shown in Figure 3. In the first step, a radiologist of our research group evaluated all the CT scans of both datasets and reported that two CT scans from Dataset 1 were identical, with the only difference being that one of them contained more slices than the other. Therefore, one of the exams was removed and Dataset 1 was left with 19 CT scans.



Figure 3. Steps used in image preprocessing.

The pixels of each slice, in both datasets, were normalized to the range 0–255 according to the MAX-MIN normalization process shown in Equation (1) [29]:

$$f_n = \frac{f_0 - f_{\min}}{f_{\max} - f_{\min}} \quad (1)$$

where

f_n —normalized pixel value;

f_0 —original pixel value;

$f_{\min}$ —minimum pixel value in the slice;

$f_{\max}$ —maximum pixel value in the slice.

All the images were resized to 512×512 pixels. In this step, one CT scan of Dataset 1, which presented a different ratio between height and width, was removed. Therefore, Dataset 1 was left with only 18 CT scans. To ensure all the CT scans to have their chests facing down and text annotations in the correct position, rotation and inversion operations were applied in the last preprocessing step. Figure 4 shows examples of CT-scan slices after all modifications were performed.

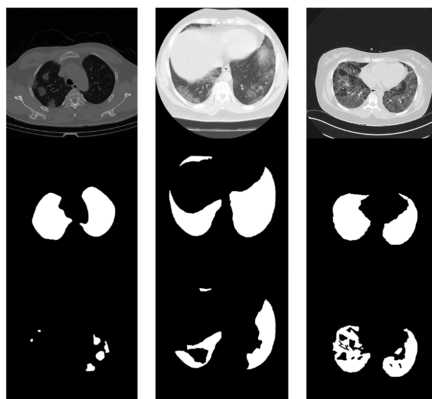


Figure 4. Examples of slices of CT scans after preprocessing. From top to bottom, a slice, a lung mask, and a COVID-19 mask.

2.2.2. DP Strategies for Training, Validation and Testing

For classifier evaluation, two DP strategies were used to divide the datasets into training and testing sets. The first one was the CT-scan strategy, which uses different CT scans for training and testing a classifier. This strategy is also known as the independent subject strategy because the CT scans of a patient are only used for training or for testing, not for both. A random process was used to select the CT scans from both sets. Figure 5 shows the results of the CT-scan strategy.

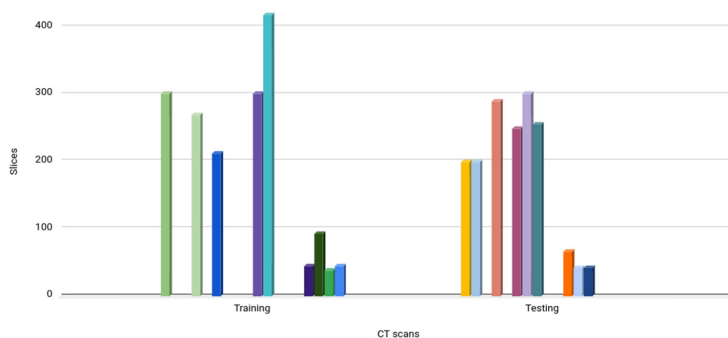


Figure 5. Training and test sets obtained with the CT-scan strategy. Different colors represent different CT-scans.

The second method was the slice strategy, which uses different slices of the same CT scan for both training and testing a classifier. This strategy is also known as the dependent subject strategy because slices from a CT scan of a patient are used for training and testing. An interleaved process was used to select the slices in each CT scan. From a given CT scan, 50% of the slices were used for training and 50% were used for testing. Figure 6 shows the results of the slice strategy.

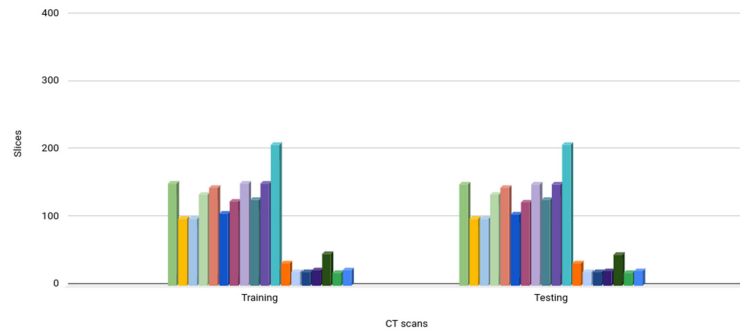


Figure 6. Training and test sets obtained with the slice strategy. Different colors represent different CT-scans.

All training sets were split at runtime to create a validation set. This procedure was carried out during the model training phase and a 75:15 ratio was adopted.

A survey of the images showed that the area affected by COVID-19 in the CT scans varied from 0.01% to 29.57%, which guaranteed a good diversity of images.

2.2.3. Classifiers

Two convolutional neural networks were used in this work for semantic segmentation. The first architecture, shown in Figure 7, was U-Net [30]. The authors evaluated this architecture for use in the segmentation of neuronal structures in electron microscopic recordings and cell segmentation in light microscopic images using 2 different datasets and achieved good performances for all the evaluated tasks. U-Net consists of an encoder with 4 subsampling stages. Each subsampling stage is composed of 2 convolution layers with a 3×3 kernel followed by a 2×2 maxpooling layer. The decoder consists of 4 oversampling stages. Each oversampling stage is composed of a 2×2 upconvolution layer, a concatenation with the corresponding feature map from the encoder, and 2 convolution layers with 3×3 kernels. This architecture connects encoder layers with decoder layers. U-Net is widely used in semantic segmentation tasks, particularly in medical image processing. It is used to perform accurate segmentation when there are few training images. In some of the experiments in our work, to improve generalization, a dropout layer was inserted between the Conv2D layer with 1024 filters and the Conv2DTranspose layer with 512 filters.

