

Filipe Mateus Moraes Rodrigues  
Orientador: Antonio Oseas de Carvalho Filho

# **Método para detecção de bacilos de tuberculose usando RetinaNet**

Picos - PI  
19 de junho de 2023

Filipe Mateus Moraes Rodrigues  
Orientador: Antonio Oseas de Carvalho Filho

## **Método para detecção de bacilos de tuberculose usando RetinaNet**

Trabalho de Monografia submetida ao Curso de Bacharelado em Sistemas de Informação como requisito parcial para obtenção de grau de Bacharel em Sistemas de Informação. Orientador: Prof. Dr. Antonio Oseas de Carvalho Filho.

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

Picos - PI  
19 de junho de 2023

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

**R696m** Rodrigues, Filipe Mateus Morais

Método para detecção de bacilos de tuberculose usando RetinaNet [recurso eletrônico] / Filipe Mateus Morais Rodrigues - 2023.  
42 f.

1 Arquivo em PDF

Indexado no catálogo *online* da biblioteca José Albano de Macêdo-CSHNB  
Aberto a pesquisadores, com restrições da Biblioteca

Trabalho de Conclusão de Curso (Graduação) – Universidade Federal do Piauí, Bacharelado em Sistemas de Informação, Picos, 2023.  
“Orientador : Prof. Dr. Antonio Oseas de Carvalho Filho”

1. RetinaNet. 2. Rede neural. 3. Tuberculose. 4. Baciloscopia de escarro. 5. Detecção de bacilos - tuberculose. I. Carvalho Filho, Antonio Oseas de. II. Título.

**CDD 005.3**

MÉTODO PARA DETECÇÃO DE BACIOS DE TUBERCULOSE USANDO  
RETINANET

FILIFE MATEUS MORAES RODRIGUES

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

Data de Aprovação

Picos – PI, 14 de agosto de 2023

Documento assinado digitalmente  
**gov.br** ANTONIO OSEAS DE CARVALHO FILHO  
Data: 14/08/2023 15:41:37-0300  
Verifique em <https://validar.iti.gov.br>

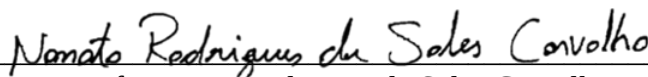
---

Prof. Antonio Oseas de Carvalho Filho



---

Prof. Edson Damasceno Carvalho



---

Prof. Nonato Rodrigues de Sales Carvalho

# Agradecimentos

Primeiramente, gostaria de expressar minha profunda gratidão a Deus, por me conceder força, saúde e sabedoria durante todo o período da graduação, por sempre estar comigo e por permitir que pessoas incríveis fizessem parte dessa jornada.

Gostaria de estender meus agradecimentos a meu pai, José Ferreira e minha mãe, Maria Nenoza, peças fundamentais em minha vida, sem eles, toda essa caminhada não seria possível, agradeço por cada sacrifício e incentivo para que eu pudesse concluir essa etapa. À todos os familiares que me apoiaram incondicionalmente, sempre me ajudando e encorajando a buscar os meus objetivos.

Ao meu orientador Antonio Oseas, que sempre me apoiou, se esforçou ao máximo para que pudesse extrair o melhor de mim, sou grato por ter sido guiado por uma excelente pessoa e profissional. Além disso, agradeço a todos os demais professores da UFPI, por compartilharem seus conhecimentos e experiências. Obrigado a todos por serem inspiração ao longo dessa trajetória.

Sou grato aos amigos que fiz durante o curso, pois compartilhamos diversos momentos de aprendizado, descontração e superações, fortalecendo nossos laços. Em especial aos colegas de Iniciação Científica, Francisco José e Mateus Assis, realizamos muitos projetos, trabalhos e atividades juntos, não poderia deixar de esclarecer minha imensa gratidão por tanto. Também, aos demais amigos de graduação: Eurico Delmondes, Lucas Lopes, Jeanderson Gomes, Marcos Eduardo e Welison Andrade. Obrigado por tornarem toda graduação mais leve e prazerosa.

E por fim, aos amigos, Ryan Sales e Vitória Carvalho, veteranos da Iniciação Científica, por toda a paciência e dedicação que tiveram em ensinar diversas coisas e se mostrarem sempre disponíveis a tirarem dúvidas e ajudar. Muito obrigado.

*Charles Chaplin*

*A persistência é o caminho do êxito.*

# Resumo

A tuberculose é uma doença infecciosa transmitida pela *Mycobacterium tuberculosis*, podendo afetar diversos órgãos do corpo humano, mas frequentemente atingindo o pulmão. Sendo a principal forma de diagnóstico a baciloscopia de escarro, exame no qual se analisa a expectoração do paciente através do microscópio em busca de bacilos, sendo essa uma técnica tanto para o diagnóstico como para acompanhamento da doença. A detecção precoce de doenças assume um papel crucial na preservação da vida dos pacientes, uma vez que o tratamento tardio pode resultar em consequências fatais. Neste cenário, abordagens baseadas em métodos computacionais têm sido objeto de pesquisa para auxiliar os profissionais de saúde na detecção da tuberculose através da análise de imagens de baciloscopia, a fim de aprimorar a precisão do diagnóstico e aumentar a probabilidade de tratamento oportuno, reduzindo complicações futuras. Nesse trabalho, é apresentado uma metodologia para detecção automática de bacilos da tuberculose em imagens de baciloscopias de escarro, utilizando a RetinaNet, que é uma arquitetura de rede neural convolucional usada para detecção de objetos em imagens, construída em cima de redes neurais convolucionais, que são a base da *Deep Learning* para processamento de imagens. A arquitetura da RetinaNet é composta por várias camadas de convolução, seguidas por camadas de agrupamento e processamento, permitindo que a rede extraia características hierárquicas de diferentes níveis de abstração a partir das imagens de entrada. No processo, a RetinaNet foi testada com diversos pré-processamentos, com o intuito de extrair o máximo de características possíveis, combinado com os *backbones* disponibilizados pela arquitetura, como mérito de comparação. Um conjunto de 1218 imagens foram usados para a avaliação deste método. Os resultados foram promissores, obtendo um mAP de 57,5%, *precision* de 64,9%, *recall* de 70,4% e um *score f1-score* de 67,5%. Finalmente, o estudo comprovou que o método possui a capacidade de atuar como um sistema de auxílio no diagnóstico da tuberculose.

**Palavras-chaves:** Tuberculose, RetinaNet, Detecção de bacilos, Rede Neural, Baciloscopia de Escarro.

# Abstract

Tuberculosis is an infectious disease transmitted by *Mycobacterium tuberculosis*, which can affect various organs in the human body, but often targets the lungs. The main diagnostic method is sputum smear microscopy, in which the patient's expectoration is analyzed under a microscope to look for bacilli. This technique is used both for diagnosis and disease monitoring. Early detection of diseases plays a crucial role in preserving patients' lives, as delayed treatment can lead to fatal consequences. In this context, approaches based on computational methods have been the subject of research to assist healthcare professionals in tuberculosis detection through the analysis of sputum smear images. The goal is to improve diagnostic accuracy and increase the likelihood of timely treatment, thereby reducing future complications. In this work, a methodology is presented for the automatic detection of tuberculosis bacilli in sputum smear images using RetinaNet. RetinaNet is a convolutional neural network architecture used for object detection in images. It is built on top of convolutional neural networks, which form the basis of Deep Learning for image processing. The RetinaNet architecture consists of several convolutional layers, followed by grouping and processing layers, enabling the network to extract hierarchical features at different levels of abstraction from the input images. During the process, RetinaNet was tested with various pre-processing techniques, aiming to extract as many features as possible, combined with the backbones provided by the architecture for comparison purposes. A dataset of 1218 images was used to evaluate this method. The results were promising, achieving an mAP (mean average precision) of 57.5%, precision of 64.9%, recall of 70.4%, and an f1-score of 67.5%. Finally, the study confirmed that the method has the capability to act as an aid in the diagnosis of tuberculosis.



# Lista de ilustrações

|  |    |
|--|----|
| Figura 1 – Exemplo de uma imagem resultante da baciloscopia. Fonte: (KAGGLE, 2020). . . . .                  | 16 |
| Figura 2 – Exemplo do resultado do segundo pré-processamento. Imagens disponíveis no (KAGGLE, 2020). . . . . | 18 |
| Figura 3 – Resultado do merge. Fonte: (KAGGLE, 2020). . . . .  | 19 |
| Figura 4 – Rede neural simples. Adaptado de Yiu (2019). . . . .  | 20 |
| Figura 5 – Arquitetura da RetinaNet. Adaptado de Lin et al. (2020). . . . .                                  | 21 |
| Figura 6 – Esboço da metodologia proposta. Fonte: Autoria própria. . . . .                                   | 26 |
| Figura 7 – Exemplo de comparação da marcação de especialistas em relação ao modelo em imagem RGB. . . . .    | 31 |
| Figura 8 – Exemplo de imagem RGB marcada por especialistas e detecção do modelo. . . . .                     | 32 |
| Figura 9 – Imagem em tons de cinza com comparação da marcações de especialista e detecção do modelo. . . . . | 33 |
| Figura 10 – Exemplo de comparação da marcações de especialistas e do modelo. . .                             | 34 |

# Lista de tabelas

|  |    |
|--|----|
| Tabela 1 – Resumo dos trabalhos relacionados em comparação a proposta. . . . .                 | 25 |
| Tabela 2 – Experimentos realizados. . . . .  | 30 |
| Tabela 3 – Resultados obtidos com as imagens em RGB. . . . .                                   | 31 |
| Tabela 4 – Resultados obtidos com as imagens <i>GRAY</i> . . . . .                             | 32 |
| Tabela 5 – Resultados obtidos com as imagens no canal $Luv(u)$ + gaussiano +<br>CLAHE. . . . . | 33 |
| Tabela 6 – Resultados obtido das imagens com o merge . . . . .                                 | 34 |
| Tabela 7 – Comparação entre os trabalhos relacionados e a abordagem proposta. .                | 35 |

# Lista de abreviaturas e siglas

|            |   |
|------------|---|
| (AP)       | <i>Average Precision</i>                                    |
| (BA)       | Baciloscopia de escarro                                     |
| (BAAR)     | álcool-ácido-resistente                                     |
| (CLAHE)    | Equalização de histograma adaptativo limitado por contraste |
| (CNN)      | <i>Convolutional Neural Network</i>                         |
| (DL)       | <i>Deep Learning</i>  |
| (F1)       | <i>f1-Score</i>   |
| (FN)       | Falsos Negativos  |
| (FP)       | Falsos Negativos  |
| (FPN)      | <i>Feature Pyramid Network</i>                              |
| (ML)       | <i>Machine Learning</i>                                     |
| (mAP)      | <i>Mean Average Precision</i>                               |
| (Pre)      | <i>Precision</i>  |
| (PR curve) | <i>Precision-Recall</i>                                     |
| (RCNN)     | rede neural convolucional baseada em região mais rápida     |
| (RELU)     | Unidade Linear Retificada                                   |
| (Rec)      | <i>Recall</i>   |
| (SSD)      | <i>Single Shot Multibox Detector</i>                        |
| (TB)       | Tuberculose   |
| (VN)       | Verdadeiros Negativos                                       |
| (VP)       | Verdadeiros Positivos                                       |
| (XML)      | <i>Extensible Markup Language</i>                           |

