



UNIVERSIDADE FEDERAL DO AMAZONAS  
CAMPUS SENADOR ARTHUR VIRGÍLIO  
FACULDADE DE TECNOLOGIA - FT



Paulo Vitor de Castro Freitas

ANÁLISE DE ARQUITETURAS DE REDES NEURAIIS PARA DETECÇÃO DE  
BACILOS EM BACIOSCOPIA DE CAMPO CLARO USANDO RADIOMICS

MANAUS- AM

Janeiro de 2026

Paulo Vitor de Castro Freitas

ANÁLISE DE ARQUITETURAS DE REDES NEURAIIS PARA DETECÇÃO DE  
BACILOS EM BACILOSCOPIA DE CAMPO CLARO USANDO RADIOMICS

Monografia de Graduação apresentada à  
Coordenação de Engenharia da  
Computação, UFAM, da Universidade  
Federal do Amazonas, como parte dos  
requisitos necessários à obtenção do  
título de Engenheiro da Computação.

Orientador: Prof. Dr. Cícero Ferreira Fernandes Costa Filho

MANAUS- AM

Janeiro de 2026

ANÁLISE DE ARQUITETURAS DE REDES NEURAIIS PARA DETECÇÃO DE  
BACILOS EM BACIOSCOPIA DE CAMPO CLARO USANDO RADIOMICS

Paulo Vitor de Castro Freitas

MONOGRAFIA SUBMETIDA AO CORPO DOCENTE DO CURSO DE ENGENHARIA  
DA COMPUTAÇÃO DA UNIVERSIDADE FEDERAL DO AMAZONAS COMO PARTE  
DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE  
ENGENHEIRO.

Aprovada por:

Assinado por:

*Cícero Ferreira Fernandes Costa Filho*

34A2B40CC5114BA...

Prof. Dr. Cícero Ferreira Fernandes Costa Filho

DocuSigned by:

*Frederico da Silva Pinagé*

F0F3100F7F24428...

Prof. Dr. Frederico da Silva Pinagé

Assinado por:

*Marly Guimarães F Costa*

A0BFD18CDDD14AF...

Profa. Dra. Marly Guimarães Fernandes Costa

Manaus

Janeiro de 2026

# Agradecimentos

Sou grato ao divino mestre arquiteto universal, minha família e amigos.

# Resumo

A tuberculose (TB) continua sendo um grave problema de saúde pública global, especialmente em países de média e baixa renda. O método diagnóstico mais amplamente empregado é a baciloscopia de campo claro do escarro, técnica de baixo custo, porém dependente da experiência do microscopista, demorada e sujeita à variabilidade intra e interobservadora. Nesse cenário, abordagens automatizadas de análise de imagens têm potencial para aumentar a precisão, padronizar resultados e reduzir a carga de trabalho em ambientes clínicos. Entre essas abordagens, técnicas de aprendizado profundo e métodos baseados em extração radiômica destacam-se por capturar padrões estruturais e texturais relevantes. Este trabalho realiza uma análise comparativa entre diferentes arquiteturas de redes neurais para detecção de bacilos em baciloscopias de campo claro, utilizando tanto imagens brutas quanto características radiômicas extraídas com PyRadiomics. O conjunto de dados utilizado contém 31.484 *patches* de imagens de microscopia, sendo 2.636 contendo bacilos (positivas) e 28.848 correspondentes ao fundo ou artefatos (negativas), sendo posteriormente balanceado para 2.635 *patches* por classe. Foram avaliados dois *pipelines* radiômicos: (i) extração usando máscaras segmentadas apenas nas imagens positivas e usando uma máscara de dimensão 40x40 (máscara “cheia”) nas negativas e (ii) extração usando uma máscara cheia em todas as imagens. Em ambos os casos, aplicou-se seleção de características por remoção de variância zero, filtragem por correlação e avaliação de importância com Random Forest. No primeiro *pipeline*, observou-se desempenho perfeito (acurácia, precisão, sensibilidade e *F1-score* iguais a 1.00) tanto no MLP quanto no RF. Entretanto, essa performance não reflete o comportamento real do problema: ela foi causada por vazamento de informação, principalmente por características como *diagnostics\_Mask-original\_VoxelNum*, cujo valor permaneceu constante nas imagens negativas devido ao uso de máscara cheia, criando um padrão artificial fortemente correlacionado ao rótulo. No *pipeline* com máscaras cheias para todas as imagens, o desempenho do RF (0.89) e do MLP (0.88) refletiu uma avaliação mais realista da abordagem radiômica. A rede convolucional treinada diretamente nas imagens brutas apresentou o melhor desempenho real e confiável, alcançando acurácia de 0.99, superando os métodos radiômicos após correção do vazamento.