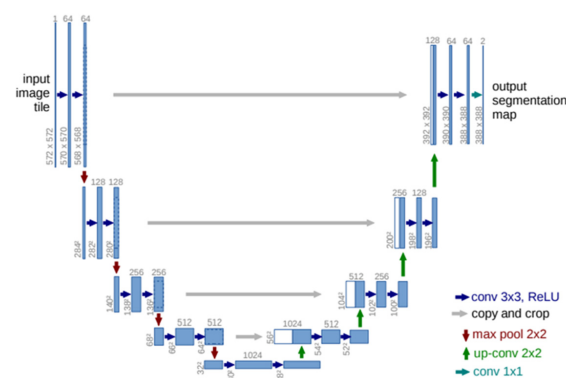


Figure 7. Convolutional neural network architecture used for semantic segmentation of lungs and COVID-19—U-Net [29].

The second architecture, CNN2, is shown in Figure 8 and was proposed by our research group [31] in order to segment the lumen in intravascular optical coherence tomography images. Like U-Net, this architecture can also be divided into two parts: downsampling and upsampling sections. The downsampling section is composed of blocks of the convolution layer–batch layer–ReLU layer. The convolution layers have filters of size 3×3 and padding 1. These blocks are followed by a convolution layer with a filter size of 1×1 and a MaxPool layer with a filter size of 2 and stride of 2. At the downsampling end, there is a dropout layer. At the architecture output, there is a Softmax layer and a classification layer. At the upsampling section, between the transposed convolution layers, there is a block composed of a convolution layer with a filter size of 3×3 –batch layer–ReLU layer–convolution layer with a filter size of 1×1 . At the output, there is a Softmax layer and a classification layer.

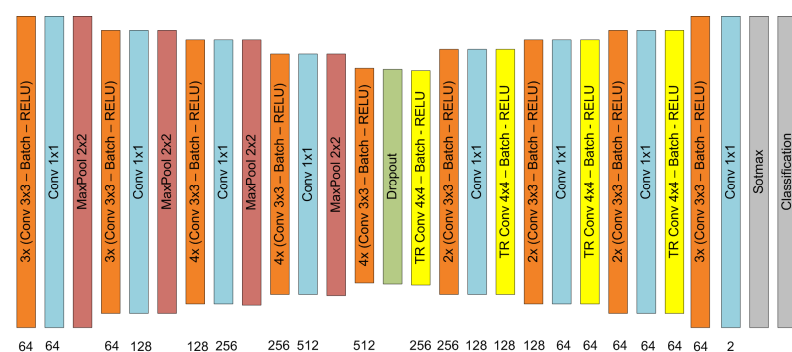


Figure 8. Convolutional neural network architecture used for semantic segmentation of COVID-19—CNN2.

The optimal hyperparameters chosen, after some experiments, to train both architectures are shown in Table 3. During the training phase, we randomly evaluated some options for the following hyperparameters: optimizer (RMSProp, ADAM and gradient descent with momentum); batch size (8, 16, 32); initial learning rate (10^{-3} , 10^{-4}); dropout rate (0.4, 0.5 and 0.6); L_2 regularization (10^{-2} , 10^{-3} , 10^{-4}).

Table 3. Hyperparameters used for classifier training.

Hyperparameter	Value
Optimizer	ADAM
Batch size	32
Training stop method	Early stop
Initial learning rate	10^{-3}
Reduction learning rate factor	10^{-1}
Interval between learning rate reductions	20 epochs
Minimum learning rate	10^{-45}
Maximum number of training epochs	600
Loss function	Binary cross entropy
Dropout rate	0.5
L_2 regularization factor	10^{-3}
Data augmentation	10 rotations of 36°

Deep learning models, with their complex architectures and numerous parameters, are susceptible to overfitting, especially when trained on small or imbalanced datasets. This occurs when the model learns the training data too well, leading to poor generalization of unseen data. Techniques like regularization, dropout, data augmentation, and early stop can mitigate this issue, but it remains a significant challenge. In this work, we used these

four techniques to mitigate the overfitting. The regularization method used was the L_2 technique. The data augmentation consisted of applying 10 rotations of 36° to the original images of the training sets. The early stop had a patience of 3.

2.2.4. Lung and COVID-19 Segmentation

For lung segmentation, only the U-Net architecture was employed. For each DP strategy, 8 experiments were performed, combining dropout (D), L_2 regularization (L), and data augmentation (A). For COVID-19 segmentation, both U-Net and CNN2 architectures were employed. For each strategy, 8 experiments were also performed using the same combination mentioned. COVID-19 segmentation was performed via both automatic and manual lung segmentation. For manual lung segmentation, the radiologists' lung labels were used.

2.3. Simulation Environment and Evaluation Metrics

The environment used to write the codes for training and testing the models was a cloud service hosted by Google, namely, Google Colab. The software used to implement the codes was Keras, a high-level deep learning API written in Python. The machine learning kernel used with Keras was TensorFlow 2.12.0, an open-source platform.

The version of Python installed on the cloud computer was 3.9.16. All the generated codes and models were stored on Google Drive. To expand Google Drive's storage capacity, Google One was added to the environment. This work adopted a public cloud architecture with private resources. The configuration of the public cloud architecture was as follows: storage capacity, 100 GB; RAM memory, 35 GB; processor, Intel(R) Xeon(R) CPU @ 2.30 GHz; operational system, Linux with Ubuntu 20.04.3 LTS; tensor processor unit, TPU v2; and RAM memory of the tensor unit, 16 GB. To gain access to continuous code execution, increase the amount of RAM, and access the TPU, Google Colab Pro was added. The TPU consists of four TPU chips and 16 GB of high-bandwidth memory (HBM).

