

Metodologia Automática para Classificação de Bacilos Utilizando a Arquitetura Faster R-CNN

Automatic Methodology for Classification of Bacillos Using the Faster R-CNN

Augusto Almondes Teixeira
Departamento de Sistemas de
Informação
Universidade Federal do Piauí, Brasil
augusto7666@ufpi.edu.br

Antonio Oseas de Carvalho
Filho
PPG em Engenharia Eletrica,
Mestrado e Doutorado em
Ciência da Computação
Depto. de Sistemas de Informação
Universidade Federal do
Piauí, Brasil
antoniooseas@ufpi.edu.br

Bruno Guedes Alcoforado
Aguiar
Depto. de Medicina Comunitaria/CCS,
Universidade Federal do

Piauí, Brasil
guedesaguiar@ufpi.edu.br
Flávio Henrique Duarte De
Araújo
PPG em Engenharia Eletrica
Doutorado em Ciência da Computação
Depto. de Sistemas de Informação
Universidade Federal do
Piauí, Brasil
flavio86@ufpi.edu.br

Paulo H. de Carvalho Silva
Depto. de Sistemas de
Comunitária/CCS, Informação,
Universidade Federal do
Piauí, Brasil
paulo103henrique@ufpi.edu.br

Romuer Rodrigues
Veloso e Silva
PPG em Engenharia Eletrica
Mestrado e Doutorado em
Ciência da Computação
Depto. de Sistemas de Informação
Universidade Federal do
Piauí, Brasil
romuere@ufpi.edu.br

ABSTRACT

Tuberculosis, caused by the bacterium *Mycobacterium tuberculosis*, primarily affects the lungs and is transmitted through the air. It is one of the leading causes of death from infectious diseases worldwide, especially in developing countries such as Brazil. The most common diagnostic method is bacilloscopy, a manual analysis of sputum samples under a microscope, which can be slow and delay timely treatment of patients. This work proposes the use of Computer Vision and Deep Learning to automate the detection and classification of bacilli in bacilloscopy, accelerating the diagnosis of tuberculosis. The results obtained demonstrate a high potential for the approach, with an average Recall of 89.65%, an average F1-Score of 71.82%, and an average Precision of 64.20%.

Keywords: Tuberculosis; Bacilloscopy; Faster R-CNN;

RESUMO

A tuberculose, causada pela bactéria *Mycobacterium tuberculosis*, afeta principalmente os pulmões e é transmitida pelo ar. É uma das principais causas de morte por doenças infecciosas no mundo, especialmente em países em desenvolvimento, como o Brasil. O método mais comum de diagnóstico é a baciloscopia, uma análise manual de amostras de escarro ao microscópio, que pode ser lenta e comprometer o tratamento rápido dos pacientes. Este trabalho propõe o uso de Visão Computacional e *Deep Learning* para automatizar a detecção e classificação de bacilos em baciloscopias, agilizando o diagnóstico da tuberculose. Os resultados obtidos demonstram um alto potencial da abordagem, com um Recall

médio de 89,65%, F1-Score médio de 71,82% e uma Precisão média de 64,20%.

Palavras-chave: Tuberculose; Baciloscopia; Faster R-CNN;

1. INTRODUCTION

A Tuberculose (TB) é uma doença infecciosa causada pelo *Mycobacterium tuberculosis* (MT). Essa doença afeta em grande parte os pulmões, mas podendo infectar outras partes do corpo (rins, ossos, sistema nervoso e gânglios linfáticos). Transmitida, em sua maior parte, por meio do ar quando uma pessoa infectada tosse ou espirra, liberando pequenas gotículas contendo a bactéria. Sua existência é datada a muito tempo, existindo registros de casos de infecção pela doença em múmias egípcias, além de ser tema de estudo na Grécia Antiga [7]. A TB é considerada a segunda doença infecciosa que mais matou adultos e adolescentes em todo o mundo, ficando para trás somente da Covid-19 no período da pandemia, e é considerada um problema importante de saúde pública [5].

