

FÁBIO ARTHUR SOARES ARAUJO

**CLASSIFICAÇÃO DAS FASES DO *PLASMODIUM* EM IMAGENS DO EXAME DE
GOTA ESPESSA PARA MALÁRIA**

MANAUS- AM

2023

FÁBIO ARTHUR SOARES ARAUJO

CLASSIFICAÇÃO DAS FASES DO PLASMODIUM EM IMAGENS DO EXAME
DE GOTA ESPESSA PARA MALÁRIA

Pesquisa desenvolvida durante a disciplina de Trabalho de Conclusão de Curso e apresentada à banca avaliadora do Curso de Engenharia Elétrica Eletrônica da Universidade Federal do Amazonas, como pré-requisito para obtenção do título de Engenheiro Elétrico Eletrônico.

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

MANAUS - AM
2023

Ficha Catalográfica

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

A663c Araujo, Fábio Arthur Soares
Classificação das fases do plasmodium em imagens do exame de gota espessa para a malária / Fábio Arthur Soares Araujo . 2023
48 f.: il.; 31 cm.

Orientadora: Marly Guimarães Fernandes Costa
Coorientador: Cícero Ferreira Fernandes Costa Filho
TCC de Graduação (Engenharia Elétrica - Eletrônica) -
Universidade Federal do Amazonas.

1. Malária. 2. Diagnóstico. 3. Microscopia de gota de sangue. 4.
Aprendizado profundo. 5. Inteligência artificial. I. Costa, Marly
Guimarães Fernandes. II. Universidade Federal do Amazonas III.
Título

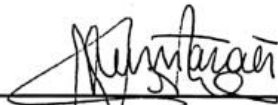
Fábio Arthur Soares Araujo

Classificação das Fases do Plasmodium em Imagens do Exame de Gota Espessa para
Malária

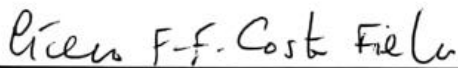
Monografia apresentada à Coordenação do
Curso de Engenharia Elétrica da Universi-
dade Federal do Amazonas, como parte dos
requisitos necessários à obtenção do título de
Engenheiro Eletricista.

Aprovado em 08 de novembro de 2023

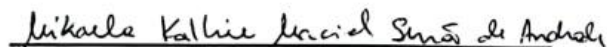
Banca Examinadora:



Prof^a. Dr^a. Marly Guimarães Fernandes Costa
Universidade Federal do Amazonas - UFAM



Prof. Dr. Cícero Ferreira Fernandes Costa
Universidade Federal do Amazonas - UFAM



Prof^a. MSc. Mikaela Kalline Maciel Serrão
Universidade Federal do Amazonas - UFAM

MANAUS-AM
2023

AGRADECIMENTOS

Gostaria de agradecer a minha família que durante toda a minha jornada na graduação me deram o apoio que eu precisava para continuar avançando e que sempre poderei contar para todas as escolhas que fiz durante minha vida.

Aos meus web-amigos do nosso grupo intitulado “TAVERNA” que durante a graduação e escrita desta monografia estiveram comigo, trazendo a diversão que faz mesmo um dia mais cansativo terminar bem. Em especial ao Chaves, meu melhor amigo, que esteve comigo em praticamente todas as noites dos meus dias durante os últimos 5 anos. Além dele, agradeço a Sailent, Meat Puppet e Gomes., quatro pessoas especiais.

Agradeço aos amigos que a graduação me apresentou. Obrigado a Luiz Gustavo, Carlos Santos, Gabriel Freitas, Josekson, Marcos Pontes e Paulo Freitas. Pessoas incríveis que conviveram com as dificuldades do curso, assim como eu, que também estavam comigo nos outros momentos que a Faculdade proporciona, sendo companhia para almoçar e conversar durante horários vagos.

Agradeço as amigadas que fiz perto do fim da graduação que já considero muito como Dar’c e Alen, pessoas que mudaram um pouco meu jeito de viver a vida.

Agradeço às pessoas do laboratório de controle *e-Controls*, onde fiz meu estágio, que fizeram esse período algo especial e inesquecível. Um laboratório de pessoas bem-humoradas e divertidas e que me proporcionou diversos conhecimentos através da troca de informação entre alunos e professores.

E por fim agradeço aos meus professores que me guiaram nessa jornada em IA, a área que decidi seguir meus conhecimentos. Agradeço ao professor Frederico Pinagé, aquele quem me apresentou a inteligência artificial no projeto SUPER ICT-Júnior, a professora Marly Guimarães e o professor Cícero Costa Filho que durante o Super ICT-Senior me guiaram nesta área, confirmando meus objetivos em seguir nela para a vida.

RESUMO

A malária é uma doença potencialmente fatal que se espalha para os seres humanos através da picada das fêmeas do mosquito *Anopheles* infectadas com espécies de parasitas *Plasmodium*. O diagnóstico rápido da malária é recomendado pela OMS para todos os pacientes com suspeita de malária, antes de receberem tratamento. A microscopia de esfregaço de gota de sangue é um dos exames diagnósticos recomendados. No entanto, a microscopia de esfregaço de gota de sangue requer experiência, é demorada e está sujeita à variabilidade intra e inter microscopista. Por outro lado, a microscopia automatizada tem potencial de superar os problemas mencionados. Neste contexto, alguns métodos computacionais baseados em aprendizado de máquina para detecção e classificação de objetos foram desenvolvidos. Este trabalho propõe a detecção e identificação das fases do *plasmodium* em imagens de microscopia de campo claro de esfregaços de sangue. Foi analisado nove modelos de redes profundas através de um estudo de ablação em um processo de 2 etapas: a primeira realizando uma classificação binária e a segunda classificando os parasitas detectados na primeira etapa em uma de suas 4 fases. Os modelos foram testados com um conjunto de dados de imagens de domínio público. Os melhores resultados foram obtidos com a rede profunda Efficient Net B6 com otimizador RMSProp. Com esta rede obteve-se uma acurácia de 89,91% e um F1-score de 83,88% o que superou os resultados apresentados em estudos anteriores.

Palavras-chave: Malária, Diagnóstico, Microscopia de gota de sangue, Aprendizado Profundo, Inteligência Artificial, Efficient Net.

ABSTRACT

Malaria is a potentially fatal disease that spreads to humans through the bite of female *Anopheles* mosquitoes infected with *Plasmodium* parasite species. Prompt diagnosis of malaria is recommended by the WHO for all patients with suspected malaria before they receive treatment. Blood smear microscopy is one of the recommended diagnostic tests. However, blood smear microscopy requires experience, is time-consuming and is subject to intra- and inter-microscopic variability. Automated microscopy has the potential to overcome these problems. In this context, some computational methods based on machine learning for object detection and classification have been developed. This work proposes the detection and identification of *Plasmodium* stages in brightfield microscopy images of blood smears. Nine deep network models were analyzed through an ablation study in a 2-stage process: the first performing a binary classification and the second classifying the parasites detected in the first stage into one of its 4 phases. The models were tested with a dataset of public domain images. The best results were obtained with the Efficient Net BX deep network. With this network, an accuracy of 89,91% and an F1-score of 83,88% were obtained, which surpassed the results presented in previous studies.

Palavras-chave: Malaria, Diagnosis, Deep Learning, *Plasmodium*, blood smear microscopy, Artificial Intelligence, Efficient Net.

LISTA DE FIGURAS

Figura 1: Ilustração de balanceamento de amostras	26
Figura 2: Camada de convolução	28
Figura 3: Ilustração das camadas de <i>pooling</i>	28
Figura 4 - Metodologia para classificação do <i>Plasmodium</i>	31
Figura 5 - Imagem de microscopia da base de dados	32
Figura 6 - Exemplo de células: (a) célula não infectada e (b-e) células infectadas mostrando o ciclo de vida do plasmodium (b) anel, (c) trofozoíto, (d) esquizonte e (e) Gametócito	32
Figura 7 - Comparação da arquitetura das redes (a) EfficientNet B0 e (b) EfficientNet B7	36
Figura 8 - Arquitetura da rede DenseNet121	37
Figura 9 - Arquitetura da rede CNN proposta	37
Figura 10: Matriz de confusão da classificação binária	40
Figura 11: Matriz de confusão do modelo multiclasse para a parcela de teste	42
Figura 12: Gráficos de acurácia e perda para o treinamento multiclasse	43

LISTA DE TABELAS

Tabela 1. Sumário dos artigos que utilizam o mesmo conjunto de dados deste trabalho (BBBC01v1)	13
Tabela 2. Quantidade de células RBC por classe, extraídas de (Ljosa et al, 2012)	33
Tabela 3: Divisão da base de dados para a tarefa de classificação binária	34
Tabela 4: Divisão da base de dados para a tarefa de classificação binária	34
Tabela 5: Resultados obtidos para classificação binária	40
Tabela 6: Comparação do desempenho da classificação multiclasse entre as redes EfficientNet, DenseNet e CNN proposta, com os três otimizadores	41
Tabela 7: Comparação do desempenho da classificação binária com a rede EfficientNet B7 e otimizador RMSProp com outros trabalhos da literatura	44
Tabela 8: Comparação do desempenho da classificação multiclasse com a rede EfficientNet B6 e otimizador RMSProp com outros trabalhos da literatura	44

LISTA DE SIGLAS

CNN	Rede neural convolucional (do inglês <i>Convolutional Neural Network</i>)
RBC	Glóbulos Vermelhos (do inglês <i>Red Blood Cell</i>)
SSC	Classificação em um estágio (do inglês <i>Single Stage Classification</i>)
TSC	Classificação em dois estágios (do inglês <i>Two Stage Classification</i>)

SUMÁRIO

1	Introdução	14
1.1	Objetivos Gerais	17
1.2	Objetivos específicos	17
1.3	Organização do Trabalho	17
2	Revisão Bibliográfica	18
2.1	<i>Applying Faster RCNN for Object Detection on Malaria Images</i> (Hung et al, 2017)	18
2.2	<i>A comparative analysis of deep learning architectures on high variation malaria parasite classification dataset.</i> (Rahman et al, 2021)	19
2.3	<i>Malaria Detection and Classification.</i> (Manku et al, 2020)	19
2.4	<i>Deep learning based malaria-infected cell detection and parasite life stage classification method.</i> (Chen et al, 2022)	20
2.5	<i>A dataset and benchmark for malaria life-cycle classification in thin blood smear images.</i> (Arshad et al, 2022)	20
3	Referencial Teórico	24
3.1	Malária	24
3.2	Pré-processamento de dados	25
3.2.1	Redimensionamento de Imagens	25
3.2.2	Balanceamento por Subamostragem	26
3.2.3	Definição de Pesos por Classe	26
3.3	Redes Neurais Convolucionais	27
3.3.1	Camada Convolutiva (<i>Convolutional Layer</i>)	27
3.3.2	Camada de Subamostragem (<i>Pooling</i>)	28

3.3.3 Camada <i>Dropout</i>	29
3.3.4 Camada ReLU (<i>Rectified Linear Unit</i>)	29
3.3.5 Regularização L2	29
3.4 Otimizadores	29
3.4.1 Estimativa Adaptativa de Momentos (ADAM)	30
3.4.2 Propagação da Raiz Média Quadrática (RMSProp)	30
3.4.3 Gradiente Descendente Estocástico com Momento (SGDM)	30
4 Materiais e Métodos	31
4.1 Base de Dados	31
4.2 Pré-processamento	33
4.3 Conjuntos de treino, validação e teste	34
4.4 Modelos Implementados	35
4.4.1 EfficientNet	35
4.4.2 DenseNet-121	36
4.4.3 CNN Proposta	37
4.5 Métricas de Avaliação	38
4.6 Ambiente de Desenvolvimento	39
5 Resultados e Discussões	39
5.1 Resultados classificação binária	39
5.2 Resultados classificação multiclasse	40
5.3 Discussões	43
5.4 Comparação com a literatura	44
6 Conclusão	44

INTRODUÇÃO

A malária é uma doença febril aguda potencialmente fatal, causada pelo parasita *Plasmodium*. É transmitida às pessoas através da picada de mosquitos Anopheles fêmeas infectadas. Encontra-se sobretudo nos países tropicais. Dos cinco tipos de *Plasmodium* que infectam os seres humanos, dois são os mais comuns: *P. vivax* e *P. falciparum*. De acordo com o último relatório mundial sobre a malária, o número estimado de mortes por malária em 2021 foi de 619.000. Durante os dois anos de pico da pandemia (2020-2021), as perturbações relacionadas com a COVID-19 provocaram mais 63.000 mortes por malária. O diagnóstico precoce e exato é essencial para o controle eficaz do desenvolvimento da doença (WHO, 2021 e 2022).