# Sumário

|          |  |           |
|----------|--|-----------|
| <b>1</b> | <b>Introdução</b>                        | <b>12</b> |
| 1.1      | Objetivos                                | 13        |
| 1.2      | Organização do Trabalho                  | 14        |
| <b>2</b> | <b>Referencial Teórico</b>               | <b>15</b> |
| 2.1      | Tuberculose                              | 15        |
| 2.2      | Pré-processamento de imagens             | 17        |
| 2.2.1    | Conversão para tons de cinza             | 17        |
| 2.2.2    | Conversão para espaço de cores LUV       | 18        |
| 2.2.3    | Combinação de imagens                    | 18        |
| 2.3      | Aprendizado profundo                     | 19        |
| 2.4      | RetinaNet                                | 21        |
| 2.5      | Métricas de Valiação                     | 22        |
| <b>3</b> | <b>Trabalhos Relacionados</b>            | <b>24</b> |
| <b>4</b> | <b>Metodologia</b>                       | <b>26</b> |
| 4.1      | Aquisição de Imagens                     | 27        |
| 4.2      | Pré-processamento                        | 27        |
| 4.3      | Detecção e Treinamento                   | 28        |
| 4.4      | Validação                                | 29        |
| <b>5</b> | <b>Resultados e Discussões</b>           | <b>30</b> |
| 5.1      | Resultados do Método Proposto            | 30        |
| 5.1.1    | Experimento 1                            | 30        |
| 5.1.2    | Experimento 2                            | 32        |
| 5.1.3    | Experimento 3                            | 33        |
| 5.1.4    | Experimento 4                            | 34        |
| 5.2      | Comparação com os Trabalhos Relacionados | 35        |
| 5.3      | Discussão                                | 35        |
| <b>6</b> | <b>Conclusão</b>                         | <b>37</b> |
| <b>7</b> | <b>Publicações</b>                       | <b>38</b> |
|          | <b>Referências</b>                       | <b>39</b> |

# 1 Introdução

Frequentemente confundida como uma doença do passado, a tuberculose (TB) é uma doença infecciosa que afeta com mais frequência os pulmões, sendo essa a forma pulmonar, e pode afetar outras partes do corpo, como ossos e o sistema nervoso, sendo essa a forma extrapulmonar. A transmissão da tuberculose acontece por via respiratória, pela eliminação de aerossóis produzidos pela tosse, fala ou espirro de uma pessoa com tuberculose ativa (SAÚDE., 2022), ela é causada pelo *Mycobacterium tuberculosis* ou bacilo de Koch. A doença pode ser tratada, chegando ao nível de cura, e ser prevenida através de determinadas medidas, pois a bactéria possui a capacidade de ser transmitida de pessoa a pessoa (FERRI et al., 2014).

Em todo o mundo, a tuberculose é a 13<sup>a</sup> principal causa de morte e a segunda principal causa de morte infecciosa após o COVID-19 (acima do HIV/AIDS) (WHO, 2023). Em 2021, um total de 1,6 milhão de pessoas morreram de tuberculose no mundo, enquanto cerca de 10 milhões sofreram com a doença, destes, 1,1 milhão eram crianças, sendo a TB infantil a forma mais difícil de diagnosticar e tratar. A TB atinge na maioria dos casos adultos. Contudo, esta é uma doença que afeta todas as faixas etárias, sobretudo, indivíduos com um sistema imunológico debilitado, como portadores do vírus HIV, pessoas em estado de desnutrição e diabéticos, possuem mais chances de desenvolver a TB (WHO, 2023).

Devido à ausência de sintomas nos estágios iniciais, a perspectiva positiva da tuberculose está associada à identificação precoce da doença, posteriormente com a iniciação do tratamento (FURG, 2019). Os exames mais usados para o diagnóstico da tuberculose são o de baciloscopia de escarro (BA), cultura para micobactéria, teste rápido para tuberculose e o de radiologia de tórax (CAMPOS, 2006). Sendo a BA, um dos métodos mais utilizados, tanto para o diagnóstico como principalmente para o controle do tratamento da doença. Esta técnica consiste na observação da lâmina de escarro do paciente através do microscópio, onde é feita a pesquisa pelo bacilo, álcool-ácido-resistente (BAAR) (GUERRA; REGO; CONDE, 2008).

O propósito do tratamento da tuberculose é alcançar a cura e minimizar prontamente a disseminação da doença. Para atingir tal objetivo, é crucial que os medicamentos empregados tenham a capacidade de rapidamente diminuir a população de bacilos (interrompendo, assim, a transmissão) (SOTGIU et al., 2016). Diante disso, métodos que agilizassem nos resultados desses exames, seriam de extrema importância para os centros de tratamento e pacientes, ajudando na detecção e acompanhamento da TB.

Os especialistas analisam as imagens produzidas pelas BA, para assim efetuar o diagnóstico da doença. Esse processo é feito de maneira manual, muitas vezes suscetível a falha. Uma maneira de agilizar esse processo, tornando menos exaustante para o profissi-

onal, seria a automação do processo de detecção do bacilo da tuberculose. Nesse contexto, atualmente, estão em progresso o desenvolvimento de diversas técnicas computacionais que visam auxiliar os profissionais especializados na tomada de decisão. Essas técnicas têm o propósito de aumentar a precisão dos resultados e melhorar as possibilidades de o paciente receber tratamento adequado de forma rápida, evitando maiores complicações futuramente.

Dentre as abordagens computacionais no suporte ao diagnóstico da TB, destaca-se, o aprendizado profundo, do inglês, *deep learning* (DL), uma subárea do aprendizado de máquina, do inglês, *machine learning* (ML). Nessa abordagem, os modelos computacionais identificam padrões em imagens para realizar classificações precisas, permitindo discriminar adequadamente entre as diferentes classes em questão. (LECUN; BENGIO; HINTON, 2015). Em comparação ao aprendizado de ML, o DL é muito superior, pois pode aprender a partir de dados brutos e possui várias camadas ocultas que permitem aprender abstrações com base em entradas (MIOTTO et al., 2018).

Dessa forma, a arquitetura RetinaNet é uma DL que usa um *backbone* que atua como extrator de características e duas sub-redes, onde a primeira trabalha para classificação dos objetos e a segunda para regressão das *bounding boxes* (LIN et al., 2020). Além disso, a RetinaNet usa uma função *Focal Loss*. Esta função reduz o desequilíbrio das classes, tornando-a mais rápida do que outras arquiteturas de dois estágios, pois não há necessidade de classificar as regiões candidatas em um estágio posterior (LIN et al., 2020). Finalmente, como apresentado em Jiao et al. (2019), a RetinaNet pode detectar objetos pequenos e médios em uma imagem com uma superioridade relativa sobre outras arquiteturas.

No que diz respeito às imagens de BA, a análise das relações entre aquelas que contêm bacilos e as que não contêm pode ser relevante para a detecção das imagens que indicam a presença de contaminação ou as saudáveis. Com base nessa premissa, o presente estudo tem como objetivo aprofundar a análise desses aspectos no que se refere à detecção da TB em imagens de baciloscopias de escarro. Para atingir esse propósito, foi avaliado o desempenho da arquitetura RetinaNet em um conjunto de dados público contendo 1218 imagens.

## 1.1 Objetivos

Conforme o problema relatado, o objetivo desse trabalho é desenvolver um método para detectar os bacilos contido em exames baciloscópicos usando técnicas de pré-processamento e abordagem *deep learning*, que seja capaz de auxiliar no diagnóstico da tuberculose e no acompanhamento do paciente durante o tratamento da doença.

Os objetivos específicos são:

1. Construir uma abordagem promissora que aplica RetinaNet no problema da TB;

2. Analisar métodos de pré-processamento nas imagens visando ressaltar os atributos distintivos da doença;
3. Desenvolvimento de uma técnica automática para detecção de bacilos em imagens baciloscópicas;
4. Produzir uma abordagem relevante que ajude os especialistas no diagnóstico e auxilia no tratamento da doença.

## 1.2 Organização do Trabalho

A estrutura do trabalho está da seguinte forma: no Capítulo 2 é tratado sobre os conceitos gerais para o entendimento da abordagem proposta; o Capítulo 3 explora os trabalhos relacionados pertinentes ao tema abordado; o Capítulo 4 apresenta os métodos e procedimentos utilizados para alcançar os objetivos estabelecidos neste trabalho; o Capítulo 5 é apresentado e discutidos os resultados obtidos com a aplicação da metodologia proposta; e, o Capítulo 6 conclui o trabalho.

## 2 Referencial Teórico

Este capítulo tem o propósito de apresentar os conceitos fundamentais que embasam a proposta deste trabalho. Para isso, está dividido em cinco subseções: a subseção 2.1 trata da tuberculose, incluindo as abordagens atuais para diagnosticar a doença; a segunda, 2.2 explora a etapa do pré-processamento de imagens; a terceira, 2.3 oferece um contexto sobre o aprendizado profundo; a quarta subseção, 2.4 explana sobre a arquitetura proposta, RetinaNet; por fim, a última subseção, 2.5 trata sobre as métricas adotadas para avaliar os resultados da detecção das imagens.

### 2.1 Tuberculose

Um total de 1,6 milhão de pessoas morreram de tuberculose em 2021 (incluindo 187.000 pessoas com HIV) (WHO, 2023). Diante deste fato, a TB ainda demonstra uma ameaça à saúde pública, sua atuação no passado já foi bem mais atuante e relevante, porém, nos dias atuais, esquecida por muitos. Em 2020, os 30 países com alta carga de TB representaram 86% dos novos casos de TB. Oito países respondem por dois terços do total, com a Índia liderando a contagem, seguida pela China, Indonésia, Filipinas, Paquistão, Nigéria, Bangladesh e África do Sul (WHO, 2023). Atualmente a tuberculose afeta na maioria dos seus casos, países com alta densidade demográfica e países que ainda estão em estado de desenvolvimento.

A tuberculose é uma doença infecciosa que afeta diversos órgãos e aflige com mais frequência os pulmões, podendo levar a morte se não for tratada. A detecção da TB no estágio inicial é essencial para que os pacientes busquem tratamento. Embora, o risco de adoecimento seja maior nos primeiros dois anos após a primeira infecção, uma vez infectada a pessoa pode adoecer em qualquer momento de sua vida (SESA, 2022).

A doença possui algumas características distintas, como um extenso intervalo entre a infecção inicial e o aparecimento dos sintomas clínicos da doença. O período de incubação, em média, varia de 4 a 12 semanas antes que as primeiras lesões sejam detectadas. A maioria dos novos casos de doença pulmonar surge aproximadamente um ano após a infecção inicial. A transmissibilidade é máxima durante o período em que o paciente está eliminando bacilos e antes de iniciar o tratamento (NOGUEIRA et al., 2012).

Em pessoas que a doença desenvolve na forma ativa, os sintomas como febre, tosse, perda de peso ou sudorese noturna podem ser leves por um longo tempo (CAMPOS, 2006). Isso pode acarretar no atraso na procura de atendimento, implicando na transmissão da bactéria para mais pessoas. Uma pessoa com tuberculose ativa pode infectar de 5 a 15 pessoas mantendo contato próximo a elas, geralmente em convivência familiar, ao longo de um ano (CRUZ, 2023). Caso não haja um tratamento adequado, a probabilidade de



morte é em média de 45% para pessoas HIV-negativa, e quase todas pessoas HIV-positivo acabam chegando a óbito (WHO, 2023).

O diagnóstico da tuberculose pode ser feito de diversas maneiras, como teste de Sensibilidade aos fármacos, cultura, radiografia de tórax, sendo o que se utiliza em maior escala por ser barato comparado com os outros métodos a BA. Estima-se que 66 milhões de vidas foram salvas através do diagnóstico e tratamento da TB entre 2000 e 2020 (WHO, 2023). A Figura 1 mostra um exemplo de imagem digital gerada a partir de um exame baciloscópico contendo bacilos marcados por especialistas.



Figura 1 – Exemplo de uma imagem resultante da baciloscopia. Fonte: (KAGGLE, 2020).