Segundo o relatório global da Organização Mundial da Saúde (OMS), no ano de 2022 foi totalizado 1,3 milhões de mortes por conta da doença, no qual 167 mil desses possuíam HIV [13]. No Brasil, o número de casos da enfermidade ainda é bastante elevado. Conforme dados do Ministério da Saúde (MS), no País são notificados, por ano, aproximadamente 80 mil novos casos de infecção e cerca de 5,5 mil óbitos devido à TB [4]. Tal mal se transforma a cada ano em um problema cada vez mais grave, por conta do crescimento no número de óbitos. Os casos da doença ocorrem, em grande parte, em países que possuem uma Renda Nacional Bruta (RNB) média e baixa [13].

Apesar dos avanços tecnológicos terem introduzido novos métodos de diagnóstico e tratamento da TB, muitos deles não são amplamente adotados devido aos altos custos de produção e execução [8]. Como meios para o diagnóstico da TB, pode-se citar: análise microscópica, radiografia do tórax, radiografia de

pele, cultura de TB e teste de liberação de interferon-gama. No entanto, o método mais utilizado é a baciloscopia de escarro, uma análise microscópica do esfregaço de escarro do paciente. A detecção e enumeração dos MTs são realizadas por leitura microscópica de lâminas por técnicos devidamente treinados [5]. Por mais que essa maneira de diagnóstico seja a mais utilizada, ainda requer uma melhoria na sua eficiência, isso porque o processo do exame é feito manualmente. Além disso, essa atividade necessita tempo, paciência, e um profissional preparado para realizá-la, principalmente no estágio inicial da enfermidade, onde existe a dificuldade na detecção da doença.

Em virtude do exposto, e com o avanço da tecnologia, diversos métodos inovadores vêm surgindo para o auxílio no combate à doença e no decorrer do tratamento, como a aplicação de Visão Computacional (VC) baseado em Deep Learning (DL). A aplicação dessa tecnologia na área vem crescendo com o passar do tempo. No diagnóstico da TB, estes métodos podem ser utilizados para automatizar e melhorar o processo de detecção de MT em amostras de escarro, tornando-o mais preciso e eficiente. Esta abordagem implica em auxiliar no diagnóstico precoce da doença.

Nesse contexto, este artigo propõe um modelo para a classificação de MT com base na baciloscopia de escarro, visando sua aplicação em sistemas de apoio ao diagnóstico precoce da TB. Utilizou-se uma base de dados composta por imagens de esfregaços de escarro fixados em lâminas. A detecção dos bacilos será realizada por meio do treinamento de uma rede neural convolucional utilizando o modelo Faster R-CNN.

2. TRABALHOS RELACIONADOS

Nesta seção, serão revisados estudos e pesquisas realizados anteriormente integrando as áreas de visão computacional e detecção de TB com algoritmos avançados de Aprendizado de Máquina (AP).

No trabalho [5], é proposto o desenvolvimento de um sistema automatizado para detecção e contagem de BAAR no diagnóstico da TB. Inicialmente, foram utilizadas técnicas de visão computacional e processamento digital de imagens, posteriormente evoluindo para o uso de AP e aprendizado profundo. A pesquisa inclui a criação de um banco de imagens e testes com diversas configurações. Conclusões preliminares indicam que a abordagem de inteligência artificial é mais precisa que métodos anteriores. O estudo envolve a comparação entre leituras manuais de lâminas e o algoritmo de deep learning, com 340 lâminas preparadas para esse fim. Na literatura [8], o autor propõe um modelo de identificação do bacilo da TB usando técnicas de processamento digital de imagens, visando um diagnóstico mais rápido e eficiente. Foi desenvolvido um protótipo chamado DetectBAAR, utilizando a biblioteca OpenCV, para processar e analisar imagens de amostras coloridas pela técnica de Ziehl-Neelsen. O protótipo mostrou-se eficaz, mas sensível a condições de iluminação não controladas, indicando a necessidade de aprimoramentos futuros. Os resultados experimentais comprovam a eficácia do modelo, sugerindo seu potencial para a detecção de outras doenças através de metodologias similares.