A microscopia de campo claro é utilizada por rotina para o diagnóstico parasitológico da malária. Trata-se de um exame de esfregaço (espesso ou fino) de sangue efetuado por um microscopista que procura pelos parasitas da malária em campos microscópicos do esfregaço. Quando bem executado, fornece um diagnóstico altamente sensível e específico da malária e permite a quantificação dos parasitas da malária e a identificação das espécies infectantes. Exige do microscopista conhecimentos especializados e muita atenção. O tempo de execução é alto, além de, estar sujeito a variabilidade intra e inter-microscopista. A análise automatizada de imagens microscópicas de campo tem o potencial de suplantar os problemas mencionados com respeito ao exame. Diferentemente da análise visual realizada pelo microscopista, que em função do volume de exames pode apresentar diferenças diagnósticas causadas pela fadiga visual, os métodos automatizados apresentam repetibilidade, independente do volume de imagens de campos microscópicos realizados. Podem ocorrer erros, mas estes são sistemáticos e podem ser contornados.

Numa pesquisa bibliográfica, é possível notar dois tipos de estudos que realizam a análise automatizada de imagens de campos microscópicos de esfregaços de sangue, utilizando técnicas de aprendizagem profunda para emular a visualização e análise do microscópio óptico. No primeiro grupo, os autores tinham como objetivo fazer uma classificação binária (distinguir entre eritrócitos infectados ou não infectados com parasitas da malária). Os trabalhos de Hung *et al* (2017) e Rahman *et al* (2021) são exemplos deste grupo. O trabalho de Hung *et al* alcançou uma acurácia de 98% e um F1-score de 99,11% enquanto o de Rahman *et al* alcançou uma acurácia de 99,35% e um F1-score de 96,85%. Considerando que a identificação das

diferentes morfologias dos parasitas em todo o ciclo de vida do *Plasmodium* é crucial para a realização de um diagnóstico correto através do exame microscópico de esfregaços sanguíneos, para além da detecção da presença de parasitas, os autores do segundo grupo efetuaram uma classificação da fase do ciclo de vida do *plasmodium* (anel, trofozoíto, esquizonte e gametócito). Como exemplos, citamos o trabalho (Manku *et al*, 2020), que fez uma classificação de cinco classes (eritrócito não infectado e as quatro classes correspondentes ao ciclo de vida do parasita no eritrócito infectado). O desempenho da proposta foi: acurácia de 82% e F1-score de 82%. O trabalho (Chen *et al*, 2022) apresentou uma acurácia de 86,60% e F1-score de 83,30% e o trabalho (Arshad *et al*, 2022), obteve uma acurácia de 79,61% e F1-score de 82,04%. Nestes trabalhos, os autores utilizam as seguintes arquiteturas de redes neurais convolucionais (CNN): VGG16, VGG19, ResNet50v2, DenseNet169, DenseNet201, Faster R-CNN e Xception.

Apesar do seu potencial, a EfficientNet ainda não foi utilizada nesta tarefa. A EfficientNet é uma rede neural profunda que busca otimizar os parâmetros de largura, profundidade e resolução utilizando um método de escalonamento composto. Intuitivamente, para que uma rede neural processe uma imagem maior, será necessária mais profundidade (número de camadas). Esse método consiste em, de maneira equilibrada, ajustar simultaneamente a profundidade, largura e resolução da rede. A profundidade, que se refere ao número de camadas, influencia a capacidade da rede de aprender características mais complexas; a largura, ou o número de neurônios por camada, afeta sua habilidade de capturar mais características; e a resolução, o tamanho da imagem de entrada, determinando o nível de detalhe visível à rede. Usando o coeficiente composto, o escalonamento é feito de forma uniforme em todas essas dimensões, resultando em um aumento proporcional do tamanho da rede, o que é uma abordagem mais sistemática em comparação com o ajuste isolado de cada dimensão (Tan *et al*, 2019).

As camadas da rede EfficientNet são formadas principalmente por blocos *MBConv* (*Mobile inverted bottleneck convolution*) com a adição de blocos de otimização *squeeze-and-excitation*. A partir da primeira arquitetura, B0, o método de escalonamento composto é aplicado para gerar as demais arquiteturas: B1 a B7. B0 é a arquitetura de menor complexidade e B7 a arquitetura de maior complexidade, concebida para tratar imagens de maior resolução.

Neste trabalho, compara-se o desempenho de todas as redes pré-treinadas EfficientNet desde a B0 até a B7, com a rede pré-treinada DenseNet121 e com uma CNN cuja arquitetura é proposta. Em contraste com as arquiteturas EfficientNet e DenseNet, esta arquitetura proposta é treinada de raiz. O processo de classificação, tal como nos trabalhos do segundo grupo, tem

dois níveis: o primeiro compreende uma classificação binária em que as classes anel, trofozoíto, esquizonte e gametócito são consideradas como uma única classe, de eritrócitos infectados (glóbulos vermelhos, RBCs), onde nesta tem como objetivo ter uma menor perda de quantidade de imagens quando houver o balanceamento de dados levando em conta a maior proporção de células saudáveis. No segundo nível, os eritrócitos infectados são classificados de acordo com o estágio do parasita em quatro classes (anel, trofozoíto, esquizonte e gametócito), assim com uma menor discrepância de quantidade de dados é possível fazer um balanceamento mais eficiente.

1.1 Objetivo Geral

Comparar desempenho de redes neurais profundas, entre elas redes convolucionais pré-treinadas de diversas complexidades e uma rede proposta na classificação dos ciclos de vida do *Plasmodium*.

1.2 Objetivos específicos

- Realizar estudo de ablação com as redes EfficientNet desde a arquitetura B0 mais simples até a B7 de maior complexidade, com a Rede DenseNet121 e uma rede CNN, utilizando 3 otimizadores ADAM, RMSProp e SGDM na tarefa de classificação dos ciclos de vida do *Plasmodium*.
- Realizar *benchmark* de trabalhos propostos na literatura e as redes treinadas neste trabalho.

1.3 Organização do Trabalho

- Capítulo 1: Introdução
- Capítulo 2: Revisão Bibliográfica
- Capítulo 3: Referencial Teórico
- Capítulo 4: Materiais e Métodos
- Capítulo 5: Resultados e discussão

O Capítulo 1 consiste em introduzir o trabalho, citando o processo de inspeção visual realizado na classificação do *Plasmodium* contextualizando e justificando a proposta, além de uma breve revisão dos artigos que já se debruçaram sobre o tema. Por fim, apresenta-se os objetivos deste trabalho.

O Capítulo 2 consiste na apresentação de trabalhos que foram publicados na literatura na área de classificação do *Plasmodium* utilizando as ferramentas de aprendizado profundo, em especial as CNNs pré treinadas.

O Capítulo 3 consiste em apresentar os principais fundamentos teóricos utilizados no desenvolvimento deste trabalho: *Plasmodium*, pré-processamento de dados, utilização de pesos para balanceamento de dados, redes neurais artificiais, redes neurais convolucionais e os métodos de otimização.

No Capítulo 4 apresenta-se o banco de dados utilizado e os métodos, como o pré-processamento dos dados, a arquitetura proposta, as arquiteturas pré treinadas e o diagrama de fluxo do treinamento.

O Capítulo 5 apresenta os resultados obtidos durante o período da pesquisa realizada, assim como comparações de resultados com os presentes na literatura, assim podendo esclarecer o diferencial do trabalho.

2 Revisão Bibliográfica

Este capítulo é dedicado a apresentar as técnicas presentes na literatura para fazer a classificação da fase do *Plasmodium*, destacando a metodologia e materiais utilizados nos trabalhos, realizando-se análises técnicas dos mesmos. Ao fim da análise é apresentada uma tabela de comparação dos trabalhos.

2.1 Applying Faster RCNN for Object Detection on Malaria Images (Hung *et al*, 2017)

Neste trabalho apresenta-se uma base de dados constituídas de 1300 imagens de campos microscópicos de lâminas de gota sangue espessa, A partir delas são extraídos cerca de 100.000 *patches* (sub imagens contendo uma célula de sangue infectada ou não com *Plasmodium vivax*, cujas amostras foram coletadas de pacientes de Manaus, Brasil e Tailândia. Os esfregaços de sangue foram corados com o reagente Giemsa. Cada patch é rotulado por um especialista em um dos 7 rótulos apresentados a seguir: 1. RBC (células vermelhas do sangue), 2. leucócitos, e as células infectadas com as diferentes fases do *plasmodium* (3. gametócito, 4. anel, 5. trofozoíto, 6. esquizonte) e 7. *patch* difícil de classificar. RBC e leucócitos são do tipo saudável, logo células normais de encontrar no sangue, enquanto as demais são células infectadas e, por fim, difíceis de classificar, células cuja fase não é facilmente distinguível, portanto, recebendo esse rótulo e conseqüentemente não sendo usadas para treinamento de redes. Após a rotulação, a base de dados resultou em classes desbalanceadas com 97% do total de células encontradas na base de dados sendo de RBCs não infectadas.

Os autores realizaram uma classificação dos *patches* do tipo binária, com objetivo de classificar cada *patch* como RBC normal ou RBC infectada. Essa tarefa foi realizada em duas etapas, uma com uma rede Faster R-CNN para detecção de objetos e a outra com a rede AlexNet para classificação. Os experimentos resultaram nas seguintes métricas: Acurácia = 98%, Precisão = 98,83%, Especificidade = 99,41%, Sensibilidade = 98,43%, F1-score = 99,11%.