A interpretação desses exames está relacionado a análise visual e interpretação do profissional. Sendo assim, distintos especialistas podem ter conclusões diferentes para o mesmo exame. Além do fato desse processo ser repetitivo, é preciso ter um alto nível de atenção e precisão sobre os detalhes presentes nas imagens, pois nas (BA) existem várias estruturas parecidas com o bacilo, tornando um processo com grandes chances de erro, principalmente quando um mesmo especialista analisa diversas lâminas consecutivamente. A detecção automática de bacilos da tuberculose por meio de imagens de microscopia de campo claro foi inicialmente proposta em 2008 por Costa et al. (2008). Nesse contexto, a análise desses exames por meio da computação tem sido de grande relevância para gerar uma segunda opinião sobre o diagnóstico e também no acompanhamento do paciente no tratamento da doença.

## 2.2 Pré-processamento de imagens

O objetivo do pré-processamento é a tentativa de obter mais informações do que os dados originais disponibilizam. Refere-se a um conjunto de técnicas e etapas aplicadas a imagens antes do treinamento de um modelo. Essas técnicas são essenciais para melhorar a qualidade dos dados de entrada e tornar a extração de informações mais precisa e eficiente (SZELISKI, 2022). Desse modo, torna-se essencial realizar o tratamento dos dados antes de empregá-los na elaboração do modelo.

De acordo com Russ (2006), as técnicas de pré-processamento são utilizadas com o objetivo de aprimorar imagens, realizando correções de iluminação, contraste, distorções e nitidez. Algumas dessas técnicas incluem a conversão do canal original de cor para a escala de cinza, aplicação de filtros para aumentar o contraste, equalização de histograma, entre outras. Um exemplo específico é a normalização de imagens para alimentação da entrada de uma *Convolutional Neural Network* (CNN). Nesse processo, as imagens são ajustadas para terem o mesmo tamanho, resolução e orientação, garantindo que todas as imagens de entrada possuam características consistentes.

Devido à sua simplicidade e rapidez, a técnica de coloração de Ziehl-Neelsen é tradicionalmente utilizada em laboratórios para o diagnóstico de micobacterioses (TAN-SUPHAISIRI; KLADPHUANG, 2002), esse método é comumente usado sobre a lâmina de esfregaço de escarro, com o objetivo de aprimorar a detecção do BA, ainda assim, muitas vezes isso se mostra insuficiente para estabelecer um nível de contraste que permita a rápida detecção do bacilo.

Nesse trabalho, foram realizados três tipos de pré-processamento nos canais de cores da imagem original RGB (do inglês *red*, *green*, *blue*): a subseção 2.2.1 relata da transformação das imagens em tons de cinza, a segunda subseção 2.2.2 na conversão de cores para o LUV e aplicação de filtros, e a subseção 2.2.3 o merge. Para cada pré-processamento, diversos modelos foram criados.

### 2.2.1 Conversão para tons de cinza

O primeiro métodos amplamente utilizados em aplicações de processamento de imagem foi a conversão das imagens para tons de cinza. Esse processo é essencial para simplificar a representação visual de uma imagem e reduzir a complexidade dos dados, tornando mais fácil o seu processamento e análise. A conversão de uma imagem colorida para tons de cinza envolve a remoção das informações de cor, mantendo apenas o brilho, o contraste, de cada pixel (LU; XU; JIA, 2014). Uma vez convertida para tons de cinza, a imagem resultante ocupa menos espaço de armazenamento e requer menos recursos computacionais para processamento em comparação com a imagem colorida original.

### 2.2.2 Conversão para espaço de cores LUV

Já o segundo melhoramento, trata-se de uma conversão da imagem RGB para o espaço de cores LUV. O espaço Luv foi projetado especificamente para cores emissivas, que correspondem a imagens captadas por uma câmera ou programa de renderização de computação gráfica, o L representa a luminância, U e V representam as informações de cor (crominância) (SHEEDY, 2005), os canais L e V são descartados, mantendo-se apenas o canal U, em seguida, é aplicado um filtro gaussiano, o (elemento estruturante utilizado é igual a 3), com o intuito de reduzir o ruído, na qual, esse filtro é baseado na função de distribuição gaussiana (também conhecida como distribuição normal), desenvolvida pelo matemático e astrônomo Carl Friedrich Gauss (DUNNINGTON; GRAY; DOHSE, 2004), por fim é aplicada uma técnica de equalização de histograma adaptativo limitado por contraste (CLAHE) (REZA, 2004). A Figura 2 mostra cada transformação que a imagem recebeu.

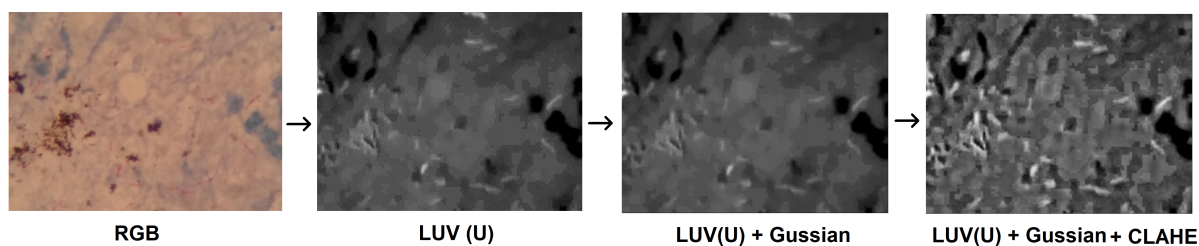


Figura 2 – Exemplo do resultado do segundo pré-processamento. Imagens disponíveis no (KAGGLE, 2020).

### 2.2.3 Combinação de imagens

Para o terceiro pré-processamento, são utilizadas três cópias da mesma imagem, sendo aplicado um tipo de processo em cada uma delas. No final, as imagens são unidas. Na primeira cópia, a imagem original em RGB é convertida para o espaço de cores LUV, e apenas o canal U é mantido. A segunda cópia é obtida a partir da primeira cópia. Ou seja, a imagem é convertida para o canal U do esquema LUV e em seguida, é aplicado um filtro gaussiano (com elemento estruturante igual a 3), para redução do ruído. Para gerar a terceira cópia, os mesmos processos da segunda cópia são aplicados, e além disso, é adicionado o filtro CLAHE. A partir dessas três cópias da mesma imagem, todas elas são combinadas, sendo esse processo denominado como "merge". Na Figura 3 representa o resultado final de todas as etapas dos pré-processamentos.

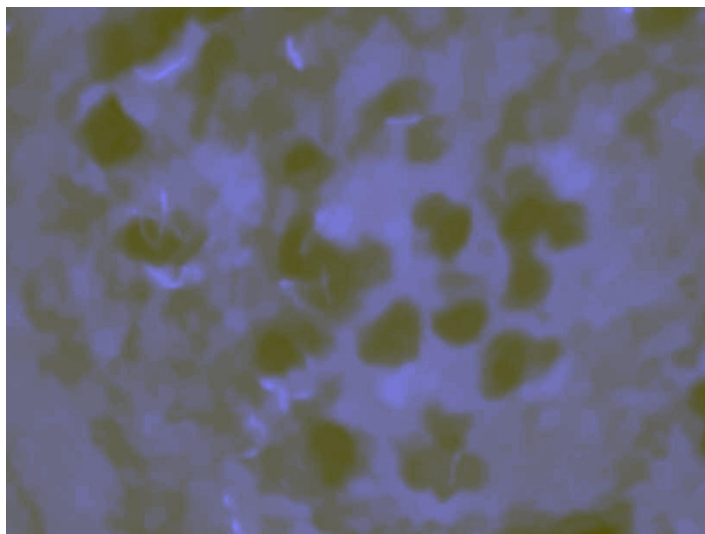


Figura 3 – Resultado do merge. Fonte: (KAGGLE, 2020).

## 2.3 Aprendizado profundo

Entende-se o aprendizado profundo, comumente conhecida como DL, um ramo do *Machine Learning* (aprendizado de máquina), o qual se fundamenta em um conjunto de algoritmos que procuram modelar abstrações de alto nível utilizando-se de um grafo profundo com múltiplas camadas de processamento, compostas por diversas transformações lineares e não lineares (CHAGAS, 2019). As contribuições da *deep learning*, são de grande notoriedade, podendo ajudar em diversas áreas como, no reconhecimento visual, reconhecimento de fala, detecção de objetos, dentre outras possibilidades (GRACE et al., 2018).

O aprendizado profundo, em sua essência, representa uma percepção de inteligência artificial que se assemelha à capacidade humana e é capaz de gerar resultados com base no conhecimento adquirido por essa assimilação (PACHECO; PEREIRA, 2018). Segundo Bengio et al. (2009), o processo de aprendizado das arquiteturas de DL é composto por diversos níveis de abstração, abrangendo desde um conceito simples até um conceito extremamente complexo, nesse contexto, uma das principais atribuições do aprendizado profundo é a extração de características de nível mais elevado, compostas pelas camadas de nível inferior. Dessa forma, os algoritmos de DL têm a capacidade de analisar dados não-estruturados sem a necessidade de qualquer tipo de pré-processamento ou supervisão (GOODFELLOW; BENGIO; COURVILLE, 2016).

O processo da DL visa fazer a identificação de uma grande e complexa estrutura em um conjunto de dados usando o algoritmo de retropropagação para que a rede possa alterar seus parâmetros ocasionando a diminuição da taxa de erro. Segundo CHAGAS (2019), os dados são transmitidos através de cada camada, com a saída da camada anterior concedendo entrada para a próxima camada. A primeira camada em uma rede é denominada de camada de entrada e a última é a camada de saída. As camadas intermediárias são cha-

camadas de camadas ocultas, sendo que cada camada da rede é formada por um algoritmo simples e uniforme engloba uma espécie de função de ativação.

Na Figura 4 é apresentado um rede neural simples, nela está contida apenas duas camadas ocultas, para que essa rede de inteligência artificial simples possa se tornar uma rede de aprendizado profundo, a quantidade de camadas ocultas deverá aumentar e em paralelo a isso, mais camadas ocultas totalmente conectadas deveram ser adicionadas a ela, aumentando assim a quantidade de camadas intermediárias e tornando essa rede em uma *deep learning*, mas nem sempre aumentar o número de camadas será uma boa solução (ARNOLD et al., 2011).

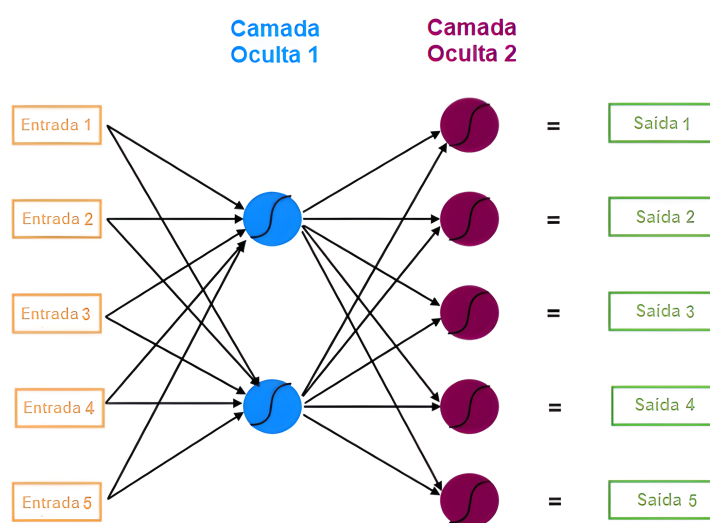


Figura 4 – Rede neural simples. Adaptado de Yiu (2019).

De modo geral, as redes neurais mais profundas são modelos com mais camadas, ou seja camadas totalmente conectadas em grande quantidade (ANALITICA, 2018). Essa teia de comunicação entre as camadas é o que permite a rede aprenda o que é imposto, ocasionando na tomada de decisão da rede, no problema na qual ela está sendo utilizada, obtendo ações, tomando decisões e prevendo cenários futuros. Hoje, o aprendizado profundo é responsável pelos sistemas de melhor desempenho em quase todas as áreas de pesquisa de inteligência artificial (RADAR, 2022).

