

# Metodologia automática para detecção de bacilos de tuberculose utilizando RetinaNet e modelos de cores

Filipe M. M. Rodrigues<sup>1</sup>, Francisco J. S. Reis<sup>1</sup>, Mateus A. Veloso<sup>1</sup>,  
João O. B. Diniz<sup>4</sup>, Romuere R. Veloso<sup>1,2,3</sup>, Antonio O. C. Filho<sup>1,2,3</sup>

<sup>1</sup> Departamento de Sistemas de Informação - Universidade Federal do Piauí - Picos, Brasil.

<sup>2</sup> Programa de Pós-graduação em Engenharia Elétrica e Ciência da Computação - Universidade Federal do Piauí - Teresina, Brasil.

<sup>3</sup> Programa de Pós-Graduação em Ciência da Computação associação UFPI - Teresina / UFMA - São Luís, Brasil.

<sup>4</sup> Instituto Federal do Ciência e Tecnologia do Maranhão - Grajaú, Brasil

{filipemts, franciscojose17, mateusveloso, antoniooseas, romuere}@ufpi.edu.br,  
joao.bandeira@ifma.edu.br

**Abstract.** Tuberculosis is the most deadly bacterial infectious disease globally, with about 1.5 million people dying every year. The disease is caused by *Mycobacterium tuberculosis*. The main form of diagnosis is sputum bacilloscopy, an exam in which the patient's sputum is analyzed under a microscope searching for bacillus, making this a technique for both diagnosis and monitoring the disease. Therefore, this work aims to develop a methodology for automated detection of the bacillus using RetinaNet. A set of 1218 images was used to evaluate this method. The results were encouraging, with an accuracy of 64.9%, a recall of 70.4%, and an F1 score of 61%. Finally, we believe that our method can act in diagnosing tuberculosis.

**Resumo.** A tuberculose é a doença infecciosa bacteriana que mais mata no mundo, cerca de 1,5 milhão de pessoas morreram todos anos. A doença é causada pelo *Mycobacterium tuberculosis*, 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. Diante disso, este trabalho tem por objetivo desenvolver uma metodologia para detecção automatizada do bacilo usando a RetinaNet. Um conjunto de 1218 imagens foi usado para a avaliação deste método. Os resultados foram animadores, obtendo uma precisão de 64,9%, recall de 70,4% e um score F1 de 61%. Finalmente, acreditamos que nosso método possui a capacidade de atuar no diagnóstico da tuberculose.

## 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 pulmo-

nar, 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 (da Saúde, 2022), ela é causada pelo *Mycobacterium tuberculosis* ou bacilo de Koch, sendo uma doença evitável e curável.

Em todo o mundo, a tuberculose é a 13ª principal causa de morte e a segunda principal causa de morte infecciosa após o COVID-19 (acima do HIV/AIDS) (WHO, 2022). Em 2020, um total de 1,5 milhão de pessoas morreram de tuberculose no mundo, enquanto cerca de 10 milhões sofreram com a doença, destes, 1,1 milhão eram de crianças, sendo a TB infantil a forma mais difícil de diagnosticar e tratar. Ocorrendo em 95% dos casos em países em desenvolvimento, a TB atinge na maioria dos casos adultos em seus anos de maior produtividade. 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.

Os exames mais utilizados para o diagnóstico da tuberculose são o de baciloscopia de escarro (BA), cultura para micobactéria, teste rápido para tuberculose e o de radiografia de tórax. Sendo a BA o método mais amplamente utilizado tanto para o diagnóstico como 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). Desde que executada corretamente em todas as suas fases, permite detectar de 60% a 80% dos casos, com resultado em até 48 horas (SESA, 2022). Ainda hoje esse exame é feito de maneira manual, através da inspeção humana, em um processo longo e cansativo, muitas vezes suscetível a falha. Uma maneira de agilizar esse processo o tornando menos extenuante para o profissional responsável pela análise, seria a automação do processo de detecção do bacilo da tuberculose.