No estudo [10] foi criada uma solução usando Reconhecimento Óptico de Caracteres (OCR) com Aprendizagem

Profunda para superar dúvidas sobre a veracidade dos dados coletados por soluções de IoT (*Internet of Things*, ou em português, Internet das Coisas). Criou-se um conjunto de dados com imagens reais de hidrômetros e treinou-se 22 modelos baseados na Faster R-CNN. O modelo com melhor desempenho alcançou mais de 90% de precisão na leitura dos primeiros cinco dígitos dos hidrômetros. O autor destaca que soluções comerciais de OCR não são sempre adequadas para domínios específicos. A abordagem de detecção de objetos usando Faster R-CNN com Resnet50 mostrou-se robusta, alcançando 93% de assertividade, superando os desafios de desalinhamento dos caracteres e interferências. Este resultado valida a eficácia do dataset criado e a precisão do modelo em diversas condições operacionais.

No trabalho [2], é destacado o potencial do modelo Faster R-CNN na detecção de pêssegos, visando a implementação de um sistema eficaz de estimativa de produção em pomares. O modelo demonstra grande potencial para a detecção de pêssegos, podendo ser uma ferramenta valiosa para estimar a produção em pomares, auxiliando no planejamento da colheita e armazenamento. No trabalho, os resultados preliminares indicam um desempenho bom, mesmo para frutos agrupados e oclusos. Para melhorar a detecção, sugere-se aumentar a quantidade de imagens de treino e ajustar os limites de oclusão dos frutos.

O artigo [1] avalia o uso de modelos de detecção de objetos para identificar núcleos corados com nitrato de prata em pacientes com câncer de colo de útero, utilizando o método AgNOR. Comparando diferentes backbones do modelo Faster-RCNN, o trabalho demonstra que a aprendizagem profunda pode melhorar significativamente o processamento de imagens citológicas. O melhor desempenho alcançou acurácia de 0,66, precisão de 0,79 e revocação de 0,80, validando a eficácia do método para detectar núcleos AgNOR. Essa abordagem automatizada pode auxiliar na rotina clínica, tornando a identificação de núcleos mais objetiva e eficiente, em contraste com os métodos manuais subjetivos e laboriosos.

Esses trabalhos exploram diferentes abordagens para a detecção de objetos, como a identificação de bacilos da TB, a estimativa de produção em pomares e a identificação de núcleos em imagens citológicas. Todos os trabalhos utilizaram Visão Computacional para a detecção e classificação de objetos, diferindo somente a estrutura e suas características, com ênfase na Faster R-CNN. Nesse artigo será realizado a otimização do modelo para a detecção de objetos pequenos, além do aprimoramento do modelo através de ajustes nos parâmetros de treinamento e na utilização de um conjunto de dados completo. Estes aprimoramentos visam aumentar a precisão e a eficácia do diagnóstico automatizado da TB, contribuindo significativamente para a área de saúde.

3. MÉTODO PROPOSTO

Nesta seção, será descrita a metodologia que foi adotada para o desenvolvimento do modelo de classificação de MTs utilizando a rede Faster R-CNN. A metodologia está dividida em cinco fases: Fase I - Revisão sistemática; Fase II Pesquisa Bibliográfica; Fase III - Aquisição e preparação de dados; Fase IV - Treinamento do modelo; Fase V - Avaliação e validação do modelo; e Fase VI - Análise dos resultados.

Fase I - Revisão sistemática: Realização de revisão sistemática da literatura sobre classificação e enumeração de MT. Foram coletados e analisados trabalhos científicos relevantes, artigos e publicações sobre visão computacional e técnicas de AP aplicadas à detecção de MT.