No que diz respeito ao poder computacional das DL, é certo que elas podem ser um tanto custosas, uma vez que suas estruturas são complexas e requerem uma grande quantidade de dados para um treinamento adequado. Isso implica na necessidade de um significativo poder computacional para aplicar suas práticas de forma eficiente.

Atualmente a quantidade de arquiteturas e algoritmos construídos e aprimorados no uso do aprendizado profundo vem crescendo exponencialmente, para cada desafio, existe uma arquitetura que age de forma mais eficiente do que outras, essas metodologias estão sendo utilizadas por todo o mundo em diversas áreas, com o intuito de facilitar, melhorar

ou até mesmo resolver determinado problema. Diante dessa grande diversidade de metodologias arquitetônicas, a RetinaNet, rede de detecção de objetos, foi escolhida para investigar o problema proposto por esse projeto.

## 2.4 RetinaNet

A arquitetura RetinaNet foi proposta por [Lin et al. \(2017\)](#) em um estudo para detecção de imagens em diferentes escalas. De acordo com [Lin et al. \(2017\)](#), o objetivo desta rede é conduzir a hierarquia de características de uma rede neural convolucional (CNN) enquanto cria uma pirâmide de características que tem uma semântica forte em todas as escalas. Como resultado, uma pirâmide de características com semântica rica em todos os níveis é obtida com uma única imagem de entrada. Assim, a RetinaNet pode criar uma pirâmide de atributos na rede sem abrir mão do poder de representação, velocidade de treinamento ou memória.

A arquitetura consiste em uma rede de backbone chamada *Feature Pyramid Network* (FPN), que é construída em um ResNet (ResNet-18, ResNet-34, ResNet-50, ResNet-101 ou ResNet-152) de forma convolucional, obtendo uma imagem de qualquer tamanho e produzindo um rico mapa de características de tamanho proporcional em vários níveis da pirâmide de recursos. A construção da pirâmide de características envolve dois caminhos, um de baixo para cima e um de cima para baixo, bem como conexões laterais ([LIN et al., 2017](#)). Como podemos analisar na Figura 5, o caminho de cima para baixo está totalmente conectada com a camada que está de baixo para cima.

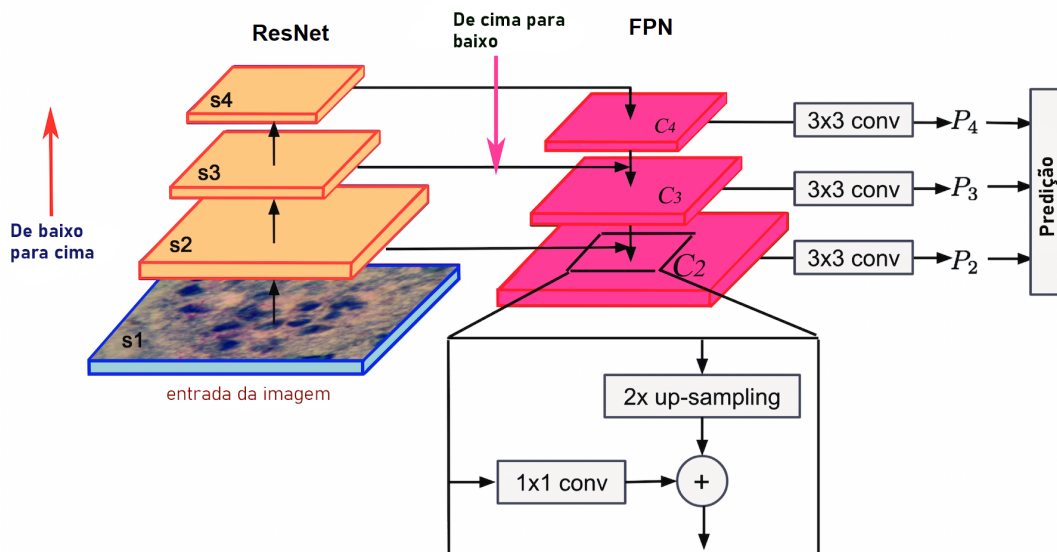


Figura 5 – Arquitetura da RetinaNet. Adaptado de [Lin et al. \(2020\)](#).

As sub-redes de classificação identificam a probabilidade da presença de um objeto em cada uma das caixas delimitadoras pertencentes a uma determinada classe de objetos.

Esta sub-rede compartilha todos os seus parâmetros nos diferentes níveis da pirâmide, consistindo assim em quatro camadas convolucionais com 256 filtros, seguidas por ativações de Unidade Linear Retificada (RELU). Em seguida, outras camadas convolucionais, seguidas por ativação sigmóide, são aplicadas (LIN et al., 2020). A sub-rede de regressão funciona em paralelo com a rede de classificação, sendo responsável por detectar o tamanho e a forma do objeto se ele existir. Tendo como saída quatro valores que correspondem à localização espacial da caixa delimitadora (LIN et al., 2020).

Para predição, a RetinaNet usa âncoras. Âncoras são caixas delimitadoras, usadas durante o treinamento de acordo com o tamanho e a posição do objeto. Em vez de usar uma única escala de âncora para cada nível, várias âncoras de aspecto são definidas para cada nível, dependendo do tamanho das caixas delimitadoras e do espaçamento entre elas, cobrindo assim várias escalas e proporções das caixas delimitadoras (LIN et al., 2017).

Os modelos de detecção de estágio único, enfrentam um problema extremo de desequilíbrio entre as classes de primeiro plano e plano de fundo devido à amostragem densa das caixas âncora (que representam possíveis localizações de objetos) (LIN et al., 2020). Para a resolução desse problema, é introduzida nessa arquitetura a função *Focal Loss*, justamente para aliviar o problema do desequilíbrio extremo entre o objeto e o fundo da imagem (HUMBARWADI, 2020). Em cada camada da pirâmide, é possível encontrar milhares de âncoras, das quais apenas algumas serão atribuídas a objetos de interesse, enquanto a grande maioria representará a classe de fundo. Embora esses exemplos resultem em valores de perda relativamente baixos, eles podem, coletivamente, sobrecarregar o modelo. Para resolver esse problema, o *Focal Loss* atua reduzindo a influência desses objetos menos relevantes e aumentando a importância de corrigir exemplos mal classificados (ARCGIS, 2023).

De modo geral, esta rede mostrou-se de extrema utilidade para a aplicação das funcionalidades deste trabalho, pois as imagens de baciloscopia, no seu fundo, contém várias estruturas parecidas com o bacilo, que por sua vez, este, é o maior desafio para a detecção. Outro atributo da RetinaNet é a sua capacidade em detectar objetos pequenos, encaixando justamente com a proposta do trabalho pois o bacilo em relação ao tamanho total da imagem, é bem pequeno. Sendo assim, a RetinaNet pode evidenciar superioridade as demais arquiteturas.

## 2.5 Métricas de Valiação

Para validar os resultados, foram empregadas métricas de avaliação estatística. Entre elas a *mean average precision* (mAP), segundo Yohanandan (2020), é uma técnica amplamente utilizada para avaliar a qualidade dos modelos de detecção de objetos. Outras técnicas foram utilizadas, tais como, *precision* (Pre), *Recall* (Rec), *F1-Score* (F1). Essas métricas foram calculadas com base na matriz de confusão. A matriz de confusão fornece

uma hipótese das medidas eficazes do modelo de classificação, mostrando o número de classificações corretas versus as classificações previstas para cada classe em um determinado conjunto de exemplos, para realizar essa validação, a matriz de confusão organiza os resultados de predição em quatro classes, sendo elas, verdadeiros negativos (VN), falsos negativos (FN), verdadeiros positivos (VP) e falsos positivos (FP) (METZ, 1978).

No problema de detecção de objetos, é necessário localizar todos os objetos presentes na imagem. Cada objeto deve ser delimitado por uma caixa de delimitadora, e cada caixa deve conter uma etiqueta indicando a que classe pertence (KARAGIANNAKOS, 2019). Diante disso, A métrica comumente usada para a detecção de objetos é o mAP, que representa um número de 0 a 100, quanto mais elevado o número, maior será a confiança no resultado da detecção. Cada caixa delimitadora possui um score, que representa o grau de certeza na classificação do objeto. com esses valores, é calculada uma curva *precision-recall* (PR curve) para cada classe de objeto. A *average precision* (AP) corresponde à área abaixo dessa curva. Primeiramente, a AP é calculada para cada classe e, em seguida, obtém-se a média de todas as classes. O resultado final é o mAP (HULSTAERT, 2018).

A *precision*, também conhecida como Valor Preditivo Positivo, representa a proporção de verdadeiros positivos (VP) em relação a todas as predições positivas, ou seja, a soma dos verdadeiros positivos e dos falsos positivos (VP + FP) como mostra na Equação 2.1. Em suma, ela mede quantas das predições positivas feitas foram corretas.

$$\text{Pre} = \frac{\text{VP}}{\text{VP} + \text{FP}} \quad (2.1)$$

O *recall*, também conhecido como Sensibilidade, expressa pela Equação 2.2, representa a proporção de verdadeiros positivos (VP) em relação à soma dos verdadeiros positivos com os falsos negativos (FN). Em outras palavras, ele mede a capacidade do classificador de acertar a classe positiva

$$\text{Rec} = \frac{\text{VP}}{\text{VP} + \text{FN}} \quad (2.2)$$

Por fim, o *f1 Score*, dada pela Equação 2.3, é a média harmônica entre *precision* (Pre) e *recall* (Rec). É uma boa métrica para avaliar a qualidade dos classificadores, pois ele considera ambas as quantidades de classificações erradas, tanto os falsos positivos, quanto os falsos negativos.

$$F1 = 2 \times \frac{\text{Pre} \times \text{Rec}}{\text{Pre} + \text{Rec}} \quad (2.3)$$



## 3 Trabalhos Relacionados

Recentemente, tem havido um crescente interesse nos estudos literários voltados para a detecção da tuberculose, especialmente na busca por aprimorar a qualidade das imagens das lâminas resultantes da baciloscopia, segmentar os bacilos presentes, classificar as lâminas de acordo com a presença de bacilos da TB e identificar esses bacilos de forma precisa. Nesta seção, será abordado o uso da RetinaNet em diferentes contextos, seguido por uma análise de trabalhos recentes e relevantes relacionados aos métodos de detecção da TB. Por fim, realizaremos uma comparação detalhada com o método proposto no presente trabalho.

Gräbel et al. (2020) faz uso do RetinaNet para a detecção de células hematopoiéticas. Para adaptar o RetinaNet ao problema proposto pelos autores, as âncoras foram modificadas para uma forma circular, que cobre a forma das células com mais precisão. Os dados utilizados para avaliar o método proposto consistem em um número total de 101 patches contendo aproximadamente 7400 células, cada patche com tamanho de  $2048 \times 2048$ , obtendo um *f1-score* de 92.5%.

Pho, Amin e Yoshitaka (2018) detecta, segmenta e identifica automaticamente diversas espécies de protozoários em micrografias em imagens de cistos e oocistos, com a RetinaNet. Os autores usam a arquitetura para prever as *bounding boxes* de protozoários nas imagens. A rede de segmentação decodifica os recursos extraídos pelo *backbone* da RetinaNet em imagens de canal único para segmentação. O conjunto de imagens utilizado pelos autores foi de 38 imagens com 43 amostras para treinamento e 31 imagens com 74 amostras para teste.