Estima-se que 66 milhões de vidas foram salvas através do diagnóstico e tratamento da TB entre 2000 e 2020 (WHO, 2020). Logo, o diagnóstico da tuberculose é de suma importância para o tratamento e a interrupção da cadeia de transmissão da doença, dito isso algumas técnicas computadorizadas vem sendo desenvolvidas para detecção do bacilo da tuberculose, entre elas está uma sub-área da inteligência artificial, o aprendizado profundo. Contudo, o maior obstáculo para detecção de bacilos está na dificuldade de delimitar o bacilo da imagem, uma vez que eles não possuem tamanho uniforme, além da existência de outros artefatos que dificultam o desempenho da arquitetura. Em nosso trabalho, usamos a RetinaNet para a tarefa de detecção do BAAR usando imagens de baciloscopia.

A arquitetura da RetinaNet 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, optamos por usar RetinaNet porque usa uma função Focal Loss proposta por (Lin et al., 2020). 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. Finalmente, como mostrado em (Jiao et al., 2019) RetinaNet pode detectar objetos pequenos e médios em uma imagem com uma superioridade relativa sobre outras arquiteturas.

Assim, acreditamos ter contribuído nos seguintes aspectos:

- Construção de uma abordagem promissora que aplica RetinaNet no problema da TB;
- Fornecimento de uma técnica automática para detecção de bacilos em imagens de BA;
- Emprega uma metodologia ágil para detecção automática do bacilo, podendo ser usada de modo a auxiliar o diagnóstico da doença.

O restante do artigo está organizado da seguinte forma: na Seção 2, discutimos trabalhos relacionados; na seção 3, apresentamos a abordagem proposta para a detecção do bacilo; na seção 4, apresentamos os resultados obtidos; na seção 5, discutimos os aspectos positivos e algumas limitações encontradas em nosso trabalho. Finalmente, na Seção 6, apresentamos a conclusão.

## 2. Trabalhos Relacionados

(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 conjunto de 101 patches de imagens

(Pho et al., 2018) usa RetinaNet para detectar, segmentar e identificar automaticamente espécies de protozoários em micrografias. Os autores usam RetinaNet 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 and Srivastana, 2018) faz uso de uma arquitetura de rede neural simples, usa uma estratégia de patchwise para a detecção de bacilos. E seu método usa um campo de visão do microscópio, pedaços de 20px por 20px da imagem de entrada. Os resultados obtidos com as métricas de recall e precisão foram de e 83,78% e 67,55% respectivamente para detecção de bacilos.

(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 (RPN). O modelo foi avaliado com 500 imagens, corada com método de ziehl neelsen, contendo 2000 objetos de treinamento, provenientes da base de imagens pública ZNSM-iDB que consiste em amostras de várias categorias.

(Swetha et al., 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. Os resultados obtidos na classificação das categorias foram de 99,5% de precisão, 94,7% de sensibilidade e 99% de especificidade.

Como podemos ver em na Secção 2, alguns trabalhos na literatura realizam a segmentação dos bacilos, mas não realizam a contagem de quantos bacilos foram detectados, resultado que pode ser alcançado com uso da RetinaNet. Além disso, acreditamos que a proposta do RetinaNet com âncoras e o uso do Focal Loss pode levar a resultados mais eficientes.

### 3. Materiais e Métodos

Nesta seção serão apresentadas as etapas desenvolvidas para a detecção automática da TB em lâminas de esfregaço de escarro usando RetinaNet. Nosso método pode ser resumido nas seguintes etapas: i) aquisição de imagens de TB; ii) foi aplicado um pré-processamento nas imagens para destacar o bacilo do fundo; iii) em seguida, realizou-se o treinamento da arquitetura RetinaNet para detectar bacilo; e, finalmente, iv) foi feita a avaliação dos resultados obtidos. A Figura 1 apresenta um resumo destas etapas.