To evaluate the model performances, the following metrics were employed: accuracy—Equation (2); global accuracy—Equation (3); Dice—Equation (4); and Jaccard—Equation (5).

$$\text{Accuracy} = \frac{\frac{TP}{TP+FN} + \frac{TN}{TN+FP}}{2} \quad (2)$$

$$\text{Global accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (3)$$

$$\text{Dice} = \frac{2TP}{2TP + FP + FN} \quad (4)$$

$$\text{Jaccard} = \frac{TP}{TP + FP + FN} \quad (5)$$

where:

TP—true positives;

TN—true negatives;

FP—false positives;

FN—false negatives.

To assess whether the differences between the mean values of two metric values were statistically significant, Student's t -test was used, with a significance level of 1%. To assess whether the Student's t -test could be applied to the results, the normality of the data was evaluated using the Shapiro–Wilk test. When the normality condition was not met, the Mann–Whitney test was applied.

3. Results

In this section, we present results concerning the following topics: lung segmentation, COVID-19 area segmentation via automatic lung segmentation, COVID-19 area segmen-

tation via manual lung segmentation, and a comparison between automatic COVID-19 segmentation and interobserver agreement.

3.1. Lung Segmentation Results

For both DP strategies, better results were obtained with U-Net + D + L (CT-scan strategy, Dice = 99.79%; slice strategy, Dice = 98.74%). Table 4 shows the results for lung segmentation in the test set when the best validation models were used.

As shown in Table 4, the results obtained with the slice strategy are better than the results obtained with the CT-scan strategy. Nevertheless, there are no statistically significant differences. For accuracy, the t -value = 2.17, $p > 0.01$ ($p = 0.019797$), 95% CI: CT scan—[96.293–98.907], and Slice—[98.395–99.125]. For global accuracy, the t -value = 2.07, $p > 0.01$ ($p = 0.024445$), 95% CI: CT scan—[98.135–99.885], and Slice—[99.572–99.748]. For Dice, the t -value = 2.42, $p > 0.01$ ($p = 0.011382$), 95% CI: CT scan—[91.886–98.694], and Slice—[97.907–98.673]. For Jaccard, the t -value = 2.44, $p > 0.01$ ($p = 0.010892$), 95% CI: CT scan—[91.805–98.495], and Slice—[97.719–98.541]. To ensure the validity of the Student's t -test, we applied the Shapiro–Wilk test to the accuracies, global accuracies, and Dice and Jaccard values, and obtained $W(27) = 0.77$, $p < 0.01$ ($p = 4.38 \times 10^{-5}$); $W(27) = 0.44$, $p < 0.01$ ($p = 4.12 \times 10^{-9}$); $W(27) = 0.53$, $p < 0.01$ ($p = 3.83 \times 10^{-8}$); and $W(27) = 0.55$, $p < 0.01$ ($p = 5.70 \times 10^{-8}$), respectively.

Table 4. Results of lung segmentation on in the test set via a U-Net network.

Strategy	Acc. (%)	Global Acc. (%)	Dice (%)	Jaccard (%)
CT scan	97.60 ± 2.00	99.01 ± 1.34	95.29 ± 5.21	95.15 ± 5.12
Slice	98.76 ± 0.76	99.66 ± 0.19	98.29 ± 0.83	98.13 ± 0.89

3.2. COVID-19 Segmentation Results

3.2.1. COVID-19 Segmentation Results with Automatic Lung Segmentation

For the U-NET network, the best results for the CT-scan and slice strategies were obtained with U-NET (Dice = 87.30%) and U-NET + A + D (Dice = 86.61%), respectively. For CNN2, the best results for the CT-scan and slice strategies were obtained with CNN2 + D (Dice = 87.30%) and CNN2 + A (Dice = 86.61%), respectively.

Table 5 shows the results obtained in the test set for COVID-19 segmentation with U-Net, which uses the best validation models. A comparison of the results of the CT-scan and slice strategies on the test set reveals that the latter presents better results. The differences between the Dice and Jaccard values are statistically significant: t -value = 3.93, $p < 0.01$ ($p = 0.00033$), 95% CI: CT scan—[66.823–79.197], Slice—[82.161–87.159] and t -value = 4.02, $p < 0.01$ ($p = 0.000308$), 95% CI: CT scan—[75.121–82.699], Slice—[84.885–88.535], respectively. Nevertheless, the differences between the accuracy and global accuracy values are not statistically significant: t -value = 0.98, $p > 0.01$ ($p = 1.67 \times 10^{-1}$), 95% CI: CT scan—[85.031–92.649], Slice—[89.003–92.477] and t -value = 0.67, $p > 0.01$ ($p = 2.56 \times 10^{-1}$), 95% CI: CT scan—[99.289–99.891], Slice—[99.538–99.842], respectively. To ensure the validity of the t -Student test, we applied the Shapiro–Wilk test to the accuracies, global accuracies, and Dice and Jaccard values, and obtained $W(27) = 0.88$, $p < 0.01$ ($p = 7 \times 10^{-3}$); $W(27) = 0.77$, $p < 0.01$ ($p = 7.61 \times 10^{-5}$); $W(27) = 0.87$, $p < 0.01$ ($p = 5 \times 10^{-3}$); and $W(27) = 0.90$, $p < 0.05$ ($p = 2.3 \times 10^{-2}$), respectively.

Table 5. Results of COVID-19 segmentation in the test set obtained with automatic lung segmentation and the U-Net method.