2.2 A comparative analysis of deep learning architectures on high variation malaria parasite classification dataset. (Rahman *et al*, 2021)

Neste trabalho foi utilizado o banco de dados BBBC041v1 uma base de dados com

1.364 imagens de campos microscópicos de lâminas de gota de sangue espessa de pacientes com suspeita de malária *p. vivax*. O Banco fornece as anotações de posição de cada célula e seus respectivos rótulos (RBC não infectada, leucócitos e as 4 fases da infecção das RBCs: gametócito, anel, trofozoíto e esquizonte). Neste trabalho foi realizada apenas uma classificação binária das RBCs (saudável ou infectada). Ressalta-se o desbalanceamento entre as classes com 56.149 amostras saudáveis e 2.099 na classe infectada para o conjunto de treinamento e para o conjunto de teste 5.098 amostras não infectadas e 299 amostras infectadas. Os autores utilizaram 4 tipos de redes, VGGNet, ResNet50, Xception, Color constancy, uma rede de transferência de aprendizado treinada com a base de dados da NIH, a TL-NIH e a CIS.

Para avaliar essas redes foram utilizadas 7 métricas: acurácia, F1-Score, sensibilidade, especificidade, valor preditivo positivo e negativo (PPV e NPV respectivamente) e AUC. Para os testes foram utilizados o otimizador Adam e o melhor resultado foi utilizando a rede TL-NIH: Acurácia = 99,35%, F1-Score = 96,85%, sensibilidade = 92,31%, especificidade = 99,76%, PPV = 95,83%, NPV = 99,55%, AUC = 96,03%.

2.3 *Malaria Detection and Classification.* (Manku et al, 2020).

Neste trabalho foi utilizado também o banco de dados BBBC041v1, já descrito no item 2.2. Para o treinamento desta rede foi utilizado uma metodologia em 2 etapas onde a primeira etapa consistiu em uma classificação binária (célula saudável e célula infectada) utilizando uma rede Faster R-CNN e a rede pré-treinada ResNet-50, para a classificação multiclasse das células infectadas nas diversas fases de vida do plasmodium. Para minimizar o desbalanceamento entre classes foi realizado uma sub-amostragem do banco resultando em 140 imagens por classe e 560 imagens no total, separadas em 90% para treino e 10% para teste. A proposta resultou no seguinte desempenho: acurácia de 82%, F1-Score de 82%, sensibilidade de 82% e precisão de 83%

2.4 *Deep learning based malaria-infected cell detection and parasite life stage classification method.* (Chen et al, 2022)

Neste trabalho foi utilizado a base de dados BBBC01v1 juntamente com a base de dados de Abbas et al (2020) com o objetivo de classificar não apenas o *plasmodium vivax* mas também o *plasmodium falciparum*, assim abrangendo os dois tipos de malária mais comuns. As amostras foram divididas na proporção de 7 para 3 em treino e teste. A classificação também foi realizada em duas etapas: a primeira usando uma Faster R-CNN para fazer uma classificação binária e uma segunda, acrescida de um pré-processamento de normalização da

imagem, faz uma classificação multiclasse utilizando a rede ResNet101 como modelo de classificação. O desempenho da proposta foi: acurácia de 86,8% e F1-Score de 83,3%.

2.5 A dataset and benchmark for malaria life-cycle classification in thin blood smear images. (Arshad et al, 2022)

Neste trabalho é proposto uma base de dados chamada de IML-Malaria, esta base tem amostras coletadas no Paquistão, um país que vive da agricultura e de clima quente úmido, portanto uma região muito propícia a ter infecções de malária. A base de dados, que não é pública, contém 345 imagens microscópicas de lâminas de sangue que contém, em média, 111 células de sangue, resultando em 38.449 imagens de células separadas (patches). Todas foram rotuladas como saudáveis ou, no caso de células infectadas, como anel, trofozoíto, esquizonte e gametócito. Suas respectivas posições na imagem do campo microscópico correspondente, são também fornecidas. Uma característica da base é o número pequeno de imagens das classes trofozoíto e esquizonte, resultado em uma base com um desbalanceamento entre as classes.

A tarefa de classificação foi feita em duas abordagens. A primeira, classificação multiclasse. Uma classe é a de célula saudável e as demais, as 4 fases do *plasmodium*, totalizando 5 classes. A rede utilizada foi a DenseNet201, a qual apresentou o seguinte desempenho: acurácia de 74,56% e F1-Score de 75,52%. A segunda abordagem foi implementada em duas etapas: uma classificação binária (células saudáveis e infectadas), seguida por uma classificação multiclasse das células infectadas. Os melhores desempenho tanto na classificação binária quanto a multiclasse foi obtida com a rede ResNet50v2 (79,61% de acurácia e F1-Score de 82,04%). Os autores também desenvolveram uma aplicação em dispositivo móvel da proposta.

Tabela 1. Sumário dos artigos que utilizam o mesmo conjunto de dados deste trabalho (BBBC01v1)

Citação	Métodos		Melhor resultado (%)	
	Pré-processamento	Classificação/Técnica		
Arshad <i>et al</i> (2022)	Remoção da região escura das imagens: corte das imagens e aplicação do método de limiar	Classificação multiclasse em uma etapa: treinado com diferentes modelos CNN (pré-treinadas na imageNet, VGG16, VGG19, ResNet50v2, DenseNet169, DenseNet201 e uma arquitetura proposta)	Classificação multiclasse em duas etapas : 1a. etapa (classificação binária - saudáveis e infectados com malária) 2a. etapa (5 classes) : diferentes modelos CNN (ResNet50v2, VGG16, VGG19, DenseNet169 and DenseNet201 e uma arquitetura proposta	classificação em duas etapas; Arquitetura: ResNet50v2 pré-treinada; Acurácia Média=79,61 e F1-Score=82,04
Hung J <i>et al</i> (2017)	Aumento de dados: rotações, giros verticais e horizontais, mudança de canal de cores e mudança de escala.	SSC para uma classificação multiclasse: Faster R-CNN usando a CNN AlexNet com 7 camadas para detectar objetos vs. não objetos e para determinar a classe dos objetos.	Classificação em 2 estágios (TSC): (classificação binária- RBC ou outro) as detecções marcadas como outros (não RBC) foram utilizadas em uma rede AlexNet para obter um recurso dimensional 4096 vetor usado para classificá-los em categorias mais refinadas.	Classificação em duas etapas: Arquitetura R-CNN Acurácia média=98%
Manku <i>et al</i> (2020)	Limitando o número de imagens para cada classe em 140 imagens por classe para obter uma base de dados mais equilibrada	SSC para uma classificação binária: uma Faster R-CNN foi usada para detectar em RBC ou não RBC	Classificação em dois estágios (TSC): (classificação multiclases) foram inseridas na segunda camada as imagens classificadas na primeira etapa na rede pré-treinada ResNet-50 . Esse modelo obteve um vetor de recursos de tamanho 2048.	Arquitetura: ResNet50v2 pré-treinada; Acurácia média=82, F1-Score=82, Sensibilidade=82, precisão=83.

Citação	Métodos		Melhor resultado (%)
	Pré-processamento	Classificação/Técnica	
Chen <i>et al</i> (2022)	normalização RGB	SSC realizando uma classificação multiclasse: ResNet101 com uma Rede Pirâmide de Recursos foi obtido para a entrada da R-CNN para obter uma melhor extração de recursos.	Arquitetura: ResNet101 pré-treinada; Acurácia média=86,6, F1-Score=83,3
A Rahman <i>et al</i> (2021)	Ponderação das classes	Classificação binária: treino com diversas redes CNN (VGG16, VGG19, ResNet50, Xception, Color constancy, transferência de aprendizado com a base NIH (TL-NIH), Conditional image synthesis (CIS))	(Classificação binária SSC - 2 rótulos (RBC vs não infectado)) - : modelos CNNs treinados na ImageNet (VGG16, VGG19, ResNet50, Xception, Color constancy, transferência de aprendizado com a base NIH (TL-NIH), Conditional image synthesis (CIS)) Arquitetura: TL-NIH pré-treinada; : Acurácia=99,35, F1-Score=96,85, Sensibilidade=92,31, Especificidade=99,76

3 Referencial Teórico

Inicialmente serão apresentadas informações sobre a Malária (seção 3.1), com objetivo de entender melhor este trabalho. Na seção 3.2 apresenta-se as técnicas de pré-processamento de dados para tratar inconstâncias na base de dados. Em seguida as redes neurais convolucionais (CNNs), o tipo de rede utilizada durante o desenvolvimento deste trabalho (seção 3.3). Por fim, são apresentados os otimizadores que foram usados durante o treino das CNNs (seção 3.4).

3.1 Malária

A malária é uma doença parasitária transmitida por mosquitos que continua sendo um importante desafio de saúde pública em todo o mundo. O agente causador da malária é o *Plasmodium*, um parasita intracelular obrigatório que passa por um ciclo de vida complexo, envolvendo diferentes estágios no hospedeiro humano e no mosquito. A classificação das fases do *Plasmodium* é essencial para entender a biologia do parasita e desenvolver estratégias de prevenção e controle da malária (WHO, 2021 e 2022).

De acordo com Bardin (2011), o ciclo de vida do *Plasmodium* pode ser dividido em duas fases distintas: a fase assexuada e a fase sexuada. A fase assexuada, que é fundamental para a patogenia da malária, ocorre no interior dos glóbulos vermelhos do hospedeiro humano e é composta pelas seguintes etapas:

1. Fase Anel: Após a invasão dos glóbulos vermelhos, o *Plasmodium* assume uma forma inicial conhecida como estágio anel, devido à sua aparência morfológica semelhante a um anel. Neste estágio, o parasita começa a metabolizar a hemoglobina e se estabelece dentro da célula.
2. Trofozoíto: À medida que o parasita cresce, ele se transforma na forma trofozoítica, uma fase ativa em que ocorre intenso consumo de nutrientes da célula hospedeira e uma rápida multiplicação de sua massa celular.
3. Esquizonte: O trofozoíto amadurece em um esquizonte, onde ocorre a divisão nuclear múltipla sem citocinese, preparando-se para produzir numerosos merozoítos.
4. Gametócito: Algumas das células infectadas diferenciam-se em gametócitos, que são as formas sexuadas do parasita, prontas para serem ingeridas por um mosquito na próxima picada, continuando o ciclo de vida no vetor.

Essa fase é responsável pela sintomatologia da doença e representa a forma predominante do parasita no hospedeiro.

A fase sexuada, por outro lado, como destacado por Bardin (2011), ocorre no interior do mosquito, geralmente do gênero *Anopheles*. Esta fase é crucial para a transmissão da malária, uma vez que os parasitas sexuados se fundem no trato digestivo do mosquito, dando origem a esporozoítos que podem ser transmitidos a novos hospedeiros humanos durante a picada do mosquito. A interrupção eficaz do ciclo de vida do *Plasmodium* nessa fase pode ser uma estratégia importante para o controle da transmissão da malária.

A disseminação da doença está intrinsecamente ligada à capacidade do *Plasmodium* de completar seu ciclo de vida no mosquito e, assim, transmitir a infecção a novos hospedeiros humanos (Freitas *et al.*, 2016).