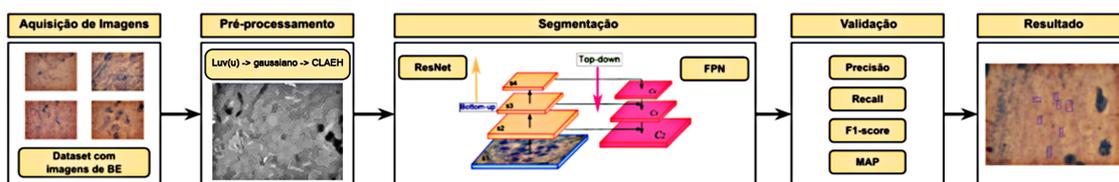


Figura 1. Método proposto.

#### 3.1. Aquisição de imagens

Este trabalho utiliza imagens de TB disponibilizadas pelo AI Research and Automated Laboratory Diagnostics através do kaggle, para fins de pesquisa e educação. Conjunto contém 1218 imagens de escarro com caixas delimitadoras de 3734 bacilos (Kaggle, 2020). Às imagens possuem uma resolução de  $1224 \times 1623$ , e cada imagem está relacionada a um arquivo XML, que contém detalhes das coordenadas de cada caixa delimitadora dos bacilos presentes na imagem. A Figura 2 apresenta um exemplo de imagens contendo as marcações dos bacilos.

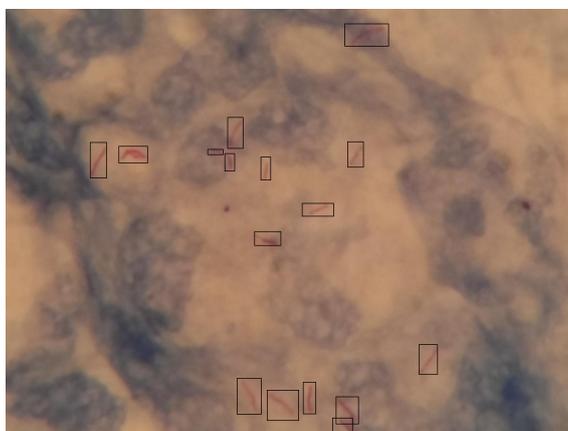
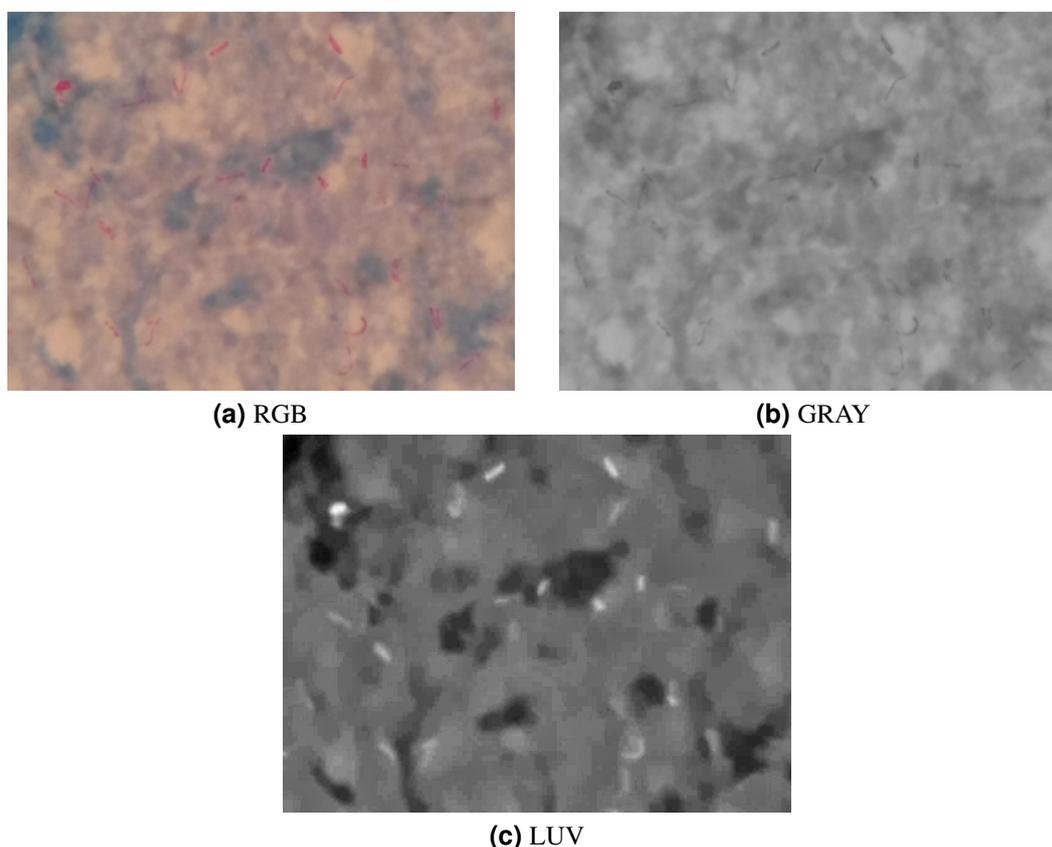