**Palavras-chave:** Tuberculose, Baciloscopia, Radiomics, PyRadiomics, Redes Neurais, CNN, Detecção de Bacilos, Inteligência Artificial.

# Abstract

Tuberculosis (TB) remains a major global public health challenge, particularly in low- and middle-income countries. The most widely used diagnostic method is sputum light-field microscopy (bacilloscopy), a low-cost technique that is, however, highly dependent on the microscopist's expertise, time-consuming, and subject to intra- and inter-observer variability. In this context, automated image analysis approaches have the potential to increase accuracy, standardize results, and reduce the workload in clinical settings. Among these approaches, deep learning techniques and radiomic extraction methods stand out for capturing relevant structural and textural patterns. This work performs a comparative analysis between different neural network architectures for bacillus detection in light-field microscopy, utilizing both raw images and radiomic features extracted with PyRadiomics. The dataset used contains 31,484 microscopy image patches, with 2,636 containing bacilli (positives) and 28,848 corresponding to the background or artifacts (negatives), subsequently balanced to 2,635 patches per class. Two radiomic pipelines were evaluated: (i) extraction using segmented masks only for positive images and a 40x40 full mask for negatives, and (ii) extraction using a full mask for all images. In both cases, feature selection was applied through zero-variance removal, correlation filtering, and importance evaluation with Random Forest. In the first pipeline, perfect performance was observed (accuracy, precision, sensitivity, and F1-score equal to 1.00) for both MLP and RF. However, this performance does not reflect the real behavior of the problem: it was caused by data leakage, primarily due to features such as `diagnostics_Mask-original_VoxelNum`, whose value remained constant in negative images due to the use of a full mask, creating an artificial pattern strongly correlated with the label. In the pipeline with full masks for all images, the performance of RF (0.89) and MLP (0.88) reflected a more realistic evaluation of the radiomic approach. The convolutional neural network trained directly on raw images showed the best real and reliable performance, achieving an accuracy of 0.99, surpassing the radiomic methods after leakage correction.

**Keywords:** Tuberculosis, Bacilloscopy, Radiomics, PyRadiomics, Neural Networks, CNN, Bacillus Detection, Artificial Intelligence.

# Lista de Figuras

Figura 1: Funcionamento da camada convolutiva. Fonte: Nvidia (2025).	29
Figura 2: Ilustração do funcionamento da camada de Max Pooling e Average Pooling. Fonte: Muhamad Yani et al (2019).	30
Figura 3: Gráfico da função de ativação ReLU.	31
Figura 4: Método Dropout, proposto por Srivastava et al. (2014).	32
Figura 5: Topologia RF. Fonte: Researchgate (2025).	34
Figura 6: Topologia MLP.	35
Figura 7: PyRadiomics pipeline. Fonte: PyRadiomics (2017).	36
Figura 8: Amostra do dataset positivo e suas respectivas máscaras segmentadas com intensidade dos pixels já invertida.	45
Figura 9: Amostra do dataset negativo.	45
Figura 10: Fluxograma da metodologia	46
Figura 11: Arquitetura da rede CNN utilizada	49
Figura 12: Curva de convergência de acurácia da rede CNN	51
Figura 13: Curva de convergência de perda da rede CNN	52
Figura 14: Matriz de confusão dos resultados com a rede CNN	52
Figura 15: Rank das 10 melhores características por importância em % com dataset usando máscaras segmentadas.	54
Figura 16: Rank das 10 melhores características por importância em % com dataset usando máscaras segmentadas.	55