Kant e Srivastava (2018) Usa uma arquitetura de rede neural simples, uma estratégia de patchwise para a detecção de bacilos. O método proposto usa um campo de visão do microscópio, pedaços de 20px por 20px da imagem de entrada. Foram utilizados bases de dados diferentes, totalizando um total de 202 imagens. Os resultados obtidos com as métricas de *recall* e *precision* foram de 83,78% e 67,55% respectivamente para detecção de bacilos.

Santos, Brito e Filho (2023) apresenta uma abordagem para a detecção de bacilos da TB através da utilização da arquitetura EfficientDet, com a variação de *backbones* em seus experimentos. Para esse propósito, foi utilizado o mesmo conjunto de dados desse trabalho (KAGGLE, 2020), e durante o treinamento, foram explorados 4 modelos de cores diferentes. A avaliação desses modelos foi realizada por meio de validação cruzada com 5 *folds*, proporcionando uma análise abrangente das performances alcançadas na detecção dos bacilos da TB. Esses resultados contribuem para o avanço das técnicas de detecção e diagnóstico da doença, oferecendo perspectivas para futuras pesquisas na área.

El-Melegy et al. (2019) propôs uma abordagem automático para detecção de bacilo em

imagens de tuberculose adquiridas com microscopia de campo claro, usando uma estrutura de rede neural convolucional baseada em região mais rápida (RCNN), para localização e classificação de objetos em imagens combinando CNNs e métodos de proposta de região. O modelo foi avaliado com 500 imagens, corada com método de ziehl neelsen, contendo 2500 bacilos, provenientes da base de imagens pública ZNSM-iDB que consiste em amostras de várias categorias.

Já Swetha e Sankaragomathi (2020) usaram uma técnica de processamento de imagem utilizada pelo classificador CNN para detectar imagens de baciloscopia de escarro e posteriormente classificar as imagens como leve, moderada ou grave. É realizado um pré-processamento para reduzir o ruído e modificar a intensidade da entrada tornando as imagens resultantes mais adequadas ao trabalho a ser realizado.

Conforme evidenciado nesta seção, certos estudos na literatura identificaram a presença de bacilos, mas não empregaram o RetinaNet. Este projeto propõe um método para detectar bacilos usando o RetinaNet. Além disso, a adoção da RetinaNet com âncoras e a utilização do *Focal Loss* podem conduzir a resultados mais eficientes.

Tabela 1 – Resumo dos trabalhos relacionados em comparação a proposta.

| Trabalho                       | Problema                               | Método  | Amostra de Imagens | Métricas de Validação                                   |
|--------------------------------|--|---|--------------------|---|
| Gräbel et al. (2020)           | Detecção de células hematopoiéticas    | RetinaNet   | 101                | F1 = 92,5%  |
| Pho, Amin e Yoshitaka (2018)   | Segmentação e detecção de protozoários | RetinaNet   | 69                 | mAP = 53%<br>Pre = 48%<br>Rec = 94%                     |
| Kant e Srivastava (2018)       | Detecção de bacilos da Tuberculose     | Rede neural simple + estratégia de patchwise                                | 202                | Pre = 67,55%<br>Rec = 83,78%                            |
| Santos, Brito e Filho (2023)   | Detecção de bacilos da Tuberculose     | Pré-processamento + detecção de bacilos usando EfficienDet e seus backbones | 1265               | IoU = 52,3%<br>Pre = 69,4%<br>Rec = 92,5%<br>F1 = 77,4% |
| El-Melegy et al. (2019)        | Detecção de bacilos da Tuberculose     | classificador CNN   | 500                | Pre = 85,1%<br>Rec = 98,4%<br>F1 = 91,2%                |
| Swetha e Sankaragomathi (2020) | Classificação de bacilos               | classificador CNN   |                    | Acc = 99,5%   |
| Método Proposto                | Detecção de bacilos da Tuberculose     | RetinaNet   | 1218               | mAP = 79,4%<br>Pre = 64,9%<br>Rec = 70,4%<br>F1 = 67,6% |

A Tabela 1 apresenta um resumo dos trabalhos relacionados a este projeto. É importante notar que a maioria das bases de dados utilizada nesses trabalhos são bastante reduzidas, podendo resultar em overfitting, baixa capacidade de generalização e podendo tornar o aprendizado difícil, exceto o trabalho desenvolvido por Santos, Brito e Filho (2023), no qual, utilizam conjunto de imagens similar ao método proposto e os resultados apresentados são semelhantes aos obtidos no trabalho. Sendo assim, atualmente a literatura não apresenta métodos para detecção de bacilos da tuberculose usando a arquitetura RetinaNet, tornando a proposta inovadora e com altas chances de obter de bons resultados.

## 4 Metodologia

Nesta seção será apresentada a metodologia utilizada para a detecção de bacilos em imagens de baciloscopia de escarro, por meio da *deep learning*, uma técnica computacional de inteligência artificial. A metodologia usada para a obtenção dos resultados proposto por esse projeto consiste em: aquisição das imagens, pré-processamento, detecção e validação dos resultados. A Figura 6 apresenta o fluxo da metodologia descrita para a obtenção dos resultados.

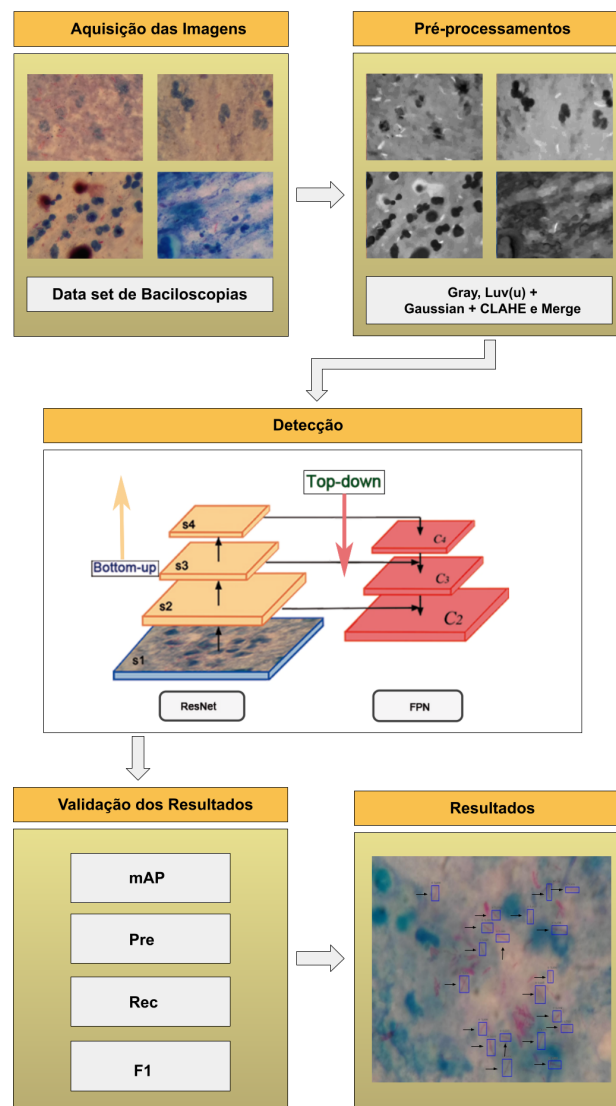


Figura 6 – Esboço da metodologia proposta. Fonte: Autoria própria.

As seções subsequentes apresentam as etapas desenvolvidas para a detecção automática da TB em lâminas de esfregaço de escarro usando RetinaNet. O método pode ser resumido nas seguintes etapas: Na Seção 4.1 relata sobre aquisição de imagens de TB;

por seguinte, na Seção 4.2 mostra a aplicação de pré-processamentos com o intuito de destacar o bacilo do fundo; na Seção 4.3 mostra o processo de treinamento da arquitetura RetinaNet para detectar bacilo; e, finalmente, na Seção 4.4 mostra como foi feita a avaliação dos resultados obtidos.

## 4.1 Aquisição de Imagens

A primeira etapa para a inicialização do trabalho prático aplicado é a aquisição das imagens. Diferentes bases são de extrema relevância para o trabalho, pois assim, o aprendizado da rede não fica limitado somente para um conjunto de imagens. Trabalhar com mais de uma base é de extrema relevância, na qual é feita a comparação entre os modelos gerados de cada base, observando e selecionando os melhores resultados e o que pode ser melhorado em um treinamento bem sucedido de uma base, para um não sucedido, no entanto, nesse trabalho, foram utilizados dados de apenas uma única base de dados.

A quantidade de imagens é um ponto a ser destacado, quanto mais imagens, maior será a precisão dos resultados. O data set utilizado nesse projeto é disponibilizada pelo *AI Research and Automated Laboratory Diagnostics* através do *kaggle*, para fins de pesquisa e educação. O conjunto contém 1218 imagens de baciloscopia de escarro com caixas delimitadoras de 3734 bacilos (KAGGLE, 2020). Às imagens possuem uma resolução de  $1224 \times 1623$ , esse foi o utilizada para entrada na RetinaNet, e cada imagem está relacionada á um arquivo *Extensible Markup Language* (XML), que contém detalhes das coordenadas de cada caixa delimitadora dos bacilos presentes na imagem.

## 4.2 Pré-processamento

Essa etapa é de suma importância, pois será determinante para a qualidade final dos dados e resultados que serão analisados como foi relatado na Seção 2.2. O objetivo do pré-processamento é a tentativa de obter mais informações do que os dados originais disponibilizam (RUSS, 2006). Essas técnicas devem ser utilizadas para o aprimoramento da base de dados.

Como visto na Seção 2.2, antes do treinamento com a RetinaNet, foram aplicados os métodos de transformações de canais de cores, o primeiro, foi transformado os dados para tons de cinza, a próxima transformação de base foi de RGB para LUV, extraíndo o canal "u" e aplicando o filtro gaussiano e CLAHE e com o "merge", com o intuito de aprimorar os aspectos estruturais do bacilo em relação ao fundo.

### 4.3 Detecção e Treinamento

A RetinaNet é responsável por fazer a detecção dos bacilos em cada imagem bacilos-cópica. A arquitetura como uma rede de detecção de objetos baseada em uma combinação de componentes de rede como *Feature Pyramid Network* (FPN) e *Single Shot Multibox Detector* (SSD), apresenta vantagens significativas na localização e classificação precisa dos bacilos em cada imagem analisada.

No entanto, para que a RetinaNet atinja o objetivo desejado, é fundamental que seus parâmetros sejam cuidadosamente ajustados e otimizados de acordo com a natureza específica do problema em questão. Dentre os parâmetros a serem definidos, destacam-se a função de ativação, que tem um papel fundamental na capacidade da rede de aprender e generalizar padrões complexos nos dados, e o *batch size*, que influencia diretamente o processo de treinamento e a eficiência do algoritmo (DEVARAKONDA; NAUMOV; GARLAND, 2017).

Além disso, de acordo com Sinha et al. (2010), a quantidade de épocas é o número de vezes que o conjunto completo de dados que é apresentado à rede durante o treinamento, é um controle importante para garantir que o modelo aprenda de forma adequada, evitando tanto o subajuste quanto o superajuste aos dados de treinamento.

Portanto, antes de alcançar um modelo bem-sucedido, são necessários testes iniciais para encontrar os parâmetros ideais que garantam o aprendizado eficiente da RetinaNet. Essa etapa de ajuste é essencial para alcançar um modelo de detecção de bacilos preciso e robusto, contribuindo para diagnósticos mais rápidos e confiáveis.

Neste trabalho, foi seguida a configuração da arquitetura RetinaNet original. Após a realização de alguns experimentos, foi definido que às imagens seriam dadas como entrada na rede com seu tamanho original, não havendo assim perda de informação. A divisão da base de imagem foi feita aleatoriamente, destinando para subconjuntos de treinamento (70%), validação (10%) e teste (20%). Assim, o resultado da divisão foi em 876 imagens para treinamento, 98 para validação e 244 para teste.