A magnitude da malária como um problema de saúde global é enfatizada pela Organização Mundial da Saúde (OMS). A OMS estima um número significativo de casos em todo o mundo. Isso ressalta a importância da pesquisa contínua sobre o ciclo de vida do *Plasmodium* e a classificação de suas fases, com o objetivo de desenvolver intervenções de saúde pública mais eficazes e reduzir a carga da malária em escala global (WHO, 2023).

Em resumo, o ciclo de vida do *Plasmodium* na malária envolve fases assexuadas e sexuadas, e a compreensão detalhada da classificação e biologia dessas fases, incluindo a esquizogônia hepática, a invasão dos glóbulos vermelhos, o desenvolvimento do trofozoíto e a esquizogônia sanguínea, é fundamental para o desenvolvimento de estratégias de prevenção, diagnóstico e controle da malária em todo o mundo.

3.2 Pré-processamento de dados

Em tarefas de classificação de dados desbalanceados, como a classificação de imagens em várias categorias, o pré-processamento desempenha um papel crucial na melhoria do desempenho dos modelos de aprendizado de máquina. Três técnicas utilizadas nesse trabalho incluem o redimensionamento de imagens, o balanceamento por subamostragem e a definição de pesos por classe. Cada uma dessas técnicas aborda desafios específicos no processamento de dados desbalanceados.

3.2.1 Redimensionamento de Imagens

O redimensionamento de imagens é uma técnica amplamente utilizada em problemas de processamento de imagens, como destacado em Gonzalez *et al* (2002). Essa técnica envolve a alteração do tamanho de uma imagem. Esse redimensionamento desempenha um papel crucial na padronização das dimensões das imagens de entrada em modelos de aprendizado de máquina, garantindo que todas as imagens tenham o mesmo tamanho e assim consigam ser treinadas em um modelo CNN.

3.2.2 Balanceamento por Subamostragem

A subamostragem da classe majoritária é uma técnica comum para lidar com conjuntos de dados desbalanceados (Gonzalez *et al*, 2002). Essa técnica envolve a remoção de exemplos da classe majoritária para equilibrar o conjunto de dados (Chollet *et al*, 2017).

Para sua execução é necessário inicialmente uma análise do conjunto de dados para determinar o grau de desequilíbrio entre as classes. Isso é geralmente feito calculando a proporção do número de amostras em cada classe, em seguida decide-se qual estratégia de subamostragem será usada. A mais simples, a subamostragem aleatória, é ilustrada na figura 1, onde amostras da classe majoritária são reduzidas através de remoção aleatória a um número de amostras próximo ao das outras classes.



Figura 1: Ilustração dos tipos de balanceamento de amostras

fonte: Pereira *et al* (2022)

3.2.3 Definição de Pesos por Classe

A definição de pesos por classe é uma outra estratégia de abordagem dos conjuntos de dados desbalanceados. Consiste em definir pesos diferentes para as classes, atribuindo maior ponderação à classe minoritária durante o treinamento do modelo (Aggarwal *et al*, 2018).

Essa abordagem permite que o modelo dê mais atenção aos exemplos da classe minoritária, reduzindo o viés em direção à classe majoritária e melhorando a precisão da classificação.

A combinação de redimensionamento de imagens, balanceamento por subamostragem e definição de pesos por classe pode contribuir para a obtenção de modelos de aprendizado de máquina mais robustos e eficazes, especialmente em cenários onde o desbalanceamento de classes é um desafio.

3.3 Redes Neurais Convolucionais

Redes Neurais Convolucionais (CNN) têm um amplo escopo de aplicações, sendo comumente empregadas no processamento de imagens, notadamente na classificação de imagens. A arquitetura de uma CNN é caracterizada por suas camadas tridimensionais. A terceira dimensão na camada de entrada representa o número de canais de cores ou planos do sinal de entrada, enquanto nas camadas subsequentes, essa dimensão reflete o número de filtros convolutivos utilizados na camada. Cada filtro convolutivo é responsável pela geração de um mapa de características por meio do processo de convolução dos valores de ativação da camada anterior (Aggarwal *et al*, 2018).

Diferentemente das redes neurais simples, nas quais os neurônios se conectam a todos os neurônios na camada anterior, em uma camada convolutiva, os neurônios se conectam apenas a um grupo restrito de neurônios. A arquitetura de uma rede convolutiva inclui várias camadas, cada uma com funções específicas. Dentre essas camadas, destacam-se a camada convolutiva, a camada de amostragem (pooling), a camada de dropout, a camada ReLU (*Rectified Linear Unit*) e a camada de regularização L2 (Aggarwal *et al*, 2018).

3.3.1 Camada Convolutiva (*Convolutional Layer*)

A camada convolutiva é essencial para a extração de características representativas das entradas. Ela consiste em vários *kernels* de convolução que calculam diferentes mapas de características. A geração do mapa de características pode ser vista na figura 2 e ocorre em duas etapas: primeiro, a entrada é convoluída com um filtro e, em seguida, uma função não linear é aplicada aos resultados da convolução. A não linearidade é de extrema importância para a detecção de características não lineares presentes nos dados de entrada (Goodfellow *et al*, 2016).

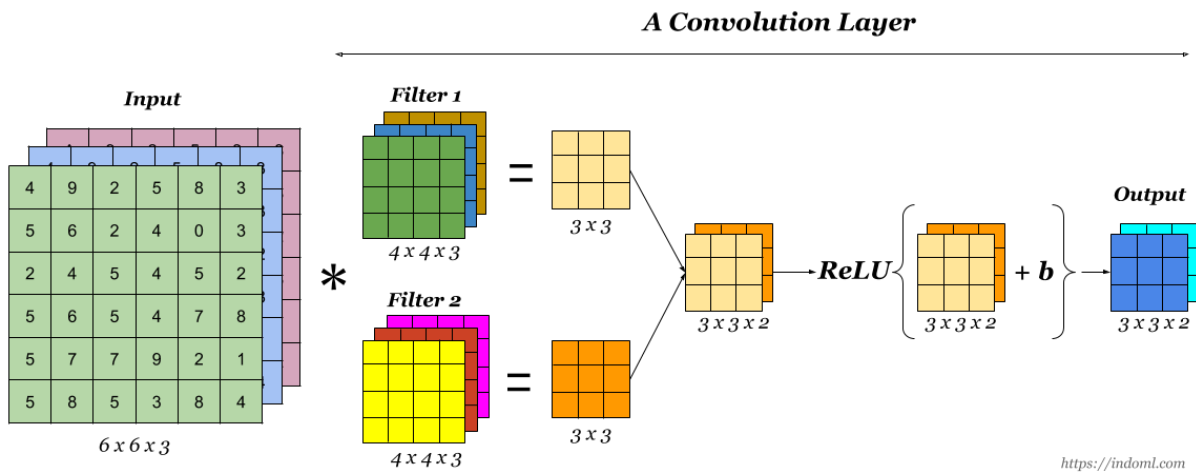


Figura 2: Cama de convolução

fonte: Beginners Guide to Convolutional Neural Networks, 2019

3.3.2 Camada de Subamostragem (*Pooling*)

A camada de *pooling* reduz a dimensionalidade das ativações lineares produzidas nas camadas anteriores. Ela opera em cada mapa de características de entrada, gerando um mapa de características reduzido como amostrado na figura 3. O *pooling* não afeta o número de mapas de características, mantendo a profundidade da imagem de saída igual à imagem de entrada na camada. Um exemplo comum de *pooling* é o max pooling, que reduz a saída em um fator pré-determinado (Serrão *et al*, 2022).

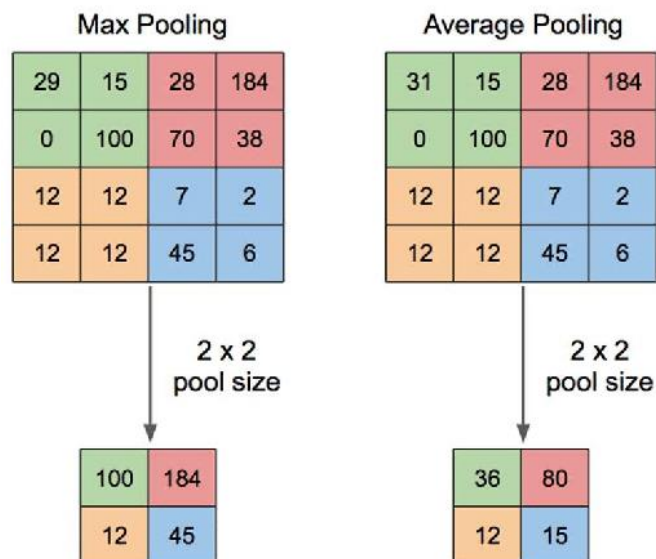


Figura 3: Ilustração das camadas de pooling

Fonte: Pooling layers in Neural nets and their variants, 2022

3.3.3 Camada *Dropout*

O *dropout* é uma técnica para melhorar a generalização da rede neural. Consiste na não utilização de uma proporção de neurônios em uma camada a cada época treinada, criando diferentes modelos de redes neurais. Se um conjunto de neurônios é eliminado, todas as conexões de entrada e saída relacionadas a esses neurônios também são removidas. Essa técnica ajuda a evitar o *overfitting*, no qual a rede se ajusta muito bem ao conjunto de treinamento, mas apresenta desempenho insatisfatório nos conjuntos de validação e teste (Aggarwal *et al*, 2018).

3.3.4 Camada ReLU (*Rectified Linear Unit*)

Nos últimos anos, a função de ativação ReLU substituiu funções de ativação como a sigmóide e a tangente hiperbólica devido à sua eficiência em termos de velocidade e precisão. A função ReLU é expressa como $f(x) = \max(0, x)$, onde x representa o valor de ativação de uma camada anterior (Aggarwal *et al*, 2018).

3.3.5 Regularização L2

A regularização L2 é uma técnica de generalização que adiciona um termo à função de custo da rede neural. Esse termo é a soma dos quadrados de todos os pesos da rede, escalonada por um fator $\lambda^2 n$, onde λ é o parâmetro de regularização e n é o tamanho do conjunto de treinamento. Essa técnica incentiva a rede a aprender pesos menores, contribuindo para a minimização da função de custo (Serrão *et al*, 2022).

Essas camadas e técnicas desempenham papéis fundamentais no funcionamento das redes neurais convolucionais, permitindo o aprendizado de características relevantes em imagens e contribuindo para a capacidade de generalização do modelo.

3.4 Otimizadores

Métodos de otimização desempenham um papel crucial no treinamento de redes neurais, sejam elas redes neurais simples ou convolucionais. Essas técnicas têm o propósito de minimizar a função de custo, o que, por sua vez, permite que as redes aprendam a realizar tarefas de classificação e regressão. Neste estudo, foram empregados diversos métodos de otimização, incluindo Estimativa Adaptativa de Momentos (ADAM), Propagação da Raiz Média Quadrática (RMSProp) e Gradiente Descendente Estocástico com Momento (SGDM).