Strategy	Acc. (%)	Global Acc. (%)	Dice (%)	Jaccard (%)
CT scan	88.84 ± 5.83	99.59 ± 0.46	73.01 ± 9.47	78.91 ± 5.80
Slice	90.74 ± 3.76	99.69 ± 0.33	84.66 ± 5.41	86.71 ± 3.95

3.2.2. COVID-19 Segmentation Results with Manual Lung Segmentation

The results for COVID-19 segmentation via the CT-scan and slice strategies that are shown in Table 6 are similar to those shown in Table 5: the Dice and Jaccard values are statistically significant, whereas the accuracy and global accuracy values are not statistically significant. For accuracy, the t -value = 0.37, $p > 0.01$ ($p = 0.355785$), 95% CI: CT-scan—[86.792–94.828], Slice—[89.800–93.320]; for global accuracy, the t -value = 0.54, $p > 0.01$ ($p = 2.96 \times 10^{-1}$), 95% CI: CT-scan—[97.559–101.741], Slice—[98.426–101.014]; for Dice, the t -value = 3.43, $p < 0.01$ ($p = 1.13 \times 10^{-3}$), 95% CI: CT-scan—[67.976–80.964], Slice—[82.606–88.094]; and for Jaccard, the t -value = 3.49, $p < 0.01$ ($p = 9.81 \times 10^{-4}$). To ensure the validity of the t -Student test, we applied the Shapiro–Wilk test to the accuracies, global accuracies, and Dice and Jaccard values, and obtained $W(27) = 0.92$, $p < 0.05$ ($p = 0.04$); $W(27) = 0.81$, $p < 0.01$ ($p = 2.80 \times 10^{-4}$); $W(27) = 0.89$, $p < 0.05$ ($p = 1.3 \times 10^{-2}$); and $W(27) = 0.88$, $p < 0.05$ ($p = 1.1 \times 10^{-2}$), respectively.

For each DP strategy, the metrics shown in Table 6 are slightly better than those shown in Table 5 with automatic lung segmentation. Nevertheless, there are no statistically significant differences for the CT-scan strategy (accuracy t -value = 0.69, $p > 0.01$ ($p = 0.260415$); global accuracy t -value = 0.32, $p > 0.01$ ($p = 3.75 \times 10^{-1}$); Dice t -value = 0.29, $p > 0.01$ ($p = 3.84 \times 10^{-1}$); and Jaccard t -value = 0.34, $p > 0.01$ ($p = 3.71 \times 10^{-1}$)) or for the slice strategy (accuracy t -value = 0.64, $p > 0.01$ ($p = 2.65 \times 10^{-1}$); global accuracy t -value = 0.25, $p > 0.01$ ($p = 4.0 \times 10^{-1}$); Dice t -value = 0.35, $p > 0.01$ ($p = 3.63 \times 10^{-1}$); and Jaccard t -value = 0.40, $p > 0.01$ ($p = 3.45 \times 10^{-1}$)).

Table 6. Results of COVID-19 segmentation in the test set obtained with manual lung segmentation and using the U-Net network.

Strategy	Acc. (%)	Global Acc. (%)	Dice (%)	Jaccard (%)
CT-scan	90.81 $\pm$ 6.15	99.65 $\pm$ 0.32	74.47 $\pm$ 9.94	79.90 $\pm$ 6.04
Slice	91.56 $\pm$ 3.81	99.72 $\pm$ 0.28	85.35 $\pm$ 5.41	87.29 $\pm$ 4.36

3.3. Segmented CT-Scan Area

The automatically segmented areas of the lungs and COVID-19 patients in test sets, expressed as percentages of the CT-scan area, are compared with the same areas when segmented by radiologists. Figure 9 shows the box graphics for both the lung and COVID-19 segmentations compared with the radiologist segmentations.

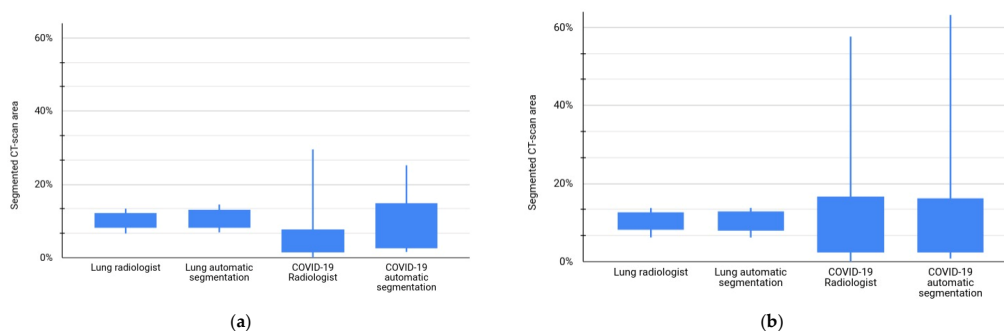


Figure 9. A comparison of the segmented areas of the lungs and COVID-19 patients in test sets, expressed as percentages of the CT-scan area, with the same area segmented by a radiologist: (a) CT-scan strategy; (b) slice strategy.

As shown in Figure 9, the percentage of area segmented using the slice strategy is very similar to the percentage of area segmented by the radiologists. These results could

be explained using Table 5, which compares automatic COVID-19 segmentation results for the CT-scan strategy and the slice strategy. As shown in this table, the metrics for the slice strategy are greater than the metrics for the CT-scan strategy, and the differences are statistically significant.

3.4. Comparison of COVID-19 Segmentation in Both Datasets with Interobserver Agreement

We compare the results of the best models for COVID-19 segmentation using the slice strategy with those of interobserver agreement. For this comparison, the 7 CT scans common to Dataset 1 and Dataset 2 were used.