No que se diz respeito aos parâmetros para o treinamento da rede, alguns foram selecionados empiricamente durante os testes. O algoritmo de otimização usado foi o Adam para uma atualização eficaz dos pesos durante o treinamento; já a taxa de aprendizagem foi fixa em 0,01 de modo a se garantir um aprendizado gradual e eficiente do modelo; o *batch size* foi definido para 2 como forma de garantir a aprendizagem da rede durante o treinamento. A seleção dos demais parâmetros foi feita baseada na literatura, seguindo assim o da rede idealizada por Lin et al. (2017). A cada época, foram realizadas 437 iterações para obter o máximo de informações possíveis, com o método *early stopping*, todos os *backbones* disponibilizados pela FPN foram utilizados.

Os testes foram realizados para cada pré-processamento mencionado na Subseção 2.2. Além disso, também foram feitos testes com as imagens originais da base, sem alteração, em formato RGB. Em seguida, os demais testes foram feitos com a base em tons de cinza,

conforme mostrado na Subseção 2.2.1. Nos próximos testes, foram aplicadas as modificações de Luv(u), ao filtro gaussiano e ao CLAHE, conforme demonstrado na Subseção 2.2.2. Por último, foram realizados os testes com as imagens denominadas por merge em 2.2.3.

## 4.4 Validação

As métricas utilizadas para a validação dos resultados foram, Precisão Média, *precision*, *recall* e *f1-Score* (F1), métricas essas que são usadas recorrentemente na literatura. As métricas serão calculadas com base na matriz de confusão. A matriz de confusão fornece uma hipótese das medidas eficazes do modelo de classificação, mostrando o número de classificações corretas versus as classificações previstas para cada classe em um determinado conjunto de exemplos (CARVALHO et al., 2021).

## 5 Resultados e Discussões

Este capítulo apresenta e discute os resultados alcançados por meio dos experimentos realizados para a detecção de bacilos em imagens de baciloscopia de escarro, utilizando a arquitetura RetinaNet, de acordo com a metodologia descrita no Capítulo 4. O capítulo está dividido em três partes: na primeira, são apresentados os experimentos dos testes realizados com a RetinaNet, na segunda, a comparação entre o método proposto e os trabalhos relacionados e por fim, as discussões.

### 5.1 Resultados do Método Proposto

Esta Seção apresenta os resultados obtidos com a execução da RetinaNet em diversos casos de teste, divididos em quatro experimentos. Os experimentos têm como objetivo encontrar a melhor representação de coloração da imagem que mais se adapte a metodologia do trabalho. A Tabela 2 mostra um resumo dos testes.

Tabela 2 – Experimentos realizados.

| Testes Realizados | Método Utilizado           |
|-------------------|----------------------------|
| Experimento 1     | Imagens originais - RGB    |
| Experimento 2     | Imagens em tons de cinza   |
| Experimento 3     | Luv(u) + gaussiano + CLAHE |
| Experimento 4     | Merge                      |

#### 5.1.1 Experimento 1

A Tabela 3 mostra os resultados da execução do modelo com as imagens no canal de cor RGB. Os testes foram realizados com todos os *backbones* (ResNet-18, ResNet-34, ResNet-50, ResNet-101 ou ResNet-152), os *backbones* desempenham um papel crucial em redes profundas, pois são responsáveis pela extração de características relevantes dos dados e influenciam a eficiência e o desempenho geral do modelo. A escolha do *backbones* adequado para uma tarefa específica pode fazer a diferença entre um modelo que atinge resultados satisfatórios e outro que alcança o estado da arte (DU et al., 2020), por esse motivo, todos eles foram submetidos a teste, com o intuito de encontrar o que alcance melhores resultados.

É apresentado os resultados para os experimentos com as imagens RGB na Tabela 3. Pode-se observar que embora a profundidade do *backbones* fosse aumentando, os resultados obtidos se mantiveram em um intervalo próximo, sendo o teste com o *backbone* ResNet

Tabela 3 – Resultados obtidos com as imagens em RGB.

| Backbones  | mAP  | Pre(%) | Rec(%) | F1(%) |
|------------|------|--------|--------|-------|
| ResNet 18  | 57.8 | 56.7   | 72.4   | 63.5  |
| ResNet 34  | 58.6 | 57.8   | 73.1   | 64.5  |
| ResNet 50  | 56.1 | 61.0   | 70.7   | 65.5  |
| ResNet 101 | 57.5 | 64.9   | 70.4   | 67.5  |
| ResNet 152 | 56.5 | 59.6   | 71.8   | 65.1  |

101, de modo geral, foi o que apresentou os melhores resultados, com um mAP de 57,5%, uma *precision* de 64,9%, um *recall* de 70,4% e um *f1 score* de 67,5%.



Figura 7 – Exemplo de comparação da marcação de especialistas em relação ao modelo em imagem RGB.

Na Figura 7, as marcações em vermelho representam as anotações feitas por especialistas, enquanto as marcações em azul foram geradas automaticamente pelo modelo treinado com imagens RGB. Observando a Figura 7, podemos notar que, embora o modelo não tenha alcançado resultados excepcionais em relação às métricas, ele conseguiu detectar os bacilos em BA na maioria dos casos. Nesse exemplo, os bacilos foram detectados corretamente em quase todas as instâncias, exceto em dois casos em que houve detecções incorretas, consideradas como falsos positivos (FP), uma possível causa pode ser à grande semelhança que os bacilos possuem com outras estruturas não identificadas como bacilos. Com relação aos FNs (Falsos Negativos) na Figura 7, consideramos que a notável similaridade entre os bacilos pode ocasionar penalizações ao modelo quando erroneamente identifica um bacilo sem a devida marcação correspondente. Devido a esse cenário, a rede busca aprender outras características, como orientação ou posição, a fim de mitigar a deficiência no aprendizado das características relevantes para a resolução do problema.





Figura 8 – Exemplo de imagem RGB marcada por especialistas e detecção do modelo.

Também existem casos na qual o modelo não consegue detectar nenhum bacilo, mesmo possuindo bacilos, como mostra na Figura 8.

### 5.1.2 Experimento 2

Já na Tabela 4, temos o resultado com as imagens que passaram pelo pré-processamento, onde foi feita a conversão das imagens RGB para escala de cinza. Nesses testes, como apresentado na Subseção 5.1.1 foram utilizados todos os *backbones* ResNet. Percebeu-se que quanto mais raso fosse os *backbones*, menos a arquitetura conseguiu detectar o bacilo.

Tabela 4 – Resultados obtidos com as imagens *GRAY*.

| <b>Backbones</b> | <b>mAP</b>  | <b>Pre(%)</b> | <b>Rec(%)</b> | <b>F1(%)</b> |
|------------------|-------------|---------------|---------------|--------------|
| ResNet 18        | <b>0.12</b> | 0.1           | <b>7.8</b>    | 3.9          |
| ResNet 34        | 0.0         | 0.2           | 0.3           | 0.2          |
| ResNet 50        | 0.0         | 0.0           | 0.0           | 0.0          |
| ResNet 101       | 22.2        | 4.9           | 62.8          | 9.10         |
| ResNet 152       | 23.5        | 4.9           | 63.4          | 9.10         |

Após analisar a Tabela 4, verificou-se que os resultados obtidos não foram satisfatórios. O modelo não apresentou desempenho promissor em nenhum dos testes realizados. Essa situação pode ser atribuída à perda de informações de cores, as quais desempenham um papel crucial na correta classificação e identificação dos objetos presentes nas imagens. Além disso, a detecção de bacilos pode exigir informações detalhadas e contextuais fornecidas pelas cores presentes nas imagens. A ausência de cores pode dificultar a distinção de objetos e padrões relevantes. Para aprimorar o desempenho, é fundamental explorar diferentes abordagens e técnicas, conforme abordado neste trabalho.

Seguindo a mesma lógica apresentada na Figura 7, a Figura 9 apresenta as marcações vermelhas realizadas por profissionais e as azuis feitas pelo modelo. Pode-se analisar que, nesse exemplo, apenas três bacilos foram identificados, e a precisão de certeza de se é realmente um bacilo ou não, é bem baixo, logo, o modelo não obteve nenhum tipo de aprendizado significativo.

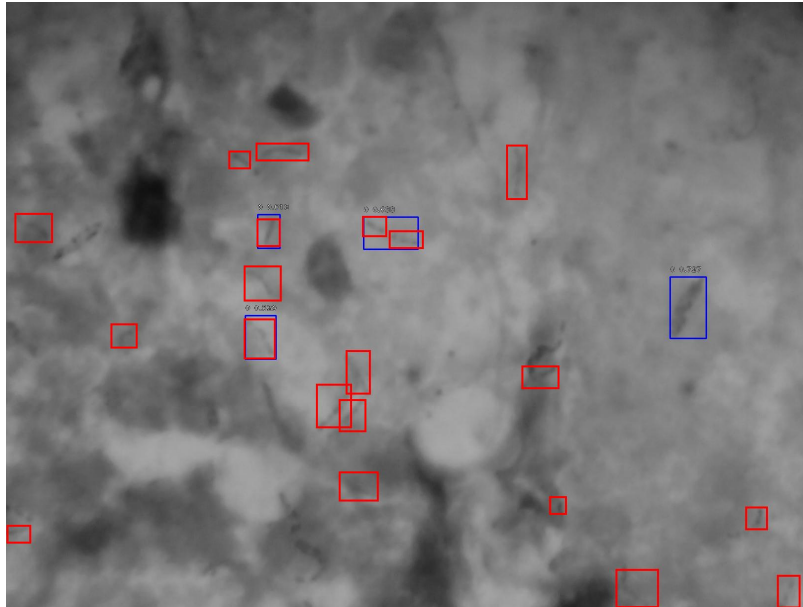


Figura 9 – Imagem em tons de cinza com comparação da marcações de especialista e detecção do modelo.

### 5.1.3 Experimento 3

O terceiro experimento apresenta os resultados das imagens que foram submetidas ao pré-processamento, onde ocorreu a conversão das imagens de RGB para o espaço de cores LUV. Nesses testes como nos demais, todos os *backbones* ResNet foram utilizados.

Tabela 5 – Resultados obtidos com as imagens no canal Luv(u) + gaussiano + CLAHE.

| Backbones  | mAP  | Pre(%) | Rec(%) | F1(%) |
|------------|------|--------|--------|-------|
| Res Net 18 | 22.3 | 15.6   | 43.2   | 22.9  |
| ResNet 34  | 23.4 | 12.3   | 45.6   | 19.3  |
| ResNet 50  | 27.5 | 13.7   | 44.5   | 21    |
| ResNet 101 | 35.6 | 34.4   | 54.6   | 42.2  |
| ResNet 152 | 37.6 | 33.4   | 64.6   | 44.0  |

Conforme observado na Tabela 5, os melhores resultados foram referentes ao *backbones* (ResNet-152), com um mAP de 37,6%, uma *precision* de 33,4%, um *recall* de 64,6% e um *f1 score* de 44%. Ao analisar os resultados, verificamos que, em comparação com as imagens no espaço de cores RGB, a utilização do espaço de cores LUV não demonstrou capacidade de fornecer as características adequadas para um aprendizado eficiente para o problema composto, por mais que o resultado do *recall* uma metrica promissora para o backbone ResNet 152, a *precision* é bastante baixa, tornando o modelo inviável para detecção de bacilos.