3.4.1 Estimativa de Dinâmica Adaptativa (ADAM)

O método Adam opera no gradiente, empregando uma média ponderada exponencial dos gradientes anteriores (V_t). Além disso, ele ajusta a taxa de aprendizado (α) dividindo-a pela raiz da média móvel (S_t) dos quadrados dos gradientes. A atualização dos pesos é realizada considerando essas médias e uma pequena constante ϵ para evitar divisão por zero. Adam também incorpora parâmetros de correção de viés (β_1 e β_2) para as médias móveis (Kingma *et al*, 2015).

3.4.2 Propagação da Raiz Média Quadrática (RMSProp)

O RMSProp é um método de otimização que adapta a taxa de aprendizado durante o treinamento. Ele utiliza uma média ponderada exponencial dos quadrados dos gradientes anteriores para ajustar a taxa de aprendizado. A atualização dos pesos considera a taxa de aprendizado adaptativa, a média dos quadrados dos gradientes (\hat{S}_t) e uma constante ϵ . O valor de β_t é responsável por ajustar a taxa de aprendizado (Ruder *et al*, 2016).

3.4.3 Gradiente Descendente Estocástico com Momento (SGDM)

O Gradiente Descendente Estocástico (SGD) é amplamente utilizado, mas pode ser lento no treinamento. Para acelerar a convergência, o SGDM introduz a noção de momento, que controla a velocidade das atualizações dos parâmetros da rede. As atualizações dos pesos levam em consideração o gradiente, a taxa de aprendizado e o momento, que suaviza as mudanças nos parâmetros. Isso resulta em uma convergência mais rápida, especialmente em regiões de gradientes constantes. A taxa de aprendizado (α) e o parâmetro de suavização (β) desempenham papéis cruciais nesse método (Goodfellow *et al*, 2016).

Esses métodos de otimização desempenham um papel fundamental no treinamento eficaz de redes neurais, permitindo que os modelos se adaptem aos dados de maneira eficiente e alcancem resultados de alta precisão em tarefas de classificação e regressão. Cada método tem suas próprias características que se adequam a diferentes cenários de treinamento e tipos de redes neurais.

4 Materiais e Métodos

Na seção anterior foram apresentados os conceitos necessários ao desenvolvimento deste projeto. Nesta seção apresenta-se os materiais e os métodos empregados, desde o pré-processamento até a realização de todos os testes das redes e as medidas de desempenho.

A tarefa de classificação dos ciclos de vida do plasmodium, proposta neste trabalho, é realizada em duas etapas, conforme a Figura 4. Preliminarmente à etapa de classificação, tem-se o pré-processamento da base de dados (balanceamento de classe e redimensionamento do tamanho das imagens de células usadas).

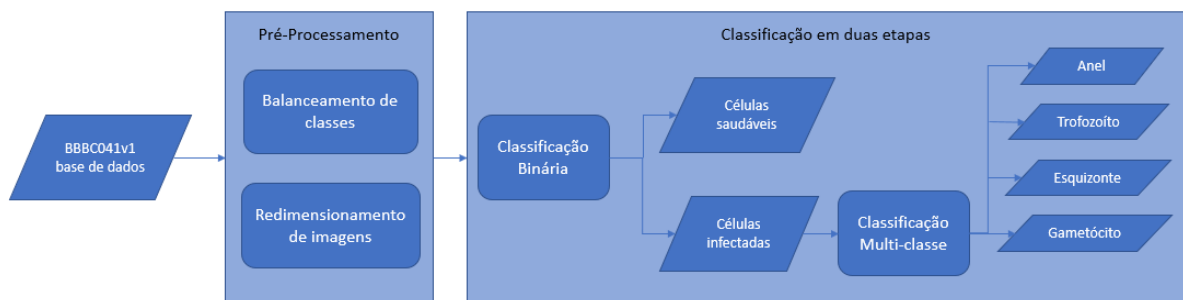


Figura 4 - Metodologia para classificação do Plasmodium

Na tarefa de classificação, serão avaliadas de diversas arquiteturas e tamanhos de redes neurais profundas, em especial as arquiteturas EfficientNet, rede com 7 modelos diferentes sendo a maior B7 a mais complexa e a B0 a menos complexa.

4.1 Base de Dados

Neste trabalho foi utilizado o banco de dados BBBC041v1 (Hung *et al*, 2017). Um banco de imagens de campos microscópicos de esfregaço de gota de sangue de pacientes com suspeita infecção por *plasmodium vivax*. Esta base de dados contém 1364 imagens de campos microscópicos de lâminas com anotações fornecidas pela Fundação de Medicina Tropical Doutor Heitor Vieira Dourado em Manaus, vide Figura 5. Dessas 1364 imagens são extraídas cerca de 80.000 sub imagens de células (*patches*). Essas imagens de células foram individualmente rotuladas em 7 rótulos, sendo 2 para células saudáveis (RBCs e leucócitos), 4 para as células infectadas com *plasmodium vivax* (anel, gametócito, esquizonte, trofozoíto) e 1 para as células difíceis de rotular.

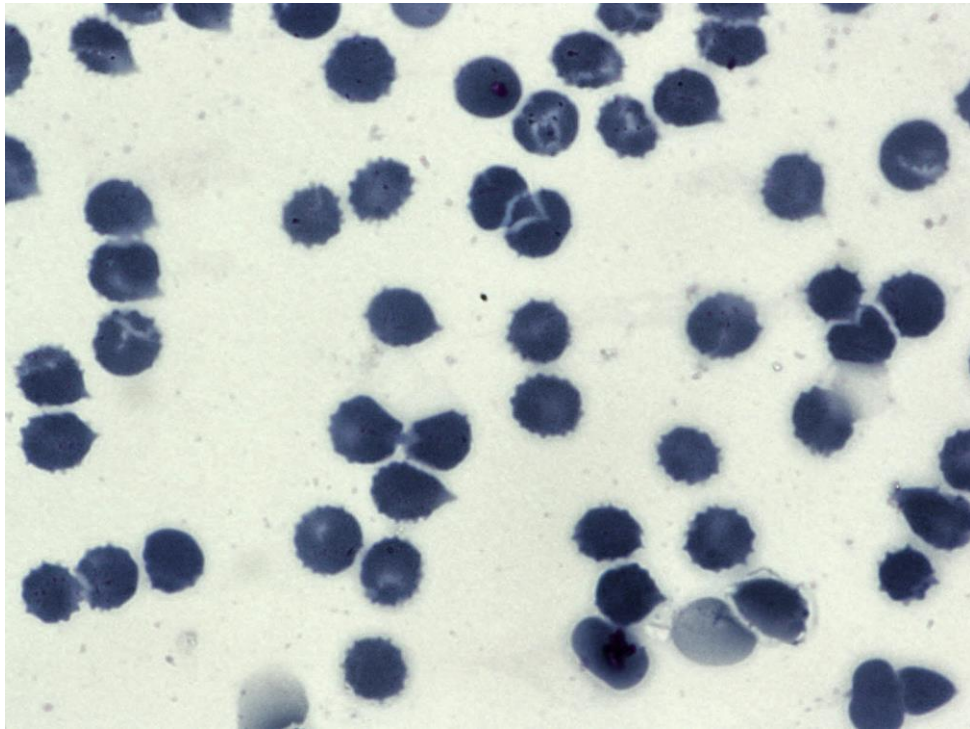


Figura 5 - Imagem de microscopia da base de dados

Fonte: (Ljosa *et al*, 2012)

A base de dados utilizada tem como sua principal característica o desbalanceamento entre as classes (vide figura 6 e Tabela 2). Isso se deve ao fato da proporção de células saudáveis ser maior que a de contaminadas em cada esfregaço. Com isso, cerca de 95% das células são da classe RBC não infectadas. Adicionalmente, há também um desbalanceamento entre as classes de células infectadas.

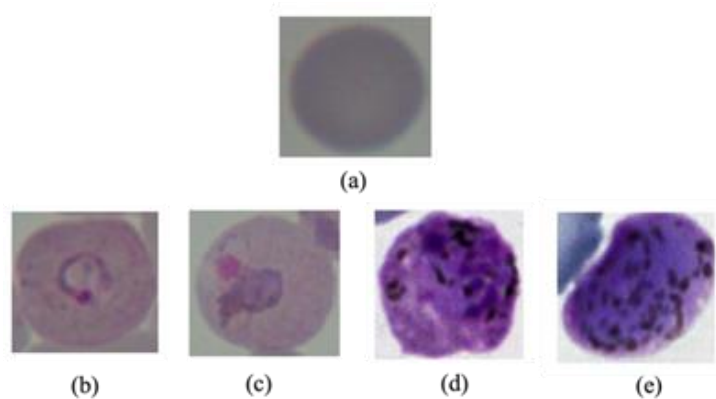


Figura 6 - Exemplo de células: (a) célula não infectada e (b-e) células infectadas mostrando o ciclo de vida do plasmodium (b) anel, (c) trofozoíto, (d) esquizonte e (e) gametócito

fonte: (Ljosa *et al*, 2012)

Tabela 2. Quantidade de células RBC por classe, extraídas de (Ljosa *et al*, 2012).

Classe da célula/ Estágio do ciclo de vida do <i>Plasmodium</i>	Quantidade
Não infectada	83,034
Infectada/Anel	522
Infectada/Trofozoíto	1,584
Infectada/Esquizonte	190
Infectada/Gametócito	156
Total	85,486

4.2. Pré-processamento

Para o pré-processamento foi inicialmente necessário fazer um redimensionamento das imagens, tendo em vista de que redes CNN precisam que as entradas da rede tenham formato padrão, então com base na literatura e nos tamanhos dos *patches* da base de dados foi utilizado um tamanho de 150x150 pixels.

Em seguida foi realizado o balanceamento das classes. Para a primeira classificação (binária) o balanceamento foi feito por subamostragem da classe majoritária. Assim, a classe de células não infectadas foi reduzida para 2500. Tal técnica foi suficiente para essa etapa de classificação.

Na segunda etapa de classificação, classificação multiclasse, optou-se por utilizar a ponderação da taxa de aprendizado das classes (Aggarwal, 2018), conforme definido a seguir:

$$\text{Ponderação do aprendizado por classe} = \frac{n_{\text{Imagens}_{total}}}{Q_{\text{Classes}} \cdot n_{\text{Imagens}}} \quad (1)$$

em que:

$n_{\text{Imagens}_{total}}$: número total de imagens.

Q_{Classes} : número de classes (neste caso as 4 fases do plasmodium).

n_{Imagens} : número de imagens presentes na classe que deseja calcular o peso.

Os seguintes pesos por classe foram obtidos: peso da classe anel = 1,19; peso da classe trofozoíto = 0,39; peso da classe esquizonte = 3,18 e peso da classe gametócito = 3,91.

Quanto maior o peso maior será a taxa de aprendizado daquela classe durante o treinamento.

4.3 Conjuntos de treino, validação e teste

Para esse trabalho foi necessário a divisão dos dados em 3 conjuntos: treino, validação e teste. O conjunto de treinamento é utilizado durante o treinamento da rede com o objetivo de extrair características das imagens. Já o conjunto de validação é usado com o objetivo de definir hiper parâmetros do modelo, assim garantindo que o aprendizado com o conjunto de treino seja eficiente. Por fim, o conjunto de teste é usado para avaliar o modelo treinado, obtendo-se as métricas de desempenho do modelo.