Fase II - Pesquisa Bibliográfica: Foi conduzido estudos aprofundados sobre processamento de imagens, visão computacional e redes neurais convolucionais. Foram revisados os principais métodos computacionais empregados no processamento de imagens e na detecção de objetos, visando escolher os mais adequados para este trabalho.

Fase III - Preparação de dados: Foi realizado a preparação dos dados essenciais para o treinamento e validação do modelo. Ajustes foram realizados nas imagens para melhorar a eficiência do treinamento e do próprio modelo.

Fase IV - Treinamento do modelo: Fase para realizar a implementação da rede Faster R-CNN para detecção e classificação de MT. O modelo será treinado usando a base de dados preparada na fase anterior, aplicando técnicas e ajustes de hiperparâmetros para melhorar o desempenho do modelo.

Fase V - Avaliação e validação do modelo: Foi feita a avaliação do desempenho do modelo treinado utilizando métricas como acurácia, precisão, recall e F1 score. Com a posse das métricas, é analisado se a pesquisa seguiria para a próxima fase ou retornava para a fase anterior (Fase IV), a fim de realizar um novo treinamento com ajustes.

Fase VI - Análise dos resultados: Nesta etapa, os resultados obtidos foram analisados detalhadamente para determinar a eficácia e precisão da ferramenta desenvolvida, para, somente assim, disponibilizar para uso clínico.

3.1 Base de dados

Foi obtido um conjunto com um total de 680 imagens de microscópio, extraídas de uma base de dados privada, montada em parceria com o Hospital Universitário da Universidade Federal do Piauí (UFPI), em Teresina. As imagens são acompanhadas de arquivos JSON (*JavaScript Object Notation*) contendo as anotações correspondentes, que são essenciais para o funcionamento da rede neural.

As imagens originais, com dimensões extensas de 3000x4000, foram cortadas em blocos menores de 640x640 pixels para reduzir o tempo de processamento da rede neural e aumentar a eficiência do treinamento. Esse procedimento não só melhora o desempenho computacional, como também expande a base de dados, proporcionando um conjunto mais diversificado e robusto. Após o corte, foi gerado um total de 8596 imagens, que foram divididas em três grupos distintos (Tabela 1). Para a validação do modelo, foram separadas 640 imagens, enquanto o conjunto de treino contém 6905 imagens e o conjunto de teste foi composto por 1051 imagens, garantindo uma distribuição adequada para cada fase do treinamento e avaliação da rede neural.

Table 1: Divisão das imagens em cada grupo após o corte.

Grupos	Quantidades	Porcentagem Aproximada
Validação	640	8%
Treinamento	6905	80%
Teste	1051	12%

A separação do conjunto de dados em sub-conjuntos é essencial para avaliar o desempenho do modelo em dados que não foram utilizados durante o processo de treinamento [12]. Isto é, essa divisão de conjuntos evita o *overfitting*, que ocorre quando um modelo de AP aprende excessivamente os detalhes e ruídos específicos do conjunto de treino, a ponto de prejudicar sua capacidade de generalizar para novos dados.

3.2 Topologia

Nos últimos anos, o avanço das tecnologias de inteligência artificial e AP tem transformado diversas áreas do conhecimento. A visão computacional é uma das que mais se destacam nesse cenário. A capacidade de sistemas computacionais interpretarem e extraírem informações valiosas de imagens e vídeos abre um vasto leque de aplicações, desde reconhecimento facial até diagnósticos médicos automatizados. Esse progresso tem sido impulsionado pela disponibilidade de grandes volumes de dados e pelo aumento do poder computacional, o que possibilita a criação de modelos cada vez mais complexos e eficientes.

Esta área tem crescido rapidamente, levando à criação de um número crescente de modelos de computação. Isso implica que, ao enfrentar um desafio de visão computacional, pode ser complicado selecionar o modelo mais adequado ou identificar quais métodos de validação e treinamento são mais adequados para o cenário específico [11]. Isso leva à necessidade de analisar fatores que auxiliam na seleção do modelo mais adequado, como estrutura da base de dados, quantidade de recursos de processamento, complexidade do problema, entre outros fatores.