A Figura 10 apresenta uma clara demonstração do desequilíbrio entre *precision* e *recall*. Nela, três setas indicam os bacilos detectados pelo modelo, no entanto, uma análise da marcação pré-definida revela que esses bacilos identificados pelo modelo são, na verdade,

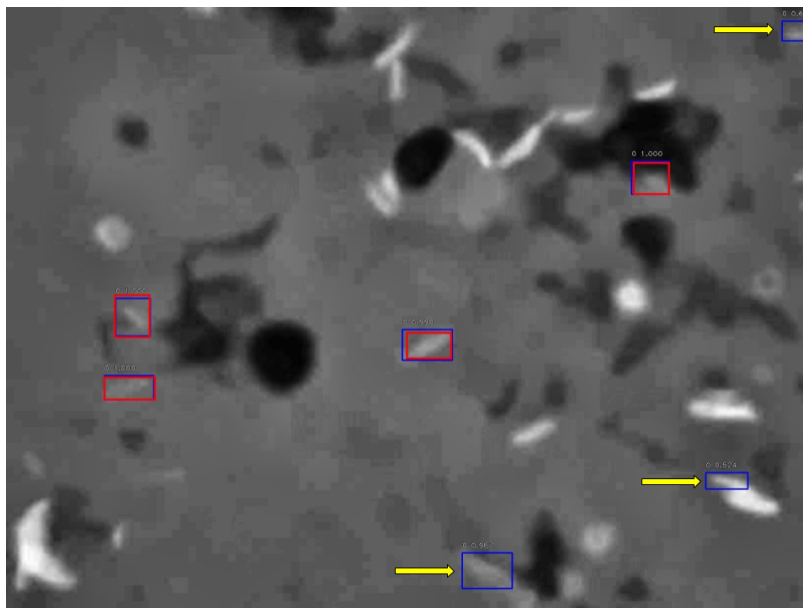


Figura 10 – Exemplo de comparação da marcações de especialistas e do modelo.

falsos positivos. Esse cenário evidencia a necessidade de aprimorar a precisão da detecção para evitar tais falsos positivos.

#### 5.1.4 Experimento 4

Por fim, a Tabela 6 mostra os resultado do último experimento realizado nesse trabalho, com as imagens que passaram pelo pré-processamento, onde foram realizados os treinamentos com o "merge", junção de 3 cópias da mesma imagem com os procedimentos. Nesses testes, foram utilizados os *backbones* (50,101 e 152) ResNet.

Tabela 6 – Resultados obtido das imagens com o merge

| Backbones  | mAP  | Pre(%) | Rec(%) | F1(%) |
|------------|------|--------|--------|-------|
| ResNet 50  | 46,3 | 38,9   | 66,7   | 49,1  |
| ResNet 101 | 45,2 | 54,6   | 59,5   | 56,9  |
| ResNet 152 | 45,9 | 58,4   | 60,4   | 59,4  |

Em relação aos demais pré-processamentos usados para alcançar os objetivos do método proposto, esse foi o melhor, mostrando resultados promissores, mas não maior do que o treinamento realizado com as imagens originais, sem alteração alguma. O modelo mais bem-sucedido foi construído com o uso do *backbone* ResNet-152, alcançando uma *precision* de 58,4% e uma taxa de *recall* de 60,4%, esses resultados promissores sugerem que o modelo tem o potencial de ser útil para a detecção de bacilos, mas ainda há margem para otimização. É importante destacar que a jornada para desenvolver modelos de detecção de objetos eficientes é iterativa e requer um processo contínuo de pesquisa e desenvolvimento.

Tabela 7 – Comparação entre os trabalhos relacionados e a abordagem proposta.

| Trabalho                       | Método   | Amostras | mAP (%) | Pre (%) | Rec (%) | F1 (%) |
|--------------------------------|--|----------|---------|---------|---------|--------|
| Gräbel et al. (2020)           | RetinaNet  | 101      |         |         |         | 92,5%  |
| Pho, Amin e Yoshitaka (2018)   | RetinaNet  | 69       | 53%     | 48%     | 94%     | 63,54% |
| Kant e Srivastava (2018)       | Rede neural simples + estratégia de patchwise                                | 202      |         | 67,55%  | 83,78%  | 74,7%  |
| Santos, Brito e Filho (2023)   | Pré-processamento + detecção de bacilos usando EfficientDet e seus backbones | 1265     |         | 69,4%   | 92,5%   | 77,4%  |
| El-Melegy et al. (2019)        | classificador CNN  | 500      |         | 85,1%   | 98,4%   | 91,2%  |
| Swetha e Sankaragomathi (2020) | classificador CNN  |          |         |         |         |        |
| Método Proposto                | RetinaNet  | 1218     | 57,5%   | 64,9%   | 70,4%   | 67,6%  |

## 5.2 Comparação com os Trabalhos Relacionados

Os melhores resultados obtidos a partir da metodologia proposta foram comparados com os estudos relacionados. Comparar esses trabalhos não é uma tarefa trivial, pois muitos fatores, como bases de imagens, métricas calculadas e métodos influenciam diretamente uma comparação justa.

A Tabela 7 apresenta uma síntese comparativa entre os resultados obtidos neste estudo e as pesquisas mencionadas no Capítulo 3. Os trabalhos que inicialmente não possuíam o valor de F1-score, mas contavam com *precision* e *recall*, tiveram esse valor adicionado de forma manual. O *f1-score* é um resultado calculado exclusivamente com base nessas duas variáveis, e sua inclusão foi realizada como uma medida para padronizar a tabela.

Ao analisar a Tabela 7, verificou-se que, entre todos os trabalhos listados, apenas o estudo de Pho, Amin e Yoshitaka (2018) apresentou resultados inferiores, apenas o *recall* está superior, porém com a *precision* em 48%, não mostra um modelo confiável. O resultado alcançado pelo método proposto apresentou uma notável semelhança com o estudo de Santos, Brito e Filho (2023), essa comparação é considerada justa, uma vez que ambos utilizaram a mesma base de dados como referência. Aos demais, foram usadas poucas imagens como base de dados para os treinamentos, podendo prejudicar seriamente o desempenho e a capacidade de generalização de uma rede neural, tornando-a menos confiável.

Em geral, pode-se afirmar que o método apresenta grande promessa, alcançando resultados próximos ao estado da arte, mesmo para problemas complexos. Além disso, a eficiência dos resultados obtidos é notável. No entanto, a otimização dos hiperparâmetros e a implementação de um pré-processamento mais robusto nas imagens são consideradas medidas essenciais para potencializar significativamente os resultados alcançados.

## 5.3 Discussão

Neste estudo, foi desenvolvido um método automatizado para a detecção de bacilos em imagens de BA. Com base nos resultados detalhados na Seção 5.1 acerca da arquitetura RetinaNet e sua integração com procedimentos de pré-processamentos específicos,

observou-se que os resultados obtidos foram encorajadores. Contudo, há oportunidades para realizar ajustes adicionais com o objetivo de aprimorar ainda mais a distinção entre os bacilos e o fundo das imagens. A seguir, destacam-se alguns pontos essenciais sobre este trabalho:

#### 1. Contribuições da metodologia proposta

- A aplicação inovadora da RetinaNet para detecção de bacilos em BA;
- Observou-se que, de forma geral, a *precision* aumenta à medida que se utiliza um *backbone* mais profundo.
- A Função *Focal Loss*, utilizada no RetinaNet, mostrou ser eficiente para que o equilíbrio entre as classes seja essencial para um resultado mais eficaz;
- Após a análise das diversas propostas de canais de cores testadas, constatou-se que a abordagem RGB foi a mais eficaz.;
- O método produz caixas delimitadoras contendo o bacilo, o que pode facilitar nas etapas seguintes da contagem do bacilo, já que a região a ser analisada é menos suscetível ao erro, mesmo por técnicas simples como a binarização;
- A abordagem apresenta resultados similares ao estado da arte, sendo validado em um conjunto de dados maior que os apresentados nos trabalhos relacionados, apresentando resultados mais confiáveis e uma capacidade de generalização maior.

#### 2. Limitações da metodologia proposta

- Ao analisar os resultados, foi percebido que no esquema de cores LUV, a arquitetura da RetinaNet não conseguiu detectar muitos bacilos. Acreditamos que ao realçamos as imagem, além do bacilo muitas regiões onde não se havia nada, poderiam começar a ser vista como regiões candidatas, dificultando o aprendizado da rede;
- conforme os resultados, alguns bacilos grandes não foram corretamente detectados. Consideramos que isso pode ser solucionando com um aumento no número de imagens para treinamento que contenham bacilos maiores e ajustes na parametrização.

## 6 Conclusão

Este trabalho propôs a aplicação da arquitetura RetinaNet para a detecção de bacilos álcool-ácido-resistentes (BAAR) em imagens de baciloscopia de escarro. Na abordagem proposta, utilizou-se uma única base de dados contendo as baciloscopias, que passaram por uma etapa de pré-processamento com transformações de canais de coloração e aplicação de filtros, com o objetivo de realçar os bacilos em relação ao fundo da imagem. Os resultados obtidos com as imagens originais no espaço de cores RGB foram consistentes e equilibrados em todas as métricas de validação. Esse enfoque apresenta potencial para aprimorar a detecção de bacilos.

A metodologia proposta apresentou resultados promissores, apesar de ser um problema complexo. Além disso, é importante destacar que esta arquitetura ainda não possui aplicações para este tipo de problema na literatura. A detecção da caixa delimitadora de bacilos pode ser relevante para etapas seguintes no processo de identificação do BAAR, como a contagem de bacilos e classificação da doença. Ainda que existam limitações, o modelo obtido conseguiu atingir uma boa generalização, mostrando que com alguns aperfeiçoamentos, a abordagem aqui proposta possui a capacidade de auxiliar os especialistas tanto no diagnóstico como no monitoramento da TB, em relação ao tratamento, acreditasse que o método é robusto e promissor, tornando este método bastante relevante para situações reais do mundo atual.

Os resultados obtidos demonstraram robustez, porém foi identificado oportunidades para aprimorar o método de detecção. Nesse contexto, com o objetivo de avançar no estudo e aprimorar os resultados obtidos, foram delineadas três propostas para trabalhos futuros :i) A utilização do aumento de dados, mostra-se fundamental para o desenvolvimento de um modelo de detecção de alta performance. Ao empregar técnicas de aumento de dados, como rotação, espelhamento pode expandir a diversidade do conjunto de treinamento, mitigar o risco de *overfitting* e permitir que o modelo generalize melhor para cenários diversos. ii) Pretende-se avaliar o modelo com esquemas de cores e filtros diversos, para obter informações valiosas sobre a flexibilidade e confiabilidade do modelo em situações reais, o que é fundamental para garantir que ele possa ser aplicado com eficácia em uma variedade de cenários e ambientes. E iii) Realizar o treinamento da RetinaNet em uma base de dados diferente, que é uma estratégia para garantir a robustez, a eficácia e a adaptabilidade do modelo em aplicações reais. Essa abordagem permite explorar um espectro mais amplo de dados e cenários, preparando o modelo para enfrentar desafios diversos e fornecer resultados mais precisos e confiáveis em tarefas de detecção de objetos.

## 7 Publicações

O método exposto neste estudo alcançou reconhecimento em publicações de âmbito internacional, sendo devidamente reconhecido e citado em 6 diferentes estudos da área:

- F. J. Dos Santos Reist et al., “BacillusNet: An automated approach using RetinaNet for segmentation of pulmonary Tuberculosis bacillus,” in Proceedings - IEEE Symposium on Computers and Communications, 2021, vol. 2021-Septe, pp. 8–11, doi: 10.1109/ISCC53001.2021.9631390.

Além desta, o método possui outra publicação, com título e revista respectivamente:

- RODRIGUES, Filipe M. M.; REIS, Francisco J. S.; VELOSO, Mateus A.; DINIZ, João O. B.; VELOSO, Romuere R.; C. FILHO, Antonio O.. Metodologia automática para detecção de bacilos de tuberculose utilizando RetinaNet e modelos de cores. In: SIMPÓSIO BRASILEIRO DE COMPUTAÇÃO APLICADA À SAÚDE (SBCAS), 22. , 2022, Teresina. Anais [...]. Porto Alegre: Sociedade Brasileira de Computação, 2022 . p. 334-345. ISSN 2763-8952. DOI: <https://doi.org/10.5753/sbcas.2022.222677>.