Para a tarefa de classificação binária, os 3 conjuntos foram divididos igualmente, conforme mostrado na Tabela 3

Tabela 3: Divisão da base de dados para a tarefa de classificação binária

classe	Número de imagens		
	Conjunto de treinamento	Conjunto de validação	Conjunto de teste
célula não infectada	833	833	834
célula infectada	817	817	818

Para classificação multi-classe foram utilizadas as imagens das classes infectadas. O conjunto de teste será composto pelas imagens de células infectadas que forem classificadas corretamente na classificação binária. O conjunto de treinamento e validação, é apresentado na Tabela 4 na proporção de 70% para treino e 30% para validação.

Tabela 4: Divisão da base de dados para a tarefa de classificação binária

classe	Número de imagens	
	Conjunto de treinamento	Conjunto de validação
Anel	244	104
Trofozoíto	740	316
Esquizonte	89	37
Gametócito	73	31

4.4 Modelos Implementados

Na pesquisa, foram utilizadas diversas arquiteturas de redes neurais convolutivas para a tarefa de classificação do *plasmodium*. Estas arquiteturas incluem as redes EfficientNet (B0 a B7), DenseNet-121 e uma rede CNN proposta. Além da análise comparativa de desempenho dos modelos implementados, buscou-se também comparar o desempenho das redes pré-treinadas com uma rede treinada somente com a base de dados utilizada neste trabalho.

4.4.1 EfficientNet

A EfficientNet é uma família de arquiteturas de redes neurais convolutivas que abrange sete variantes, da B0 à B7. Cada variante é caracterizada por uma combinação específica de tamanho, profundidade e largura da rede, otimizada para diferentes níveis de eficiência computacional e desempenho. A rede B0 é a variante mais leve, com menos parâmetros, enquanto a B7 é a mais profunda e complexa, com um número substancialmente maior de parâmetros como é possível ver na figura 7 onde é visto a diferença entre a rede mais simples e a mais complexa.

As redes EfficientNet são projetadas para escalabilidade, o que significa que podem ser ajustadas para atender a requisitos de tamanho e desempenho específicos. Isso é alcançado por meio do uso de três parâmetros principais: *width scale* (escala de largura), *depth scale* (escala de profundidade) e *resolution scale* (escala de resolução) (Tan *et al.*, 2019).

As redes EfficientNet têm se destacado em tarefas de classificação de imagens devido ao seu equilíbrio entre eficiência computacional e precisão. Portanto, empregou-se as sete variantes da EfficientNet (B0 a B7), cada uma oferecendo um nível diferente de desempenho e eficiência. Isso permitiu a comparação entre essas arquiteturas.

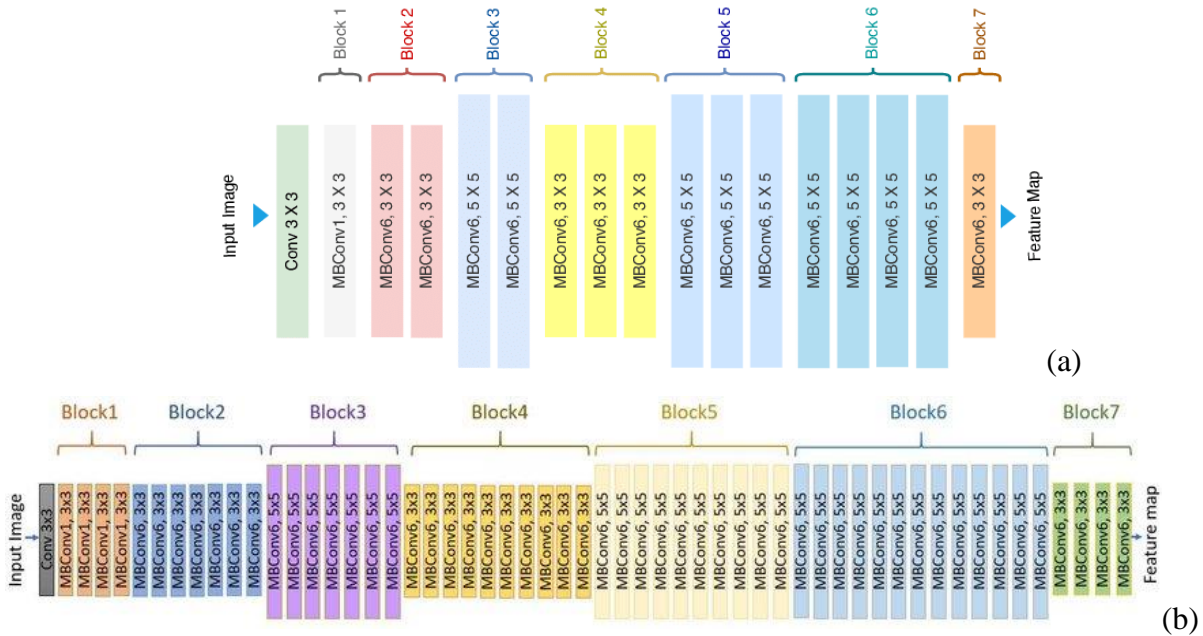


Figura 7 - Comparação da arquitetura das redes (a) efficientNet B0 e (b) efficientNet B7

Fontes: Ahmed *et al* (2020) e Baheti *et al* (2020)

4.4.2 DenseNet-121

A arquitetura DenseNet-121 é uma contribuição significativa para o campo de visão computacional, introduzida por Huang *et al* (2017). Essa arquitetura é uma evolução do conceito de redes neurais convolutivas (CNNs) que enfatiza a densa conectividade entre suas camadas, arquitetura essa que pode ser vista na figura 8.

Diferentemente das CNNs convencionais, onde as camadas são conectadas de forma sequencial, a DenseNet-121 estabelece conexões densas entre camadas subsequentes. Cada camada recebe entradas de todas as camadas anteriores, permitindo um fluxo eficiente de informações por toda a rede. Isso resulta em um modelo altamente paramétrico e poderoso.

Blocos de Densidade: A arquitetura é composta por quatro blocos densos, cada um contendo várias camadas convolutivas. Cada bloco incorpora uma camada de normalização de lote e funções de ativação ReLU. Essa estrutura densa facilita a aprendizagem de características complexas a partir dos dados de entrada, extraindo informações de várias escalas espaciais, aproveitando os recursos de diferentes camadas para melhorar a capacidade de classificação.

Apesar de sua conectividade densa, a arquitetura utiliza menos parâmetros do que outras redes com desempenho similar. Isso é possível graças ao compartilhamento de parâmetros entre camadas e ao uso eficiente de atalhos residuais, assim apresentando um desempenho notável em tarefas de classificação de imagens e sendo amplamente reconhecida

por sua capacidade de capturar informações detalhadas e complexas das imagens, se tornando uma escolha popular em aplicações de visão computacional devido à sua eficácia na extração de características e ao seu poder de generalização.

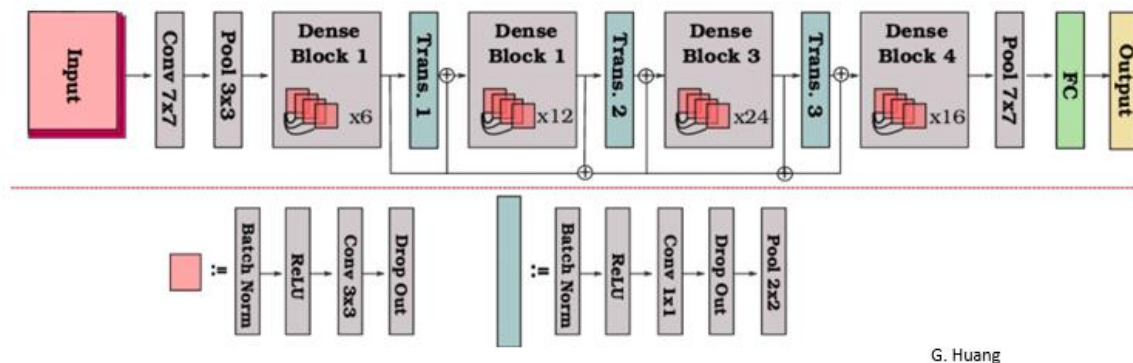


Figura 8 - Arquitetura da rede DenseNet121

Fonte: Huang (2018).

4.4.3 CNN Proposta

Também foi desenvolvida uma rede convolutiva personalizada para fazer benchmark com as pré-treinadas escolhidas, onde essa pode ser vista na figura 9 e consiste em quatro blocos de extração de recursos. Cada bloco incorpora uma camada de convolução, normalização de lote e max-pooling, e função de ativação ReLU em cada camada. E para a parte totalmente conectada tem 2 camadas densas intercaladas com camadas de dropout que contribuem para a análise e generalização dos recursos extraídos pela rede.

Feita essa arquitetura será possível fazer o *benchmark* que foi proposto para esse trabalho, onde esta rede proposta com características similares as pré-treinadas escolhidas foi montada para a comparação.

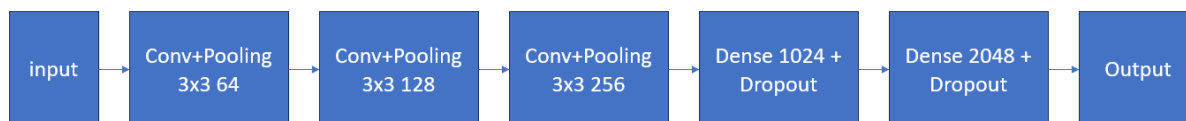


Figura 9 - Arquitetura da rede CNN proposta

4.5 Métricas de Avaliação

Os desempenhos dos modelos implementados foram avaliados através das seguintes métricas: acurácia, precisão, sensibilidade (ou *Recall*), especificidade e *F1-Score*. Essas métricas foram calculadas com base em conceitos fundamentais, a saber: Verdadeiro Positivo (VP), que ocorre quando o modelo classificar corretamente um dado de entrada como positivo; Verdadeiro Negativo (VN), que ocorre quando o modelo classifica corretamente um dado de entrada como negativo; Falso Positivo (FP), que ocorre quando o modelo erroneamente classificar um dado de entrada como positivo; Falso Negativo (FN), que ocorre quando o modelo erroneamente classificar um dado de entrada como negativo. Neste estudo, a classificação positiva representa quando a rede é capaz de classificar o *plasmodium* ou célula saudável conforme o seu rótulo presente na base de dados, enquanto a classificação negativa ocorre quando a rede tentou classificar o ciclo de vida de um *plasmodium* e classificou de forma diferente daquela anotada na base de dados. A seguir apresenta-se as expressões de cada uma das métricas referidas:

- Acurácia: avalia a proporção de previsões corretas feitas pelo modelo e é calculada da seguinte forma:

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

- Precisão: mede a proporção de previsões positivas que eram de fato positivas e é calculada da seguinte forma:

$$precisão = \frac{VP}{VP + FP} \quad (3)$$

- Sensibilidade ou *recall*: avalia a capacidade do modelo de identificar corretamente os casos positivos e é calculada como segue:

$$sensibilidade = \frac{VP}{VP + FN} \quad (4)$$

- Especificidade: mede a capacidade do modelo de identificar corretamente os casos negativos e é calculada da seguinte forma:

$$especificidade = \frac{VN}{VN + FP} \quad (5)$$

- F1-score: é um indicador do equilíbrio entre precisão e sensibilidade, fornecendo uma medida geral da qualidade do modelo e é calculado da seguinte maneira:

$$F1-Score = 2 \cdot \frac{\textit{precisão} \cdot \textit{sensibilidade}}{\textit{precisão} + \textit{sensibilidade}} \quad (6)$$

4.6 Ambiente de Desenvolvimento

Para as simulações foi utilizado um sistema com Python 3.8.16 e um hardware com processador Intel Core Xeon W-1270 CPU equipado com 64 GB de RAM, placa gráfica Nvidia Quadro 5000 GPU, rodando no sistema operacional Windows 10.