As shown in Table 7, the automatic method using the slice strategy presented better metric values than interobserver agreement did in both datasets. Nevertheless, the differences between the results of the automatic method in both the datasets and interobserver agreement were not statistically significant. In this case, the Shapiro–Wilk test showed that Student’s *t*-test could not be applied. We therefore applied the Mann–Whitney test. For Dataset 1 $\times$ interobserver, we obtained the following results: accuracy: $U = 16, p > 0.01$ ($p = 1.54 \times 10^{-1}$); global accuracy: $U = 20, p > 0.01$ ($p = 3.05 \times 10^{-1}$); Dice: $U = 13, p < 0.01$ ($p = 7.93 \times 10^{-2}$); Jaccard: $U = 13, p > 0.01$ ($p = 7.93 \times 10^{-2}$). For Dataset 2 $\times$ interobserver, we obtained the following results: accuracy: $U = 6, p > 0.01$ ($p = 1.07 \times 10^{-2}$); global accuracy: $U = 24.5, p > 0.01$ ($p = 0.47608$); Dice: $U = 17, p > 0.01$ ($p = 1.87 \times 10^{-1}$); Jaccard: $U = 13, p > 0.01$ ($p = 7.92 \times 10^{-2}$).

Table 7. A comparison between the COVID-19 segmentation of the proposed method using the slice approach with the results of interobserver agreement.

Method	Acc. (%)	Global Acc. (%)	Dice (%)	Jaccard (%)
Automatic method in Dataset 1	89.64 $\pm$ 4.68	99.51 $\pm$ 0.43	83.21 $\pm$ 7.52	85.67 $\pm$ 5.45
Automatic method in Dataset 2	99.44 $\pm$ 3.3	99.50 $\pm$ 0.04	79.30 $\pm$ 7.90	82.90 $\pm$ 6.54
Interobserver agreement	88.08 $\pm$ 5.16	99.52 $\pm$ 0.43	77.69 $\pm$ 7.53	81.76 $\pm$ 4.89

3.5. Lung and COVID-19 Segmentation Examples

Figure 10 shows an example of lung and COVID-19 segmentation for a slice of a CT scan using a U-Net trained with a CT-scan strategy and a slice strategy. In the slice shown, there is a great coincidence between the areas segmented with the two strategies. Figure 11 shows an example of COVID-19 segmentation by both groups of radiologists for a slice of a CT scan. As shown, there is a small difference between the two segmentations.

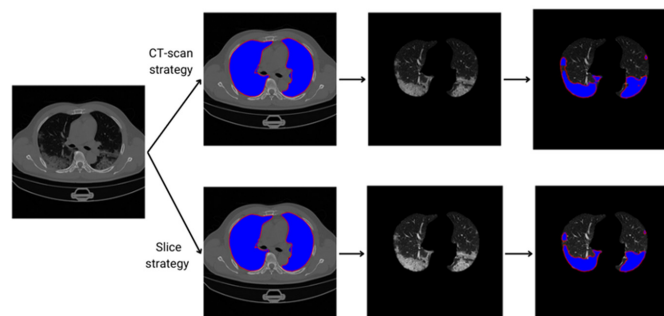


Figure 10. The first row shows the lung and COVID-19 images segmented with U-Net trained with the CT scan strategy. The second row shows the lung and COVID-19 images segmented with U-Net trained with the slice strategy. Radiologist segmentation (Dataset 1) is shown in blue, whereas automatic segmentation is shown in red.

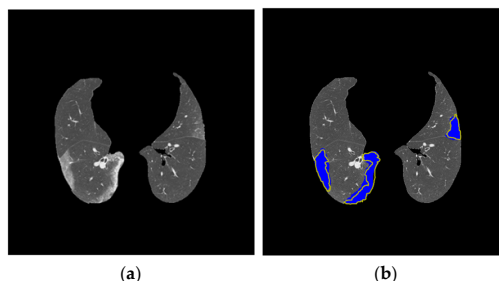


Figure 11. (a) Original image; (b) agreement between radiologists. Radiologists' segmentation in Dataset 1 is shown in blue, whereas radiologists' segmentation in Dataset 2 is shown in yellow.

In low-contrast areas in CT images, segmentation variations may occur, as shown in Figure 12. In these cases, the difficulty of distinguishing areas with similar pixel intensities compromises the identification of areas affected by COVID-19, as also identified by authors who used one-step segmentation and a CT scan strategy [5,6], one-step segmentation and a slice strategy [17], two-step segmentation and a CT strategy [20], and two-step segmentation and a slice strategy [22].

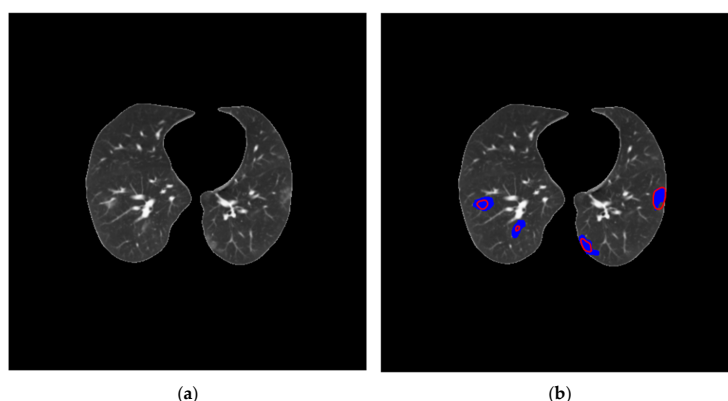


Figure 12. (a) Original image with low contrast; (b) radiologist segmentation (Dataset 1) is shown in blue, whereas automatic segmentation is shown in red.