Tendo em vista tal abordagem, a topologia selecionada foi a Faster R-CNN [9], devido sua robustez, além do seu alto desempenho em detecção, quando analisado sua precisão e eficiência computacional. Por ser um detector *twostage* (dois estágios), possui alta precisão de localização e reconhecimento de objetos [6]. Além disso, foi utilizado o framework Detectron2, desenvolvido pelo Facebook e escrito em Python. Esse framework foi escolhido devido à sua fácil compreensão e ao fato de já possuir implementações prontas, permitindo a modificação de apenas alguns parâmetros.

O modelo Faster R-CNN [9] é o aprimoramento do R-CNN (*Region-based Convolutional Neural Networks*) com a adição de uma rede para criar propostas de região (*Region Proposal Network* - RPN), substituindo o método de busca seletiva. A rede de proposta de região auxilia na diminuição do tempo de processamento [6]. O algoritmo navega pela *feature map* – representação visual das características que foram extraídas da imagem – utilizando janelas deslizantes. Cada janela deslizante contém 'âncoras', representando pontos de referência para localizar regiões de interesse. São utilizadas várias 'âncoras' com diferentes tamanhos e proporções em cada janela, permitindo assim a detecção de objetos de diversos tamanhos [3]. A Figura 1 mostra o passo a passo da rede.

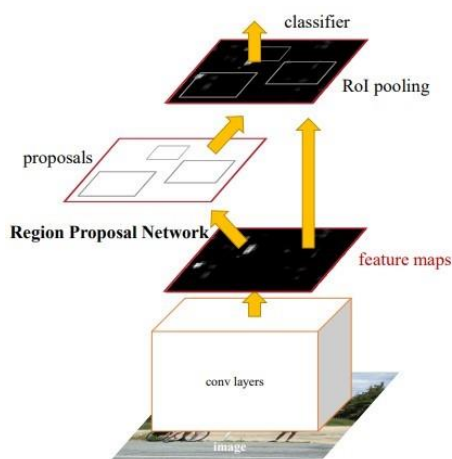


Figure 1: Modelo de funcionamento do Faster R-CNN. [9]

4. RESULTADOS E DISCUSSÃO

Nessa seção, serão abordados os resultados obtidos a partir da aplicação das tecnologias selecionadas, bem como suas implicações e relevância para o contexto da pesquisa. Serão analisados os dados coletados, discutindo-se os padrões identificados e as tendências observadas, visando fornecer uma compreensão aprofundada do tema em questão.

4.1 Métricas

A avaliação do desempenho de um modelo de aprendizado de máquina é crucial para garantir que ele esteja funcionando de maneira eficiente e eficaz, especialmente em tarefas sensíveis como a detecção de MT. As métricas de avaliação fornecem uma base quantitativa para medir a precisão e a confiabilidade das previsões do modelo, permitindo identificar suas forças e limitações. A qualidade da detecção e contagem de MT serão avaliadas usando métricas padrão em detecção de objetos, como:

- Verdadeiro Positivo (VP): Classificações corretas da classe positiva.
- Falso Positivo (FP): Classificações incorretas para a classe positiva.
- Verdadeiro Negativo (VN): Classificações corretas da classe negativa.
- Falso Negativo (FN): Classificações incorretas para a classe negativa.

Com a posse desses valores, é possível calcular as métricas de desempenho, servindo para a avaliação do modelo. Foram selecionadas três métricas essenciais: Recall, Precisão e F1 Score. A primeira mede a capacidade do modelo de detectar corretamente todas as instâncias positivas, minimizando falsos negativos. Já Precisão avalia a proporção de previsões corretas, reduzindo falsos positivos. O F1 Score equilibra Precisão e Recall, combinando ambos em uma métrica que reflete o

desempenho geral do modelo, especialmente útil quando se busca um equilíbrio entre detecção e exatidão.