Figura 2. Exemplo de uma imagem com caixa delimitadora

### 3.2. Pré-processamento

A técnica de coloração de Ziehl-Neelsen 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, devido a isso na etapa de pre-processamento foi feita duas conversões nos canais de cores da imagem RGB (Red-Green-Blue). O primeiro deles foi uma conversão muito usado em aplicações de processamento de imagem, cujo processo consiste em converter uma imagem colorida para uma imagem em tons de cinza, conforme visto na Figura 3. Já o segundo melhoramento, trata-se de uma conversão da imagem RGB para o espaço de cores LUV, onde L representa a luminância, U e V representam as informações de cor (crominância). Em seguida, aplicamos um filtro gaussiano (com elemento estruturando igual a 3) para redução do ruído na imagem, para então ser aplicado uma técnica de equalização de histograma adaptativo limitado por contraste (CLAHE) (Reza, 2004) . A Figura 3 apresenta um resultado para estes procedimentos.



**Figura 3. Exemplos de (a) imagem original em RGB; (b) imagem após a conversão de RGB para níveis de cinza; (c) imagem resultante após a conversão para o espaço de cores LUV, após a utilização do filtro Gaussiano com CLAHE.**

### 3.3. Segmentação

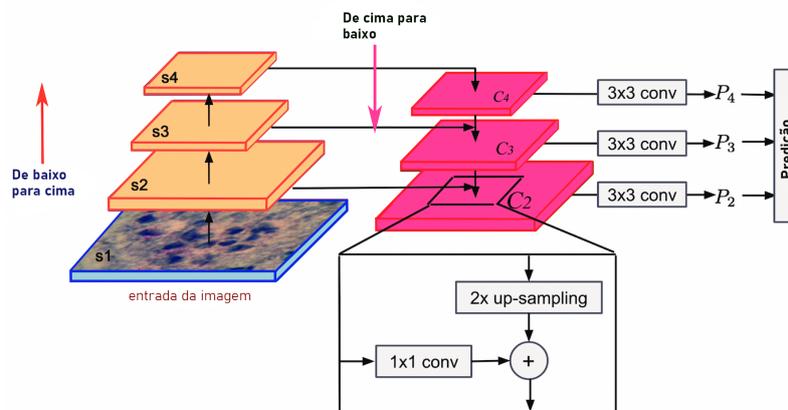
Para a tarefa de detecção dos bacilos, optamos por utilizar a RetinaNet. Essa arquitetura foi proposto por (Lin et al., 2017) em um estudo para detecção de imagens em diferentes escalas. De acordo com (Lin et al., 2017) , o objetivo desta rede é conduzir a hierarquia

de características de uma rede neural convolucional (CNN) enquanto cria uma pirâmide de características com 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, RetinaNet pode criar uma pirâmide de atributos na rede sem abrir mão do poder de representação, velocidade de treinamento ou memória.

A arquitetura RetinaNet consiste em uma rede de backbone chamada Feature Pyramid Network, 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). Neste trabalho, utilizamos a versão do ResNet-152 por ser a mais profunda e ter proporcionado os melhores resultados nos experimentos realizados.