# Lista de Tabelas

Tabela 1: Tabela de revisão da literatura	24
Tabela 2: Atributos extraídos	47
Tabela 3: Descrição das 10 melhores características selecionadas para dataset usando máscaras segmentadas	53
Tabela 4: Descrição das 10 melhores características selecionadas para dataset usando máscaras cheias	55
Tabela 5: Métricas obtidas para os modelos avaliados	56

# Lista de Siglas

<b>CNN</b>	Rede neural convolucional ( <i>Convolutional Neural Network</i> )
<b>GLCM</b>	Matriz de coocorrência de níveis de cinza ( <i>Gray Level Co-occurrence Matrix</i> )
<b>GLRLM</b>	Matriz de comprimento de execução de níveis de cinza ( <i>Gray Level Run Length Matrix</i> )
<b>GLSZM</b>	Matriz de zona de tamanho de níveis de cinza ( <i>Gray Level Size Zone Matrix</i> )
<b>IBSI</b>	Iniciativa de Padronização de Biomarcadores de Imagem ( <i>Image Biomarker Standardisation Initiative</i> )
<b>MLP</b>	Perceptron multicamadas ( <i>Multilayer Perceptron</i> )
<b>NGTDM</b>	Matriz de diferença de tom de cinza por vizinhança ( <i>Neighborhood Gray Tone Difference Matrix</i> )
<b>RF</b>	Floresta aleatória ( <i>Random Forest</i> )
<b>ROI</b>	Região de interesse ( <i>Region of Interest</i> )
<b>TB</b>	Tuberculose ( <i>Tuberculosis</i> )
<b>BAAR</b>	Bacilos álcool-ácido resistentes ( <i>Acid-Fast Bacilli</i> )
<b>SGD</b>	Gradiente descendente estocástico ( <i>Stochastic Gradient Descent</i> )
<b>ReLU</b>	Unidade linear retificada ( <i>Rectified Linear Unit</i> )
<b>BN</b>	Normalização em lote ( <i>Batch Normalization</i> )
<b>SGDM</b>	Gradiente descendente estocástico com momentum ( <i>Stochastic Gradient Descent with Momentum</i> )
<b>RGB</b>	Vermelho, verde e azul ( <i>Red, Green, Blue</i> )
<b>ZN</b>	Coloração de Ziehl–Neelsen ( <i>Ziehl–Neelsen</i> )

<b>HMLP</b>	Perceptron multicamadas híbrido ( <i>Hybrid Multilayered Perceptron</i> )
<b>IBSI</b>	Iniciativa de Padronização de Biomarcadores de Imagem ( <i>Image Biomarker Standardisation Initiative</i> )
<b>CT</b>	Tomografia computadorizada ( <i>Computed Tomography</i> )
<b>KNN</b>	K-vizinhos mais próximos ( <i>K-Nearest Neighbors</i> )
<b>SVM</b>	Máquina de vetores de suporte ( <i>Support Vector Machine</i> )
<b>LR</b>	Regressão logística ( <i>Logistic Regression</i> )
<b>NTM</b>	Micobactérias não tuberculosas ( <i>Nontuberculous Mycobacteria</i> )