4.2 Treinamento e avaliação do modelo

O treinamento do modelo foi realizado utilizando a plataforma Google Colab, que oferece recursos gratuitos para acesso às unidades de processamento gráfico (GPUs). Essas GPUs são essenciais para acelerar o processamento de grandes volumes de dados, especialmente em tarefas de treinamento de redes neurais profundas. O Google Colab disponibiliza essas unidades de computação gratuitamente, permitindo que usuários realizem treinamentos complexos em um ambiente de execução de alto desempenho.

Com as configurações dos parâmetros devidamente ajustadas, deu-se início ao treinamento do modelo que durou aproximadamente três horas. Após a conclusão, foi necessário realizar uma avaliação rigorosa para validar sua eficácia na tarefa específica de detecção de MT. A avaliação do modelo é uma etapa crucial, pois permite determinar a capacidade do modelo de generalizar para novos dados, ou seja, como ele se comporta em relação a dados que não foram vistos durante o treinamento.

Após aplicar o modelo a uma série de teste, foi obtido as métricas necessárias para avaliar o mesmo. A métrica de precisão foi obtida por meio do framework Detectron2, que possui um API capaz de calcular métricas de Precisão Média (*Average Precision*, em inglês) de diferentes medidas. As outras métricas selecionadas (Recall e o F1 Score) foram calculadas por meio de funções criadas e implementadas a API devido a ausência de suporte direto para essas métricas do Detectron2.

Ao fim da avaliação, foi obtido um Recall médio de 89,65%, F1 Score médio de 71,81% e um AP50 (*Average Precision* para um IoU de 50%. Indica o desempenho do modelo para uma tolerância de sobreposição mais baixa) de 64,20% (Tabela 2).

Table 2: Métricas de Recall, F1 Score e AP50, essa última calculada pela API.

Label	Recall médio	F1 Score médio	AP50
Bacilo	89,65%	71,82%	64,20%

Os resultados obtidos com as métricas selecionadas demonstram um desempenho robusto do modelo na tarefa de detecção de MT. O Recall médio de 89.65% indica que o modelo é altamente eficaz em identificar a maioria dos bacilos presentes nas amostras. Já o F1 Score médio de 71.81% reflete um equilíbrio entre precisão e recall, sugerindo que ainda há espaço para melhorias na precisão. Por fim, a métrica AP50 de 64.20% mostra que o modelo possui uma capacidade moderada de prever corretamente a localização dos MT com uma sobreposição de 50% entre as caixas delimitadoras, o que é satisfatório para muitos cenários clínicos onde a detecção geral é mais importante do que a exatidão na localização.

A predição foi aplicada nas imagens do dataset de validação para avaliar a robustez do modelo. A capacidade de identificar MT foi imposta corretamente, mas ainda requer melhorias no treinamento para aumento da precisão. A Figura 2 mostra previsões com a classe do bacilo e a confiança do modelo. Embora tenha ocorrido a identificação na maioria das regiões

de interesse, algumas predições apresentam baixa confiança ou erros, indicando dificuldades com variações nas imagens, como ruído, coloração e sobreposições.