3.6. Computational Cost

The time required to evaluate an image of 512×512 pixels is approximately 0.08 s for the CT-scan strategy and 0.07 s for the slice strategy. For an evaluation of a CT scan with 200 images, for example, this implies a time of 16 s using the CT-scan strategy and a time of 14 s using the slice strategy.

4. Discussion

From Table 2, which is the result of a literature review on the topic of two-step segmentation, we also observe that better Dice metric values are obtained with the slice strategy than with the CT-scan strategy. Tables 5 and 6 show that the results obtained in this study with two-step segmentation, with both DP strategies, are competitive with results previously obtained in the literature, which are shown in Table 2.

For example, for COVID-19 segmentation with the CT-scan strategy, the best values for the Dice metric obtained in the literature for two-step segmentation were 71.3% and 67.3% for [20,21], respectively. In this work, with automatic lung segmentation, a Dice

value of 73.01% was achieved. On the other hand, for COVID-19 segmentation with the slice strategy, the best values for the Dice metric obtained in the literature for two-step segmentation were 85.0% and 79.8% at [23,24], respectively. In this work, with automatic lung segmentation, a Dice value of 84.66% was achieved.

With respect to lung segmentation, the results obtained in this work are also state of the art. The best value in terms of the Dice metric obtained for the literature on lung and two-step segmentation was 97.18% [25], which was obtained with the slice strategy. Using the CT-scan strategy, the values were 87.90% for the right lung and 85.80% for the left lung in [21]. In this work, we obtained Dice values of 95.29% with the CT-scan strategy and 98.29% with the slice strategy.

U-Net networks were also used in several other previous studies on the segmentation of COVID-19 areas: [8,9,11,13,17]. In this study, although the datasets used are small, we credit the good results obtained to the benefit of combining a U-Net architecture with the following techniques to improve generalization: dropout, L_2 regularization, and data augmentation. Several combinations of these three techniques were used to obtain the best results.

In this study, we compared the performance of the deep models proposed for COVID-19 segmentation, where a slice strategy was used, with the results obtained using interobserver agreement. In the literature, to the best of our knowledge, only one study [9] has compared COVID-19 segmentation with automatic methods and interobserver agreement. For the CT-scan strategy, the authors obtained Dice values for the COVID-19-infected region segmentation results of the automatic method and those of two experienced radiologists of 0.74 ± 0.28 and 0.76 ± 0.29 , respectively, which were close to the interobserver agreement, i.e., 0.79 ± 0.25 .

In this work, using a slice-strategy, as shown in Table 7, we obtained Dice values of 83.21 ± 7.52 and 80.78 ± 9.45 for the COVID-19-infected region segmentations of the automatic method and the two groups of radiologists. These were Dataset 1 and Dataset 2, respectively. For interobserver agreement, we obtained a Dice value of 77.69 ± 7.53 . The differences between the Dice values of the automatic segmentation method in both datasets and the interobserver agreement method, however, were not statistically significant.

The Dice values of the automatic methods obtained in this work using the slice strategy are higher than the values obtained in [9] with the CT scan strategy. The interobserver Dice value obtained in this work is close to that obtained in [9]. Nevertheless, the standard deviation is greater.

Owing to the small number of common images in the two datasets, it was not possible to make a comparison via the CT-scan strategy. In future works, we intend to use larger datasets, evaluated by two groups of radiologists, to make this comparison possible.

A comparison of the results presented in Table 6 revealed that, for the detection of COVID-19 with two-step segmentation, when the U-Net network and manual lung segmentation were used, the values obtained for COVID-19 segmentation were better for the slice strategy than for the CT-scan strategy in terms of Dice and Jaccard metrics, with a statistically significant difference at the 1% level.

An evaluation of the results presented in Tables 5 and 6 revealed that, for the same metrics mentioned above, there were no statistically significant differences at the 1% significance level for the accuracy, global accuracy, and Dice and Jaccard metrics when manual or automatic segmentation of the lungs was used.

The COVID-19 pandemic has catalyzed the development of numerous diagnostic tools, including the algorithm presented in this study, which was initially designed for early detection and the risk stratification of patients. Despite overcoming the critical phase of the pandemic, the relevance of the proposed algorithm transcends the context of COVID-19, demonstrating potential application in the identification and quantification of pulmonary alterations with similar radiological characteristics (ground-glass opacities and pulmonary consolidations), often present in conditions such as pulmonary infarction, viral pneumonias, and hypersensitivity pneumonitis.

The segmentation of these areas also allows the quantification of lung involvement, providing an objective parameter for monitoring disease progression and responses to treatment. In scenarios of high demand for imaging services and overloaded health systems, the algorithm can be incorporated into screening systems, optimizing workflow and prioritizing the analysis of exams with suspected significant alterations.

We believe that the algorithm's ability to segment and quantify areas of lung alteration, with statistically non-significant differences when compared to the segmentation of two groups of radiologists, makes it a valuable tool with an impact beyond the context of COVID-19, contributing to improving the diagnosis and management of different lung diseases.

5. Conclusions

The main conclusions of this study are as follows:

1. Using U-NET, COVID-19 segmentation was slightly better for the slice strategy than for the CT-scan strategy. For Dice and Jaccard metrics, there was a difference statistically significant at the 1% significance level.
2. A comparison of the performance of the automatic COVID-19 segmentation results obtained with U-NET and the interobserver agreement in a group of 7 CT scans using the slice strategy revealed no statistically significant differences.
3. When U-Net was used, there were no statistically significant differences in terms of accuracy, global accuracy, and Dice and Jaccard metrics when manual or automatic segmentation of the lungs was performed.