Para predição, 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.

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 sigmoide, 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).



**Figura 4. Arquitetura da RetinaNet**

### 3.3.1. Treinamento

Neste trabalho, seguimos 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 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 quantidade de épocas foi variando no decorrer dos testes, até se obter o melhor cenário para o modelo. A cada época, foram realizadas 437 iterações para obter o máximo de informações possível da imagem.

### 3.4. Validação

Para validar os resultados, foram utilizadas métricas de avaliação estatística, como precisão (Pre), Recall (Rec), F1-Score (F1) e Precisão Média (MAP). 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.

## 4. Resultados

Nessa Seção serão analisados os resultados da detecção de bacilos a partir das imagens de baciloscopia descritas na Seção 3.1.

Os resultados foram avaliados considerando as imagens em: i) RGB; ii) níveis de cinza; e iii) o canal U do espaço de cores LUV em conjunto com filtro gaussiano e CLAHE. Optamos por utilizar apenas o canal U devido a boa qualidade de contraste apresentado por este canal para os bacilos. Com intuito de investigar a capacidade da RetinaNet, nós avaliamos também diferentes variações de backbones da ResNet (Target et al., 2016), no intuito de ampliar as possibilidades de sucesso na detecção dos bacilos e melhoria nos resultados.

Na Tabela 1, apresentamos os resultados para os experimentos com imagens RGB com a arquitetura proposta variando os backbones. Podemos observar que embora a profundidade do backbone fosse aumentando, os resultados obtidos se mantiveram em um intervalo próximo, sendo o teste com o backbone ResNet 101, de modo geral, foi o que apresentou os melhores resultados. Com um MAP de 57,5%, uma precisão de 64,9%, um recall de 70,4 e um F1 score de 61%.

Já na Tabela 2, temos o resultado com as imagens que passaram pelo pré-processamento, onde foi feita a conversão para das imagens RGB para escala de

**Tabela 1. Resultados obtidos das imagens RGB.**

| <b>Backbones</b> | <b>MAP</b>  | <b>Pre(%)</b> | <b>Rec(%)</b> | <b>F1(%)</b> |
|------------------|-------------|---------------|---------------|--------------|
| Res Net 18       | <b>57.8</b> | 56.7          | <b>72.4</b>   | 64.5         |
| Res Net 34       | 58.6        | 57.8          | 73.1          | 65.4         |
| Res Net 50       | 56.1        | 61.0          | 70.7          | 65.8         |
| Res Net 101      | 57.5        | 64.9          | 70.4          | 67.6         |
| Res Net 152      | 56.5        | <b>59.6</b>   | 71.8          | 65.7         |

cinza(Gray). Nesses testes, como no anterior foram utilizados todos os backbones Res-Net. Percebemos que quanto mais raso fosse o backbone, menos arquitetura conseguiu detectar o bacilo.

**Tabela 2. Resultados obtidos das imagens GRAY.**

| <b>Backbones</b> | <b>MAP</b>  | <b>Pre(%)</b> | <b>Rec(%)</b> | <b>F1(%)</b> |
|------------------|-------------|---------------|---------------|--------------|
| Res Net 18       | <b>0.12</b> | 0.1           | <b>7.8</b>    | 3.9          |
| Res Net 34       | 0.0         | 0.2           | 0.3           | 0.2          |
| Res Net 50       | 0.0         | 0.0           | 0.0           | 0.0          |
| Res Net 101      | 22.2        | 4.9           | 62.8          | 33.8         |
| Res Net 152      | 23.5        | <b>4.9</b>    | 63.4          | 34.1         |

Já na Tabela 3, apresentamos o resultado com as imagens que passaram pelo pré-processamento, onde foi feita a conversão das imagens RGB o esquema de cores LUV. Nesses testes como nos demais, todos os backbones ResNet. Ao analisamos os resultados percebe-se que em comparação com as imagens RGB, o esquema de cores LUV não foi capaz de fornecer as características para um bom aprendizado da rede.