# Referências

- ANALITICA, C. *Entenda o que é Deep Learning e Como Funciona*. 2018. Disponível em: <https://culturaanalitica.com.br/deep-learning-oquee-como-funciona/>. Acessado em 28 de junho de 2023. Citado na página 20.
- ARCGIS. *How RetinaNet works?* 2023. Disponível em: <https://developers.arcgis.com/python/guide/how-retinanet-works/>. Acessado em 10 de julho de 2023. Citado na página 22.
- ARNOLD, L. et al. An Introduction to Deep Learning. In: *European Symposium on Artificial Neural Networks (ESANN)*. Bruges, Belgium: [s.n.], 2011. (Proceedings of the European Symposium on Artificial Neural Networks (ESANN)). Disponível em: <<https://hal.science/hal-01352061>>. Citado na página 20.
- BENGIO, Y. et al. Learning deep architectures for ai. *Foundations and trends® in Machine Learning*, Now Publishers, Inc., v. 2, n. 1, p. 1–127, 2009. Citado na página 19.
- CAMPOS, H. S. Diagnóstico da tuberculose. *Pulmão RJ*, v. 15, n. 2, p. 92–99, 2006. Citado 2 vezes nas páginas 12 e 15.
- CARVALHO, E. D. et al. An approach to the classification of covid-19 based on ct scans using convolutional features and genetic algorithms. *Computers in Biology and Medicine*, v. 136, p. 104744, 2021. ISSN 0010-4825. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0010482521005382>>. Citado na página 29.
- CHAGAS, E. T. D. O. *DEEP LEARNING E SUAS APLICAÇÕES NA ATUALIDADE*. 2019. Disponível em: <https://www.nucleodoconhecimento.com.br/administracao/deep-learning>. Acessado em 26 de junho de 2023. Citado na página 19.
- COSTA, M. G. et al. Automatic identification of mycobacterium tuberculosis with conventional light microscopy. In: IEEE. *2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. [S.l.], 2008. p. 382–385. Citado na página 16.
- CRUZ., F. O. *Tuberculose*. 2023. Disponível em: <https://portal.fiocruz.br/doenca/tuberculose>. Acessado em 25 de junho de 2023. Citado na página 15.
- DEVARAKONDA, A.; NAUMOV, M.; GARLAND, M. Adabatch: Adaptive batch sizes for training deep neural networks. *arXiv preprint arXiv:1712.02029*, 2017. Citado na página 28.
- DU, X. et al. Spinenet: Learning scale-permuted backbone for recognition and localization. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. [S.l.: s.n.], 2020. p. 11592–11601. Citado na página 30.
- DUNNINGTON, G. W.; GRAY, J.; DOHSE, F.-E. *Carl Friedrich Gauss: titan of science*. [S.l.]: MAA, 2004. Citado na página 18.



EL-MELEGY, M. et al. Identification of tuberculosis bacilli in zn-stained sputum smear images: A deep learning approach. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*. [S.l.: s.n.], 2019. p. 0–0. Citado 3 vezes nas páginas 24, 25 e 35.

FERRI, A. O. et al. Diagnóstico da tuberculose: uma revisão. *Revista Liberato*, v. 15, n. 24, p. 145–154, 2014. Citado na página 12.

FURG. *DAS alerta sobre a importância do diagnóstico precoce e tratamento adequado da tuberculose*. 2019. Disponível em: <https://www.furg.br/avisos/das-alerta-sobre-a-importancia-do-diagnostico-precoce-e-tratamento-adequado-da-tub>. Acessado em 27 de junho de 2023. Citado na página 12.

GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. *Deep learning*. [S.l.]: MIT press, 2016. Citado na página 19.

GRÄBEL, P. et al. Circular anchors for the detection of hematopoietic cells using retinanet. In: IEEE. *2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI)*. [S.l.], 2020. p. 249–253. Citado 3 vezes nas páginas 24, 25 e 35.

GRACE, K. et al. When will ai exceed human performance? evidence from ai experts. *Journal of Artificial Intelligence Research*, v. 62, p. 729–754, 2018. Citado na página 19.

GUERRA, R. L.; REGO, L.; CONDE, M. B. Diagnóstico da tuberculose pulmonar em pacientes com baciloscopia direta de escarro negativa. *Pulmão RJ* 2008; 17 (2-4): 87-90, 2008. Citado na página 12.

HULSTAERT, L. *A Beginner's Guide to Object Detection*. 2018. Disponível em: <https://www.datacamp.com/tutorial/object-detection-guide>. Citado na página 23.

HUMBARWADI, S. *Object Detection with RetinaNet*. 2020. Disponível em: <https://keras.io/examples/vision/retinanet/>. Acessado em 15 de abril de 2022. Citado na página 22.

JIAO, L. et al. A survey of deep learning-based object detection. *IEEE Access*, v. 7, p. 128837–128868, 2019. Citado na página 13.

KAGGLE. *Tuberculosis Image Dataset*. 2020. Disponível em: <https://www.kaggle.com/saife245/tuberculosis-image-datasets>. Accessed June 05, 2021. Citado 6 vezes nas páginas 8, 16, 18, 19, 24 e 27.

KANT, S.; SRIVASTAVA, M. M. Towards automated tuberculosis detection using deep learning. In: *2018 IEEE Symposium Series on Computational Intelligence (SSCI)*. [S.l.: s.n.], 2018. p. 1250–1253. Citado 3 vezes nas páginas 24, 25 e 35.

KARAGIANNAKOS. *Localization and Object Detection with Deep Learning*. 2019. Disponível em: [https://sergioskar.github.io/Localization\\_and\\_Object\\_Detection](https://sergioskar.github.io/Localization_and_Object_Detection). Citado na página 23.

LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. *nature*, Nature Publishing Group UK London, v. 521, n. 7553, p. 436–444, 2015. Citado na página 13.

- LIN, T.-Y. et al. Feature pyramid networks for object detection. In: *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. [S.l.: s.n.], 2017. p. 936–944. Citado 3 vezes nas páginas 21, 22 e 28.
- LIN, T.-Y. et al. Focal loss for dense object detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 42, n. 2, p. 318–327, 2020. Citado 4 vezes nas páginas 8, 13, 21 e 22.
- LU, C.; XU, L.; JIA, J. Contrast preserving decolorization with perception-based quality metrics. *International journal of computer vision*, Springer, v. 110, p. 222–239, 2014. Citado na página 17.
- METZ, C. E. Basic principles of roc analysis. In: ELSEVIER. *Seminars in nuclear medicine*. [S.l.], 1978. v. 8, n. 4, p. 283–298. Citado na página 23.
- MIOTTO, R. et al. Deep learning for healthcare: review, opportunities and challenges. *Briefings in bioinformatics*, Oxford University Press, v. 19, n. 6, p. 1236–1246, 2018. Citado na página 13.
- NOGUEIRA, A. F. et al. Tuberculose: uma abordagem geral dos principais aspectos. *Rev. Bras. Farm*, v. 93, n. 1, p. 3–9, 2012. Citado na página 15.
- PACHECO, C. A. R.; PEREIRA, N. S. Deep learning conceitos e utilização nas diversas áreas do conhecimento. *Revista Ada Lovelace*, v. 2, p. 34–49, 2018. Citado na página 19.
- PHO, K.; AMIN, M. K. M.; YOSHITAKA, A. Segmentation-driven retinanet for protozoa detection. In: *2018 IEEE International Symposium on Multimedia (ISM)*. [S.l.: s.n.], 2018. p. 279–286. Citado 3 vezes nas páginas 24, 25 e 35.
- RADAR. *Como funcionam as redes neurais de Deep Learning*. 2022. Disponível em: <https://ilumeo.com.br/todos-posts/2020/09/13/como-funcionam-as-redes-neurais-de-deep-learning>. Acessado em 28 de junho de 2023. Citado na página 20.
- REZA, A. M. Realization of the contrast limited adaptive histogram equalization (clahe) for real-time image enhancement. *Journal of VLSI signal processing systems for signal, image and video technology*, Springer, v. 38, n. 1, p. 35–44, 2004. Citado na página 18.
- RUSS, J. C. *The image processing handbook*. [S.l.]: CRC press, 2006. Citado 2 vezes nas páginas 17 e 27.
- SANTOS, P. R. S dos; BRITO, V. de C.; FILHO, A. O. de C. Kochdet: Arquitetura profunda baseada em bifpn para detecção de bacilos de kock. *Revista de Sistemas e Computação-RSC*, v. 13, n. 1, 2023. Citado 3 vezes nas páginas 24, 25 e 35.
- SAÚDE., B. M. da. *Tuberculose*. 2022. Disponível em: <https://www.gov.br/saude/pt-br/assuntos/saude-de-a-a-z/t/tuberculose>. Acessado em 15 de fevereiro de 2022. Citado na página 12.
- SESA. *Tuberculose*. 2022. Disponível em: <https://saude.es.gov.br/neve-tuberculose>. Acessado em 25 de junho de 2023. Citado na página 15.

SHEEDY, J. E. *Field Guide to Visual and Ophthalmic Optics*. [S.l.]: LWW, 2005. Citado na página 18.

SINHA, S. et al. Epoch determination for neural network by self-organized map (som). *Computational Geosciences*, Springer, v. 14, p. 199–206, 2010. Citado na página 28.

SOTGIU, G. et al. *The ERS-endorsed official ATS/CDC/IDSA clinical practice guidelines on treatment of drug-susceptible tuberculosis*. [S.l.]: Eur Respiratory Soc, 2016. 963–971 p. Citado na página 12.

SWETHA, K.; SANKARAGOMATHI. Convolutional neural network based automated detection of mycobacterium bacillus from sputum images. In: *2020 International Conference on Inventive Computation Technologies (ICICT)*. [S.l.: s.n.], 2020. p. 293–300. Citado 2 vezes nas páginas 25 e 35.

SZELISKI, R. *Computer vision: algorithms and applications*. [S.l.]: Springer Nature, 2022. Citado na página 17.

TANSUPHAISIRI, U.; KLADPHUANG, B. Evaluation of sputum staining by modified cold method and comparison with ziehl-neelsen and fluorochrome methods for the primary diagnosis of tuberculosis. *Southeast Asian Journal of Tropical Medicine and Public Health*, Citeseer, v. 33, n. 1, p. 128–135, 2002. Citado na página 17.

WHO, W. H. O. *Tuberculosis*. 2023. Disponível em: <https://www.who.int/en/news-room/fact-sheets/detail/tuberculosis>. Acessado em 27 de junho de 2023. Citado 3 vezes nas páginas 12, 15 e 16.

YIU, T. *Understanding Neural Networks*. 2019. Disponível em: <https://towardsdatascience.com/understanding-neural-networks-19020b758230>. Acessado em 27 de fevereiro de 2022. Citado 2 vezes nas páginas 8 e 20.

YOHANANDAN, S. map (mean average precision) might confuse you. *Towards Data Science*, v. 9, 2020. Citado na página 22.



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

**Identificação do Tipo de Documento**

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

Eu, **Filipe Mateus Moraes Rodrigues**, autorizo com base na Lei Federal nº 9.610 de 19 de Fevereiro de 1998 e na Lei nº 10.973 de 02 de dezembro de 2004, a biblioteca da Universidade Federal do Piauí a divulgar, gratuitamente, sem ressarcimento de direitos autorais, o texto integral da publicação “**Método para detecção de bacilos de tuberculose usando RetinaNet**” de minha autoria, em formato PDF, para fins de leitura e/ou impressão, pela internet a título de divulgação da produção científica gerada pela Universidade.

Picos-PI 20 de Agosto de 2023.

*Filipe Mateus M. Rodrigues*

Assinatura