Author Contributions: Conceptualization, A.d.S.O., M.G.F.C. and C.F.F.C.F.; methodology, A.d.S.O., M.G.F.C. and C.F.F.C.F.; software, A.d.S.O.; validation, J.P.G.F.C. and C.F.F.C.F.; formal analysis, A.d.S.O., M.G.F.C. and C.F.F.C.F.; investigation, A.d.S.O.; resources, A.d.S.O.; data curation, A.d.S.O.; writing—original draft preparation, A.d.S.O. and C.F.F.C.F.; writing—review and editing, J.P.G.F.C. and C.F.F.C.F.; visualization, M.G.F.C.; supervision, C.F.F.C.F.; project administration, M.G.F.C.; funding acquisition, M.G.F.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was carried out within the scope of the Samsung-UFAM Project for Education and Research (SUPER), according to Article 39 of Decree n°10.521/2020, and was funded by Samsung Electronics of Amazonia Ltd.a., under the terms of Federal Law n°8.387/1991 through agreement 001/2020, signed with UFAM and FAEPI, Brazil. The authors would also like to thank CAPES and FAPEAM for their financial support.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The datasets are available in [27,28].

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Pan American Health Organization (PAHO). *COVID-19: An Informative Guide*; PAHO: Washington, DC, USA, 2020.
2. World Health Organization (WHO). *Therapeutics and COVID-19: Living Guideline*; WHO: Geneva, Switzerland, 2023.
3. World Health Organization (WHO). *Use of Chest Imaging in COVID-19: A Rapid Advice Guide*; WHO: Geneva, Switzerland, 2020.
4. Habib, M.; Ramzan, M.; Khan, S.A. A Deep Learning and Handcrafted Based Computationally Intelligent Technique for Effective COVID-19 Detection from X-Ray/CT-Scan Imaging. *J. Grid Comput.* **2022**, *20*, 23. [[CrossRef](#)] [[PubMed](#)]
5. Kitrungrotsakul, T.; Chen, Q.; Wu, H.; Iwamoto, Y.; Hu, H.; Zhu, W.; Chen, C.; Xu, F.; Zhou, Y.; Lin, L.; et al. Attention-RefNet: Interactive Attention Refinement Network for Infected Area Segmentation of COVID-19. *IEEE J. Biomed. Health Inform.* **2021**, *25*, 2363–2373. [[CrossRef](#)] [[PubMed](#)]
6. Shan, F.; Gao, Y.; Wang, J.; Shi, W.; Shi, N.; Han, M.; Xue, Z.; Shen, D.; Shi, Y. Abnormal Lung Quantification in Chest CT Images of COVID-19 Patients with Deep Learning and Its Application to Severity Prediction. *Med. Phys.* **2021**, *48*, 1633–1645. [[CrossRef](#)] [[PubMed](#)]