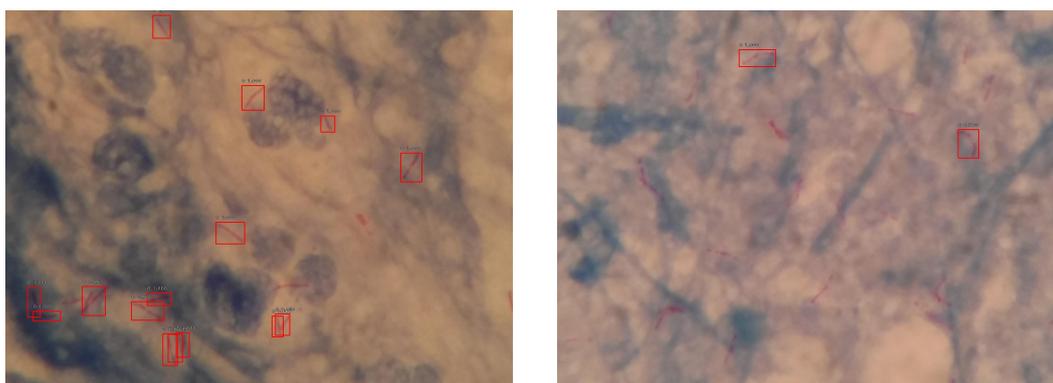
**Tabela 3. Resultados obtidos das imagens U + G + Clahe.**

| <b>Backbones</b> | <b>MAP</b>  | <b>Pre(%)</b> | <b>Rec(%)</b> | <b>F1(%)</b> |
|------------------|-------------|---------------|---------------|--------------|
| Res Net 18       | <b>22.3</b> | 15.6          | <b>43.2</b>   | 29.4         |
| Res Net 34       | 23.4        | 12.3          | 45.6          | 28.9         |
| Res Net 50       | 27.5        | 13.7          | 44.5          | 29.1         |
| Res Net 101      | 35.6        | 34.4          | 54.6          | 44.5         |
| Res Net 152      | 37.6        | <b>33.4</b>   | 64.6          | 49.0         |

Na Figura 5 temos uma representação visual dos resultados obtidos, como podemos ver em imagens RGB a arquitetura da RetinaNet conseguiu detectar razoavelmente bem os bacilos, tanto no melhor como no pior caso.

Na Figura 6 é apresentado visualmente os resultados obtidos com imagens em escala de cinza, podemos observar que no melhor caso foi detectado um número razoável de bacilos, já no pior não foi possível detectar nenhum.

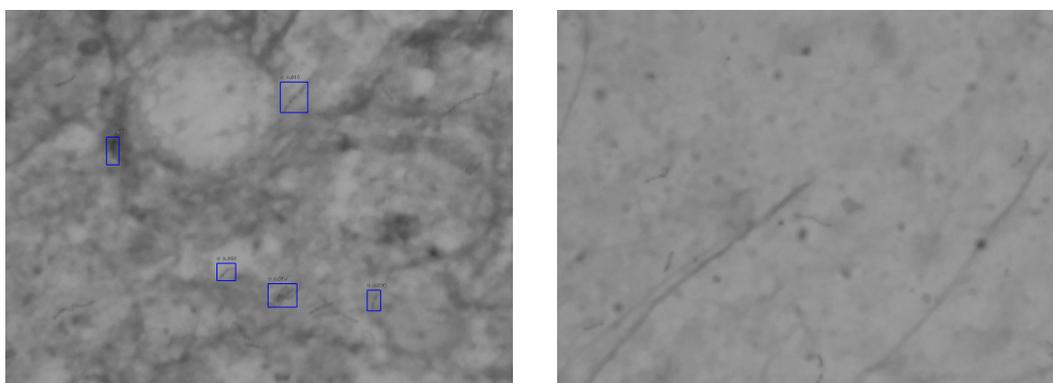
Na Figura 7 é exibido os resultados visuais que nossa arquitetura alcançou com as imagens no esquema de cores LUV, no melhor caso foi detectado apenas um bacilo, enquanto no pior não foi detectado nenhum.