Utilizou-se as seguintes bibliotecas para auxiliar no desenvolvimento: tensorflow na versão 2.2.1, Keras na versão 2.4.3. A biblioteca pandas na versão 1.0.3 foi utilizada para trabalharmos com as tabelas dos conjuntos de dados, a biblioteca matplotlib na versão 3.2.2 foi utilizada para visualização.

5 Resultados e Discussões

Esta etapa tem como objetivo apresentar os resultados obtidos por todos os modelos treinados durante o desenvolvimento do trabalho, analisá-los e, por fim compará-los com resultados de trabalhos similares presentes na literatura e analisados na seção de revisão bibliográfica deste trabalho.

Ressalta-se que a classificação binária foi implementada apenas com o modelo EfficientNet B7, enquanto para a classificação multiclasse foram implementados os modelos EfficientNet B0 a B7, DenseNet121 e a CNN proposta

5.1 Resultados classificação binária

Para a etapa de classificação binária, a Tabela 5 compara os resultados obtidos com a rede EfficientNet B7, utilizando os otimizadores escolhidos. As seguintes métricas são calculadas: acurácia, precisão, especificidade, sensibilidade e F1-Score. Os melhores resultados foram obtidos com o otimizados RMSProp. A Figura 12 mostra a matriz de confusão correspondente a esse modelo de melhor desempenho.

Tabela 5: Resultados Obtidos para classificação binária para o conjunto de teste (células saudáveis ou células infectadas com **plasmodium**)

Classificação binária EfficientNet B7

Otimizador	Acurácia	Precisão	Especificidade	Sensibilidade	F1-Score
RMSProp	99.27%	99.16%	99.39%	99.40%	99.28%
Adam	98.53%	98.78%	98.85%	98.45%	98.61%
SGDM	95.34%	96.51%	95.11%	95.78%	96.14%

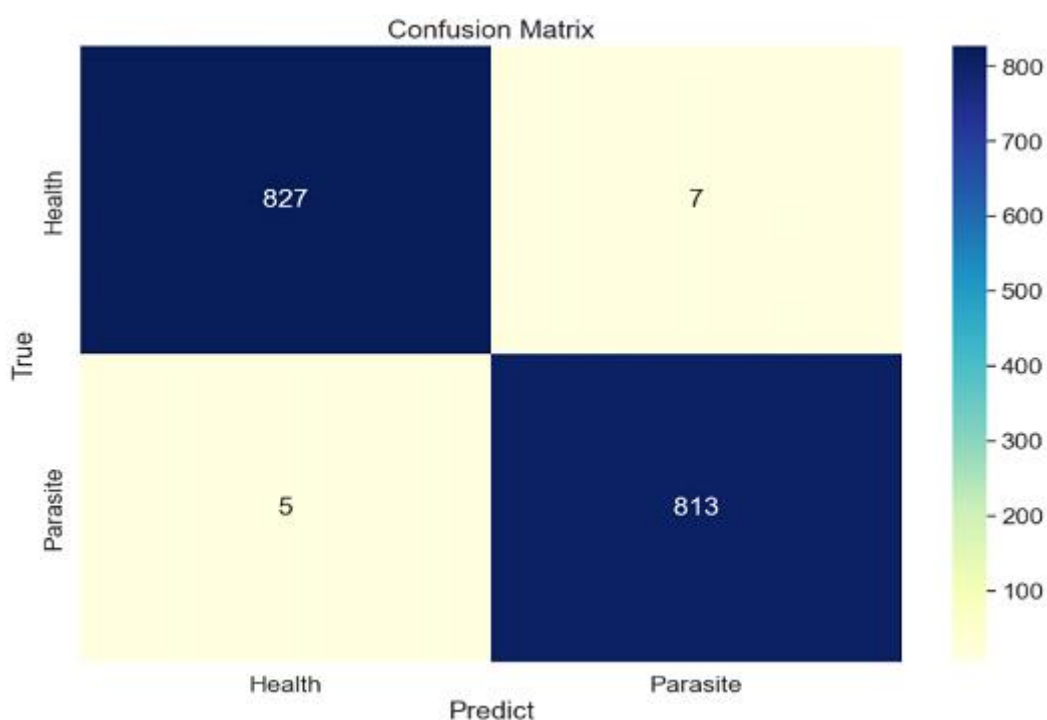


Figura 12: Matriz de confusão da classificação binária com o modelo de melhor desempenho (arquitetura EfficientNet B7 com o otimizador RMSProp)

5.2 Resultados classificação multiclasse

A Tabela 6 mostra as métricas de acurácia, sensibilidade, especificidade e *F1-Score* obtida para classificação multiclasse, com EfficientNet B0 a B7, DenseNet121 e CNN proposta, com os três otimizadores: Adam, RMSProp e SGDM.

Conforme mostrado, os melhores resultados foram obtidos com o EfficientNet B6 utilizando o otimizador RMSProp. A Figura 13 mostra a matriz de confusão correspondente.

Após os treinos e os resultados anotados na Tabela 6 é possível observar algumas características: uma delas é que a rede mais complexa não conseguiu o melhor resultado, sendo este obtido pela EfficientNet B6 com o otimizador RMSProp uma rede mais simples do que a B7. As demais redes apresentaram um resultado próximo a essa de maior complexidade.

Tabela 6: Comparação do desempenho da classificação multiclasse entre as redes EfficientNet, DenseNet e CNN proposta, com os três otimizadores para o conjunto de teste.

Model	Acurácia	Sensibilidade	Especificidade	F1-Score
EfficientNetB0 ADAM	60,39%	24,64%	75,25%	20,45%
EfficientNetB0 RMSProp	64,70%	25,00%	75,00%	19,64%
EfficientNetB0 SGDM	54,24%	29,60%	75,01%	26,74%
EfficientNetB1 ADAM	86,47%	80,70%	94,07%	79,09%
EfficientNetB1 RMSProp	64,70%	25,00%	75,00%	19,64%
EfficientNetB1 SGDM	39,73%	43,95%	80,52%	34,47%
EfficientNetB2 ADAM	83,76%	80,37%	93,28%	76,43%
EfficientNetB2 RMSProp	86,10%	80,94%	93,90%	78,69%
EfficientNetB2 SGDM	64,70%	25,00%	75,00%	19,64%
EfficientNetB3 ADAM	88,44%	83,12%	94,37%	82,48%
EfficientNetB3 RMSProp	87,58%	84,16%	94,66%	81,89%
EfficientNetB3 SGDM	87,70%	85,47%	94,84%	82,40%
EfficientNetB4 ADAM	82,41%	75,62%	92,10%	73,57%
EfficientNetB4 RMSProp	16,36%	34,55%	77,08%	16,84%
EfficientNetB4 SGDM	46,74%	54,59%	83,07%	40,90%
EfficientNetB5 ADAM	87,58%	78,68%	93,58%	79,95%
EfficientNetB5 RMSProp	87,58%	83,95%	94,53%	81,38%
EfficientNetB5 SGDM	64,58%	24,95%	74,97%	19,62%
EfficientNetB6 ADAM	89,42%	83,88%	95,10%	83,06%
EfficientNetB6 RMSProp	89,91%	83,74%	95,18%	83,88%
EfficientNetB6 SGDM	56,46%	64,53%	86,25%	51,13%
EfficientNetB7 ADAM	89,05%	81,69%	94,36%	82,85%

Model	Acurácia	Sensibilidade	Especificidade	F1-Score
EfficientNetB7 RMSProp	86,96%	79,93%	93,58%	80,30%
EfficientNetB7 SGDM	48,83%	33,43%	79,34%	30,26%
DenseNet121 ADAM	84,75%	74,46%	92,03%	85,09%
DenseNet121 RMSProp	86,10%	93,07%	80,75%	82,21%
DenseNet121 SGDM	30,26%	26,73%	62,46%	37,60%
CNN Proposta ADAM	82,53%	67,64%	91,26%	75,73%
CNN Proposta RMSProp	83,27%	63,23%	92,39%	73,44%
CNN Proposta SGDM	64,57%	30,62%	84,51%	84,25%

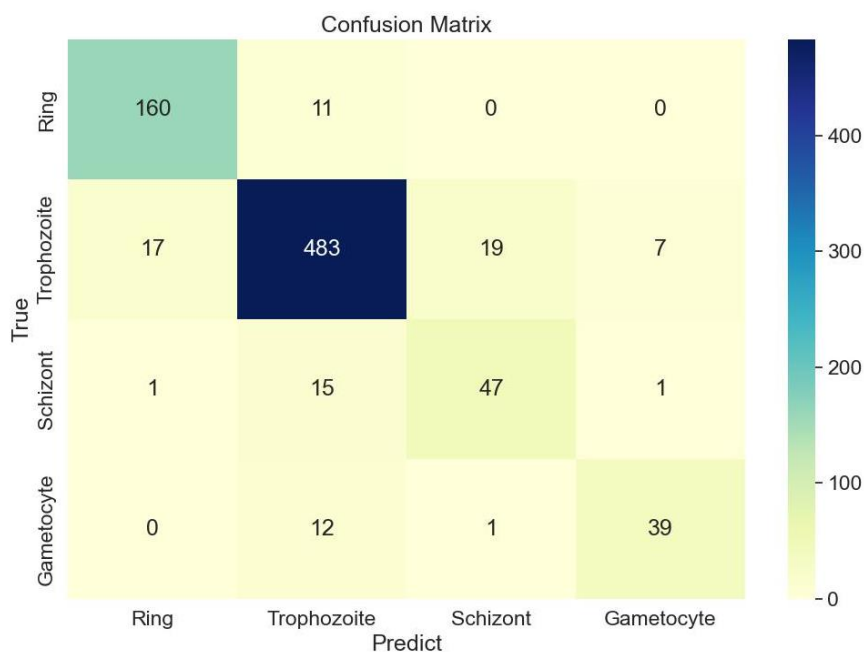


Figura 13: Matriz de confusão da classificação multiclasse do modelo de melhor desempenho (arquitetura EfficientNet B6 com o otimizador RMSProp)

A seguir, na Figura 14, apresenta-se o gráfico de acurácia e perda referente a treino desta rede, nele podemos ver que apesar de ter o melhor resultado ela apresentou *overfitting* e isso foi presente em todas as redes treinadas mas essa sendo a que apresentou menos essa característica indicando que a dimensão da rede se adequa melhor a base de dados assim gerando o melhor resultado, além dele também é apresentado a matriz de confusão para a

parcela de teste da rede, mostrando as individualidades da rede para cada fase.