7. Wang, G.; Liu, X.; Li, C.; Xu, Z.; Ruan, J.; Zhu, H.; Meng, T.; Li, K.; Huang, N.; Zhang, S. A Noise-Robust Framework for Automatic Segmentation of COVID-19 Pneumonia Lesions from CT Images. *IEEE Trans. Med. Imaging* **2020**, *39*, 2653–2663. [\[CrossRef\]](#) [\[PubMed\]](#)
8. Zhou, L.; Li, Z.; Zhou, J.; Li, H.; Chen, Y.; Huang, Y.; Xie, D.; Zhao, L.; Fan, M.; Hashmi, S.; et al. A Rapid, Accurate and Machine-Agnostic Segmentation and Quantification Method for CT-Based COVID-19 Diagnosis. *IEEE Trans. Med. Imaging* **2020**, *39*, 2638–2652. [\[CrossRef\]](#) [\[PubMed\]](#)
9. Li, Z.; Zhong, Z.; Li, Y.; Zhang, T.; Gao, L.; Jin, D.; Sun, Y.; Ye, X.; Yu, L.; Hu, Z.; et al. From Community-Acquired Pneumonia to COVID-19: A Deep Learning-Based Method for Quantitative Analysis of COVID-19 on Thick-Section CT Scans. *Eur. Radiol.* **2020**, *30*, 6828–6837. [\[CrossRef\]](#)
10. Liu, J.; Dong, B.; Wang, S.; Cui, H.; Fan, D.P.; Ma, J.; Chen, G. COVID-19 Lung Infection Segmentation with a Novel Two-Stage Cross-Domain Transfer Learning Framework. *Med. Image Anal.* **2021**, *74*, 102205. [\[CrossRef\]](#)
11. Hoang-Thi, T.N.; Vakalopoulou, M.; Christodoulidis, S.; Paragios, N.; Revel, M.P.; Chassagnon, G. Deep Learning for Lung Disease Segmentation on CT: Which Reconstruction Kernel Should Be Used? *Diagn. Interv. Imaging* **2021**, *102*, 691–695. [\[CrossRef\]](#)
12. Ranjbarzadeh, R.; Jafarzadeh Ghouschi, S.; Bendeache, M.; Amirabadi, A.; Ab Rahman, M.N.; Baseri Saadi, S.; Aghamohammadi, A.; Kooshki Forooshani, M. Lung Infection Segmentation for COVID-19 Pneumonia Based on a Cascade Convolutional Network from CT Images. *BioMed Res. Int.* **2021**, *2021*, 5544742. [\[CrossRef\]](#)
13. Amyar, A.; Modzelewski, R.; Li, H.; Ruan, S. Multi-Task Deep Learning Based CT Imaging Analysis for COVID-19 Pneumonia: Classification and Segmentation. *Comput. Biol. Med.* **2020**, *126*, 104037. [\[CrossRef\]](#)
14. Amara, K.; Aouf, A.; Kennouche, H.; Djekoune, A.O.; Zenati, N.; Kerdjidi, O.; Ferguene, F. COVIR: A Virtual Rendering of a Novel NN Architecture O-Net for COVID-19 Ct-Scan Automatic Lung Lesions Segmentation. *Comput. Graph.* **2022**, *104*, 11–23. [\[CrossRef\]](#) [\[PubMed\]](#)
15. Zhang, Q.; Ren, X.; Wei, B. Segmentation of Infected Region in CT Images of COVID-19 Patients Based on QC-HC U-Net. *Sci. Rep.* **2021**, *11*, 22854. [\[CrossRef\]](#) [\[PubMed\]](#)
16. Geng, P.; Tan, Z.; Wang, Y.; Jia, W.; Zhang, Y.; Yan, H. STCNet: Alternating CNN and Improved Transformer Network for COVID-19 CT Image Segmentation. *Biomed. Signal Process. Control* **2024**, *93*, 106205. [\[CrossRef\]](#)
17. Wu, Y.H.; Gao, S.H.; Mei, J.; Xu, J.; Fan, D.P.; Zhang, R.G.; Cheng, M.M. JCS: An Explainable COVID-19 Diagnosis System by Joint Classification and Segmentation. *IEEE Trans. Image Process.* **2021**, *30*, 3113–3126. [\[CrossRef\]](#)
18. Fan, D.P.; Zhou, T.; Ji, G.P.; Zhou, Y.; Chen, G.; Fu, H.; Shen, J.; Shao, L. Inf-Net: Automatic COVID-19 Lung Infection Segmentation from CT Images. *IEEE Trans. Med. Imaging* **2020**, *39*, 2626–2637. [\[CrossRef\]](#)
19. Fung, D.L.X.; Liu, Q.; Zammit, J.; Leung, C.K.S.; Hu, P. Self-Supervised Deep Learning Model for COVID-19 Lung CT Image Segmentation Highlighting Putative Causal Relationship among Age, Underlying Disease and COVID-19. *J. Transl. Med.* **2021**, *19*, 318. [\[CrossRef\]](#)
20. Gong, K.; Wu, D.; Arru, C.D.; Homayounieh, F.; Neumark, N.; Guan, J.; Buch, V.; Kim, K.; Bizzo, B.C.; Ren, H.; et al. A Multi-Center Study of COVID-19 Patient Prognosis Using Deep Learning-Based CT Image Analysis and Electronic Health Records. *Eur. J. Radiol.* **2021**, *139*, 109583. [\[CrossRef\]](#)
21. Ma, J.; Wang, Y.; An, X.; Ge, C.; Yu, Z.; Chen, J.; Zhu, Q.; Dong, G.; He, J.; He, Z.; et al. Toward Data-Efficient Learning: A Benchmark for COVID-19 CT Lung and Infection Segmentation. *Med. Phys.* **2021**, *48*, 1197–1210. [\[CrossRef\]](#)
22. Wu, D.; Gong, K.; Arru, C.D.; Homayounieh, F.; Bizzo, B.; Buch, V.; Ren, H.; Kim, K.; Neumark, N.; Xu, P.; et al. Severity and Consolidation Quantification of COVID-19 from CT Images Using Deep Learning Based on Hybrid Weak Labels. *IEEE J. Biomed. Health Inform.* **2020**, *24*, 3529–3538. [\[CrossRef\]](#)
23. Ilhan, A.; Alpan, K.; Sekeroglu, B.; Abiyev, R. COVID-19 Lung CT Image Segmentation Using Localization and Enhancement Methods with U-Net. *Procedia Comput. Sci.* **2022**, *218*, 1660–1667. [\[CrossRef\]](#)
24. Paluru, N.; Dayal, A.; Jenssen, H.B.; Sakinis, T.; Cenkramaddi, L.R.; Prakash, J.; Yalavarthy, P.K. Anam-Net: Anamorphic Depth Embedding-Based Lightweight CNN for Segmentation of Anomalies in COVID-19 Chest CT Images. *IEEE Trans. Neural Netw. Learn. Syst.* **2021**, *32*, 932–946. [\[CrossRef\]](#) [\[PubMed\]](#)
25. Kang, M.; Hong, K.S.; Chikontwe, P.; Luna, M.; Jang, J.G.; Park, J.; Shin, K.C.; Park, S.H.; Ahn, J.H. Quantitative Assessment of Chest CT Patterns in COVID-19 and Bacterial Pneumonia Patients: A Deep Learning Perspective. *J. Korean Med. Sci.* **2021**, *36*, e46. [\[CrossRef\]](#) [\[PubMed\]](#)
26. Saood, A.; Hatem, I. COVID-19 Lung CT Image Segmentation Using Deep Learning Methods: U-Net versus SegNet. *BMC Med. Imaging* **2021**, *21*, 19. [\[CrossRef\]](#) [\[PubMed\]](#)
27. Ma, J.; Ge, C.; Wang, Y.; An, X.; Gao, J.; Yu, Z.; Zhang, M.; Liu, X.; Deng, X.; Cao, S.; et al. COVID-19 CT Lung and Infection Segmentation Dataset. *Zenodo* 2020. [\[CrossRef\]](#)
28. MedSeg; Jenssen, H.B.; Sakinis, T. MedSeg COVID Dataset 2. *Figshare* **2021**. [\[CrossRef\]](#)
29. Gonzalez, R.C.; Woods, R.E. *Digital Image Processing*, 4th ed.; Pearson: New York, NY, USA, 2017.

30. Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. *arXiv* **2015**, arXiv:1505-04597.
31. Miyagawa, M.; Costa, M.G.F.; Gutierrez, M.A.; Costa, J.P.G.F.; Costa Filho, C.F.F. Detecting Vascular Bifurcation in IVOCT Images Using Convolutional Neural Networks with Transfer Learning. *IEEE Access* **2019**, *7*, 66167–66175. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.