(a) Melhor

(b) Pior

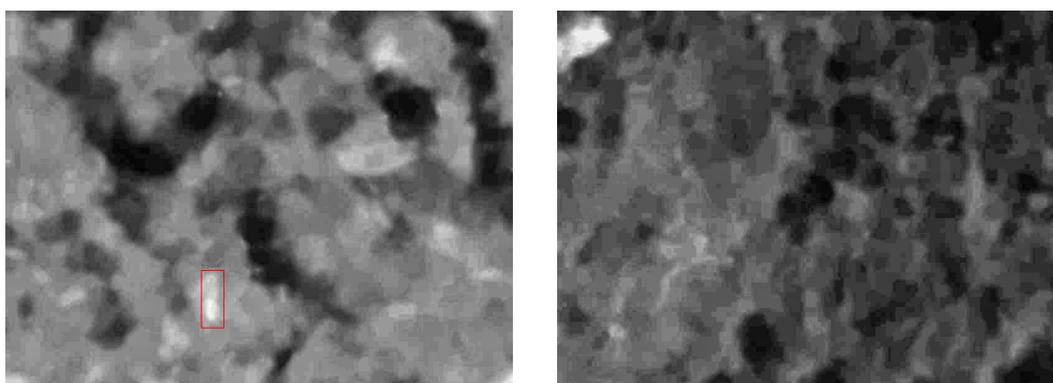
**Figura 5. Amostras qualitativas utilizando o modelo de cor RGB.**



(a) Melhor

(b) Pior

**Figura 6. Amostras qualitativas utilizando o modelo de cor gray.**



(a) Melhor

(b) Pior

**Figura 7. Amostras qualitativas utilizando o modelo de cor LUV.**

#### 4.1. Comparação com outros resultados

O trabalho de comparar os resultados obtidos com outros trabalhos não é uma tarefa simples, através dele pode se destacar o que se foi feito. Inúmeros fatores influencia para se ter uma comparação justa, como bases de imagens, métricas calculadas e métodos. Assim, a Tabela 4 apresentamos de forma quantitativa e sintetizada, os trabalhos apresentados na Seção de 2

**Tabela 4. Comparações com outros métodos.**

| <b>Trabalho</b>             | <b>Objetivo</b>                        | <b>Amostras</b> | <b>Pre(%)</b> | <b>Rec(%)</b> | <b>F1(%)</b> |
|-----------------------------|--|-----------------|---------------|---------------|--------------|
| (Kant and Srivastana, 2018) | Detecção automatizada de bacilos       | -               | 67.55         | 83.78         | -            |
| (El-melegy et al., 2019)    | Detecção automática de bacilos         | 500             | 82.6          | -             | -            |
| (Swetha et al., 2020)       | Segmentação e classificação de bacilos | -               | 99.5          | 94.7          | -            |
| <b>Proposta</b>             | <b>Detecção de bacilos</b>             | <b>1218</b>     | <b>64.9</b>   | <b>70.4</b>   | <b>61</b>    |

De modo geral, pode se dizer que nosso método é promissor, atingindo resultados muito próximos do estado da arte, uma vez que ele obtém resultados eficientes, ainda que se trate de um problema complexo. Destacamos também o número de imagens utilizadas em nosso trabalho, já que de todos os trabalhos descritos acima, o nosso foi que utilizou o maior conjunto de imagens. Além disso, nenhum dos estudos mencionaram a detecção da caixa delimitadora dos bacilos, o que pode tornar a generalização do método inferior para imagens ainda não vistas, pois delimitar especificamente o contorno dos bacilos. Finalmente, consideramos que otimizar os hiperparâmetros e promover um pré-processamento mais robusto nas imagens, ocasionará em resultados muito mais significativos.