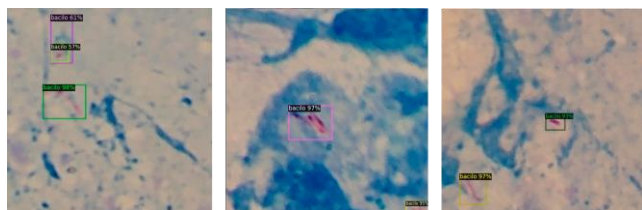


Figure 2: Predições de bacilos em imagens da base de dados de Validação

4.3 Conclusão

Com base nos resultados apresentados, é possível concluir que o modelo Faster R-CNN demonstrou uma capacidade robusta para a detecção de MT em imagens clínicas. O uso da métrica AP50, acompanhado das métricas de Recall e F1 Score, permitiu uma avaliação detalhada do desempenho do modelo, indicando que ele é altamente eficaz em identificar MT, mas ainda enfrenta desafios relacionados à precisão. Por conta disso, apesar dos resultados positivos, a análise das predições revelou que o modelo pode ser sensível a variações nas imagens, como níveis de ruído e sobreposições de bacilos, o que impacta a confiança e a precisão das detecções. Um fator importante que afetou o treinamento foi a limitação do número de iterações, devido às restrições de tempo de uso e capacidade de processamento da GPU do Google Colab. Essas limitações impediram que o modelo fosse treinado por um período mais prolongado ou com um maior número de iterações, comprometendo a capacidade de generalização e precisão nas predições. Esses fatores destacam a importância de considerar ambientes de treinamento mais robustos ou a utilização de estratégias que permitam uma melhor gestão dos recursos computacionais.

Enquanto o modelo atual demonstrou um desempenho promissor, sua eficácia pode ser significativamente melhorada com ajustes no processo de treinamento e a utilização de recursos computacionais mais adequados. Tornando isso o próximo passo para o modelo se tornar uma ferramenta ainda mais confiável e precisa para a detecção de MT em imagens clínicas, contribuindo para avanços na área de diagnósticos automatizados.

5. REFERENCES

- [1] J. G. A. Amorim, L. A. B. Macarini, A. V. Matias, A. Cerentini, F. B. D. M. Onofre, A. S. C. Onofre, and A. Von Wangenheim. Detecção de núcleos em imagens citológicas de agnir utilizando aprendizado profundo. *Journal of Health Informatics*, 12, 2020.
- [2] E. T. Assunção, P. D. Gaspar, R. Mesquita, A. Veiros, and H. Proença. Resultados preliminares de detecção de imagens de pêssegos aplicando o método faster r-cnn. *Revista da Associação Portuguesa de Horticultura-março de 2020*, (136):30–33, 2020.
- [3] L. Cerqueira, K. de Souza, S. Ternes, and

J. Camargo Neto. Usando a rede neural faster-rcnn para identificar frutos verdes em pomares de laranja.

In: CONGRESSO INTERINSTITUCIONAL DE

INICIAÇÃO CIENTÍFICA, 14., 2020. Anais ..., 2020. [4] M. da Saúde Brasil. Tuberculose, 2023.

- [5] J. V. B. Gomide, C. J. Augusto, E. A. Leal, J. P. Tarabal, N. V. de Castro Barroso, M. A. C. Soares, and B. P. Lima. Detecção e contagem automáticas de bacilos álcool ácido resistentes para o diagnóstico da tuberculose. *Código 31: revista de informação, comunicação e interfaces*, 1(1), 2023.
- [6] L. Jiao, F. Zhang, F. Liu, S. Yang, L. Li, Z. Feng, and R. Qu. A survey of deep learning-based object detection. *CoRR*, abs/1907.09408, 2019.
- [7] V. D. O. MARTINS and C. V. DE MIRANDA. Diagnóstico e tratamento medicamentoso em casos de tuberculose pulmonar: revisão de literatura. *Revista saúde multidisciplinar*, 7(1), 2020.
- [8] V. J. Pinto. *Detecção do bacilo da tuberculose através do processamento e análise de imagem microscópica*. PhD thesis, Mestrado em Sistemas de Informação e Gestão do Conhecimento, 2018.
- [9] S. Ren, K. He, R. Girshick, and J. Sun. Faster r-cnn: Towards real-time object detection with region proposal networks. *Advances in neural information processing systems*, 28, 2015.
- [10] P. J. L. d. SILVA. Faster r-cnn para extração de caracteres em hidrômetros. B.S. thesis, 2023.
- [11] D. H. C. Vassalo et al. Análises de competições presentes na plataforma kaggle para auxiliar no desenvolvimento de novas soluções para problemas de visão computacional. 2021.
- [12] L. C. Vieira and E. S. Pereira. Análise de tráfego urbano: visão computacional e inteligência artificial para otimização de semáforos. B.S. thesis, 2024. [13] W. H. O. WHO. Tuberculosis, 2023.