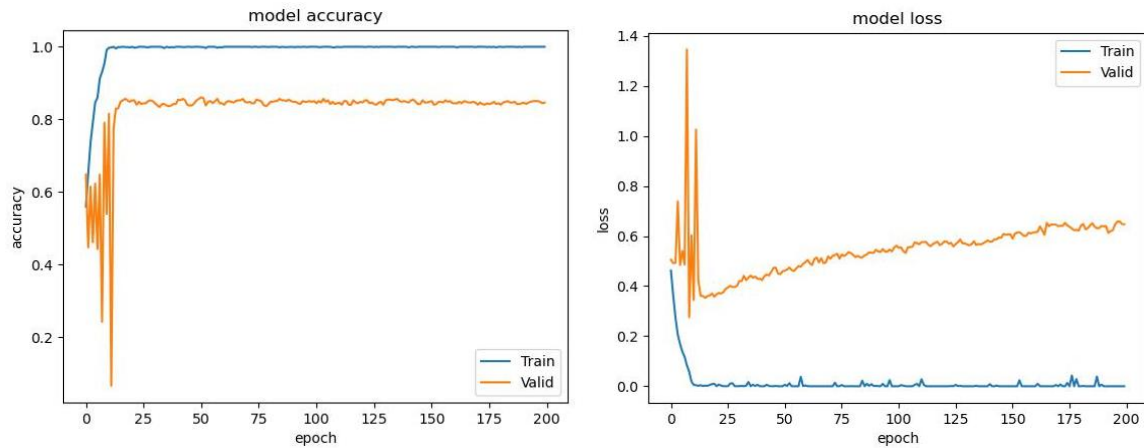


Figura 14: Gráficos de acurácia e perda para o treinamento multiclasse (arquitetura EfficientNet B6 com o otimizador RMSProp)

5.3 Discussões

Com respeito a classificação multiclasse, vemos que o uso de todas as arquiteturas da rede EfficientNet foram relevantes para o êxito dessa etapa, pois mostrou que a maior rede, não necessariamente é a que apresenta o melhor desempenho. Este também é afetado pelo conjunto de dados disponíveis. Uma análise conjunta desses dois fatores, impacta no tempo de treino e nos requisitos de hardware e pode até permitir mais ajustes de hiper parâmetros na rede.

Como pode ser observado na Tabela 6, as redes DenseNet121 e a arquitetura proposta, não conseguiram competir com as melhores arquiteturas EfficientNet.

Com relação a CNN proposta vemos que, mesmo tendo uma arquitetura similar a EfficientNet, os resultados não foram o esperado, mostrando que o uso de uma rede pré-treinada tem grande influência em relação a uma rede que foi treinada apenas com o *dataset* utilizado que não possui uma quantidade suficiente de imagens).

5.4 Comparação com a literatura

Com o fim de mostrar os avanços obtidos neste trabalho são apresentadas comparações com trabalhos publicados. Na Tabela 7 apresenta-se a comparação da primeira etapa deste

trabalho, a classificação binária.

Tabela 7: Comparação do desempenho da classificação binária com a rede EfficientNet B7 e otimizador RMSProp com outros trabalhos da literatura

Tabela de Comparação Binária					
Estudo	Acurácia	Precisão	Especificidade	Sensibilidade	F1-Score
Hung J <i>et al</i>	98%	98,83%	99,41%	98,43%	99,11%
A Rahman <i>et al</i>	99,35%	96,85%	99,76%	92,31%	92,31%
EfficientNet B7					
RMSProp	99,27%	99,16%	99,39%	99,40%	99,28%

Na Tabela 7 é possível observar que o resultado obtido com a rede EfficientNet B7 com otimizador RMSProp foi similar ao estado da arte presente na literatura. Apesar de não obter a melhor acurácia é possível ver que o F1-Score, outra métrica de importância, conseguiu superar o estado na arte. Então atingimos o objetivo para esta camada de classificação.

A comparação para a classificação multiclasse, é apresentada na Tabela 8. Neste caso a nossa melhor rede, a rede EfficientNet B6, superou o estado da arte.

Tabela 8: Comparação do desempenho da classificação multiclasse com a rede EfficientNet B6 e otimizador RMSProp com outros trabalhos da literatura

Estudo	Acurácia	F1-score
Arshad <i>et al</i>	79,61%	82,04%
Manku <i>et al</i>	82%	82%
X. -Z. Chen <i>et al</i>	86.6%	83,3%
EfficientNet B6		
RMSProp	89,91%	83,88%

6 Conclusão

O principal objetivo deste trabalho foi avaliar o desempenho de nove arquiteturas pré-treinadas, EfficientNet B0 a B7 e DenseNet 121, com o desempenho de uma arquitetura CNN treinada do zero, na classificação de células RBC saudáveis e infectadas com o parasita da malária em imagens microscópicas de esfregaço de gota de sangue espesso. Foram avaliados três otimizadores: ADAM, RMSProp e SGDM. A classificação foi realizada em duas etapas. Na primeira etapa foi obtida uma classificação binária, em células infectadas e não infectadas. Na

segunda etapa, o estágio do ciclo de vida do plasmódio foi identificado nas células infectadas. O desequilíbrio entre as classes foi resolvido tanto pela redução de amostras da classe com maior número de amostras, quanto por uma estratégia de equilíbrio de peso durante o treinamento da rede neural. Nas duas etapas de classificação, os melhores resultados foram obtidos com o otimizador RMSProp e com a rede EfficientNet B7 para classificação binária e EfficientNet B6 para uma classificação multiclasse. Os resultados obtidos demonstraram que, no problema de classificação multiclasse, o EfficientNet B6 superou resultados previamente obtidos na literatura, com precisão de 89,91% e escore F1 de 83,88%. No problema de classificação binária, o EfficientNet B7 obteve resultados comparáveis aos resultados obtidos em estudos previamente publicados na literatura.

Referências

- Abbas, S. S., & Dijkstra, T. M. H. (2020). Detection and stage classification of Plasmodium falciparum from images of Giemsa stained thin blood films using random forest classifiers. *Diagn Pathol*, 15(130), 1-11. doi: 10.1186/s13000-020-01012-7.
- Aggarwal, C. C. (2018). *Neural Networks and Deep Learning*. New York: Springer.
- AGGARWAL, C. C. (2018). *Neural Networks and Deep Learning*. New York: Springer.
- Ahmed, T., & Sabab, N. (2020). Classification and understanding of cloud structures via satellite images with EfficientUNet. doi: 10.1002/essoar.10507423.1.
- Arshad, Q., Ali, M., Hassan, S.-U., Chen, C., Imran, A., Rasul, G., & Sultani, W. (2022). A dataset and benchmark for malaria life-cycle classification in thin blood smear images. *Neural Computing and Applications*, 34. doi: 10.1007/s00521-021-06602-6.
- Baheti, B., Innani, S., Gajre, S., & Talbar, S. (2020). Eff-UNet: A Novel Architecture for Semantic Segmentation in Unstructured Environment. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, 1473-1481. doi: 10.1109/CVPRW50498.2020.00187.
- Bardin, L. (2011). *Análise de conteúdo*. São Paulo: Edições 70.
- BBBC041v1. (2012). Broad Bioimage Benchmark Collection [Ljosa et al., *Nature Methods*].
- Chen, X.-Z., Hsu, Y.-C., Chen, Y.-L., Goh, C. M., & Yu, W.-J. (2022). Deep learning based malaria-infected cell detection and parasite life stage classification method. *2022 IEEE 11th Global Conference on Consumer Electronics (GCCE)*, 420-421. doi: 10.1109/GCCE56475.2022.10014315.
- Chollet, F. (2017). *Deep learning with Python*. New York, NY: Manning Publications Co.
- Freitas, I. M., Moreira, M. A., & Silva, M. V. (2016). Alfabetização científica: uma revisão de literatura. *Ensaio Pesquisa Educação Ciência*, 20(1), 123-152.
- Gonzalez, R. C., & Woods, R. E. (2002). *Digital image processing (2nd ed.)*. Upper Saddle River, NJ: Prentice Hall.

GOODFELLOW, I., BENGIO, Y., & COURVILLE, A. (2016). Deep Learning. The MIT Press.

Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. (2017). Densely Connected Convolutional Networks. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2261-2269. doi: 10.1109/CVPR.2017.243.

Hung, J., Lopes, S. C. P., Nery, O. A., Nosten, F., Ferreira, M. U., Duraisingh, M. T., Marti, M., Ravel, D., Rangel, G., Malleret, B. B., Lacerda, M. V. G., Rénia, L., Costa, F. T. M., & Carpenter, A. E. (2017). Applying Faster RCNN for Object Detection on Malaria Images. Conf Comput Vis Pattern Recognit Workshops. doi: 10.1109/cvprw.2017.112.

KINGMA, D. P., & BA, J. (2015). A Method for Stochastic Optimization. 3rd International Conference for Learning Representations (ICLR), San Diego.

Manku, R., Sharma, A., & Panchbhai, A. (2020). Malaria Detection and Classification.

MEDIUM. (2019). Beginners Guide to Convolutional Neural Networks. Acesso em: <https://towardsdatascience.com/beginners-guide-to-understanding-convolutional-neural-networks-ae9ed58bb17d>. Acesso em: 04 de nov. de 2023.

MEDIUM. (2022). Pooling layers in Neural nets and their variants. Acesso em: <https://medium.com/aiguys/pooling-layers-in-neural-nets-and-their-variants-f6129fc4628b>. Acesso em: 04 de nov. de 2023.

Pereira, G. R.ROCHA, A. C., & Almeida, L. H. (2020). Classificação dos estados de evolução microestrutural de tubos de reforma a vapor de diferentes composições químicas utilizando inteligência artificial uma abordagem de END 4.0.

Rahman, A., Zunair, H., Reme, T. R., Rahman, M. S., & Mahdy, M. R. C. (2021). A comparative analysis of deep learning architectures on high variation malaria parasite classification dataset. Tissue Cell, 69, 101473. doi: 10.1016/j.tice.2020.101473.

RUDER, S. (2016). An overview of gradient descent optimization algorithms. arXiv:1609.04747.

SERRÃO, M. K. M. (2020). Detecção automática de bacilos em baciloscopia de campo claro usando aprendizado profundo e técnica de imagem mosaico. Dissertação (Mestrado em

Engenharia Elétrica) - Universidade Federal do Amazonas.

Tan, M., & Le, Q. V. (2019). EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. ICML 2019, 10691-10700. Acesso em: <http://proceedings.mlr.press/v97/tan19a.html>.

World Health Organization. (2021). World malaria report 2021. Geneva: World Health Organization; Licence: CC BY-NC-SA 3.0 IGO.

World Health Organization. (2022). World malaria report 2022. Geneva: World Health Organization; Licence: CC BY-NC-SA 3.0 IGO.

World Health Organization. (2023). World malaria report 2023. Geneva: World Health Organization; Licence: CC BY-NC-SA 3.0 IGO.