## 5. Discussão

Neste trabalho, foi apresentado um método para detecção automática de bacilos em imagens de BA. Dito isso, baseado nos resultados mostrados na seção 6, alguns pontos merecem ser ressaltados:

1. Contribuições da metodologia proposta
  - A aplicação inovadora da RetinaNet para detecção de bacilos em BA;
  - A Função Perda Focal, utilizada no RetinaNet, mostrou ser eficiente para que o equilíbrio entre as classes seja essencial para um resultado mais eficaz;
  - Nosso 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;
2. Limitações da metodologia proposta
  - Ao analisamos os resultados, percebemos que no esquema de cores LUV, a arquitetura da RetinaNet não conseguiu detectar muitos bacilos. Acreditamos que ao realçamos imagem, além do bacilo muitas regiões onde não se havia nada, poderiam começar a ser vista como regiões candidatas, dificultando o aprendizado da rede;
  - Como visto nos 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

Neste trabalho, apresentamos um método para detecção automática de bacilos da tuberculose em imagens microscópicas usando a arquitetura da RetinaNet. A metodologia proposta apresentou resultados promissores, apesar de ser um problema complexo. Além disso, e 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. Por fim, acreditamos que nosso método é robusto e promissor.

## Referências

Araujo, F. H. D., Silva, R. R. V., Medeiros, F. N. S., Parkinson, D. D., Hexemer, A., Carneiro, C. M., and Ushizima, D. M. (2018). Reverse image search for scientific data within and beyond the visible spectrum. *Expert Systems with Applications*, 109:35–48.

da Saúde., B. M. (2022). Tuberculose. Disponível em: <https://www.gov.br/saude/pt-br/assuntos/saude-de-a-a-z/t/tuberculose>. Acessado em 15 de fevereiro de 2022.

El-Melegy, M., Mohamed, D., and Elmelegy, T. (2019). Automatic Detection of Tuberculosis Bacilli from Microscopic Sputum Smear Images Using Faster R-CNN, Transfer Learning and Augmentation, pages 270–278.

Gräbel, P., Ozkan, O., Crysandt, M., Herwartz, R., Baumann, M., Klinkhammer, B. M., Boor, P., Brummendorf, T. H., and Merhof, D. (2020). Circular anchors for the detection of hematopoietic cells using retinanet. In 2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI), pages 249–253. IEEE.

Jiao, L., Zhang, F., Liu, F., Yang, S., Li, L., Feng, Z., and Qu, R. (2019). A survey of deep learning-based object detection. *IEEE Access*, 7:128837–128868.

Kaggle (2020). Tuberculosis image dataset. Retrieved from <https://www.kaggle.com/saife245/tuberculosis-image-datasets>. Accessed June 05, 2021.

Kant, S. and Srivastava, M. M. (2018). Towards automated tuberculosis detection using deep learning. In 2018 IEEE Symposium Series on Computational Intelligence (SSCI), pages 1250–1253.

Lin, T.-Y., Dollár, P., Girshick, R., He, K., Hariharan, B., and Belongie, S. (2017). Feature pyramid networks for object detection. In 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 936–944.

Lin, T.-Y., Goyal, P., Girshick, R., He, K., and Dollár, P. (2020). Focal loss for dense object detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 42(2):318–327.

Pho, K., Mohammed Amin, M. K., and Yoshitaka, A. (2018). Segmentation-driven retinanet for protozoa detection. In 2018 IEEE International Symposium on Multimedia (ISM), pages 279–286.

Reza, A. M. (2004). 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*, 38(1):35–44.

SESA (2022). Tuberculose. Disponível em: <https://saude.es.gov.br/neve-tuberculose>. Acessado em 8 de fevereiro de 2022.

Swetha, K., Sankaragomathi, B., and Thangamalar, J. B. (2020). Convolutional neural network based automated detection of mycobacterium bacillus from sputum images. In 2020 International Conference on Inventive Computation Technologies (ICICT), pages 293–300.

Targ, S., Almeida, D., and Lyman, K. (2016). Resnet in resnet: Generalizing residual architectures. arXiv preprint arXiv:1603.08029.

WHO, W. H. O. (2022). Tuberculosis. Disponível em: <https://www.who.int/en/news-room/fact-sheets/detail/tuberculosis>. Acessado em 05 de fevereiro de 2022.