# Sumário

1	Introdução	14
1.1	Objetivo Geral	15
1.2	Objetivos específicos	16
1.3	Organização do Trabalho	16
2	Revisão da Literatura	18
2.1	Automatic Bacillus Detection in Light Field Microscopy Images Using Convolutional Neural Networks and Mosaic Imaging Approach (Serrão et al., 2020)	18
2.2	Detection of Mycobacterium tuberculosis in Ziehl–Neelsen Stained Tissue Images Using Zernike Moments and HMLP (Osman, Mashor & Laafar, 2010)	19
2.3	Automatic Classification of Light Field Smear Microscopy Patches Using Convolutional Neural Networks for Identifying Mycobacterium tuberculosis (Pérez López et al., 2017)	20
2.4	Identification of Tuberculosis Bacilli from Bright-Field Microscopic Sputum Smear Images Using U-Net and Random Forest (Greeshma & Vishnukumar, 2023)	20
2.5	PyRadiomics: Computational Radiomics System to Decode the Radiographic Phenotype (van Griethuysen et al., 2017)	21
2.6	Differentiating Nontuberculous Mycobacterium Pulmonary Disease from Pulmonary Tuberculosis Using CT Radiomics (Yan et al., 2021)	22
2.7	Discussão das Lacunas da Literatura	23
3	Fundamentos Teórico	25
3.1	Tuberculose	25
3.2	Baciloscopia de Campo Claro	27
3.3	Redes Neurais Convolucionais (CNNs)	28
3.3.1	Camadas Convolucionais	29
3.3.2	Camadas de Pooling	29
3.3.3	Funções de Ativação	30
3.3.4	Normalização	31
3.3.5	Regularização e Dropout	31
3.3.6	Otimização e Aprendizado	32
3.4	Métodos Supervisionados	32
3.4.1	Floresta Aleatória (RF)	33
3.4.2	Perceptron Multicamadas (MLP)	34
3.5	Radiômica	36
3.5.1	Segmentação e definição da ROI	36
3.5.2	Pré-processamentos: normalização, discretização e filtros	37
3.5.3	Categorias de características radiômicas	37
3.5.4	Alta dimensionalidade e seleção de características	37
3.5.5	Aplicação neste estudo	37
3.6	Seleção de Características	38
3.6.1	Remoção de Variância Zero	38

3.6.2 Filtragem por Correlação	39
3.6.3 Seleção Baseada em Importância de Árvores (Random Forest Importance)	39
3.6.4 Operador de Seleção e Menor Contração Absoluta (LASSO)	40
3.7 Métricas de Avaliação	40
3.7.1 Matriz de Confusão	41
3.7.2 Acurácia	41
3.7.3 Precisão	41
3.7.4 Sensibilidade	42
3.7.5 F1-score	42
3.7.6 Especificidade	42
3.7.7 Curva ROC e AUC	42
4 Materiais e Métodos	44
4.1 Materiais	44
4.1.1 Conjunto de Dados	44
4.1.2 Ambiente Computacional	45
4.2 Métodos	46
4.2.1 Balanceamento do Conjunto de Dados	46
4.2.2 Estratégias de Construção das Máscaras	46
4.2.3 Extração Radiômica	47
4.2.4 Seleção e Tratamento de Características	48
4.2.5 Modelos Supervisionados Utilizados	48
4.2.6 Métricas de Avaliação	50
5 Resultados e Discussões	51
5.1 Resultados da Rede Neural Convolucional (CNN)	51
5.2 Resultados dos Modelos Radiômicos com Máscaras Segmentadas	53
5.3 Resultados dos Modelos Radiômicos com Máscara Cheia	54
5.4 Discussão Comparativa entre as Abordagens	56
6 Conclusão	58
7 Notas	59
8 Referências Bibliográficas	60

# Capítulo 1

## Introdução

A tuberculose (TB) é uma doença infecciosa de grande impacto global, causada pela bactéria *Mycobacterium tuberculosis*. Apesar dos avanços no diagnóstico e no tratamento, a TB permanece entre as principais causas de morbidade e mortalidade em diversos países, sobretudo em regiões de baixa e média renda. Segundo estimativas recentes da Organização Mundial da Saúde (WHO, 2023), milhões de novos casos continuam a ser registrados anualmente, e o controle efetivo da doença segue sendo um desafio epidemiológico relevante. O diagnóstico precoce desempenha papel fundamental na interrupção da cadeia de transmissão e na redução de complicações clínicas, reforçando a necessidade de métodos diagnósticos acessíveis, precisos e reprodutíveis.

O método diagnóstico mais difundido para detecção da tuberculose pulmonar é a baciloscopia de campo claro do escarro, amplamente utilizada devido ao baixo custo e à simplicidade de implementação em laboratórios de rotina. No entanto, trata-se de uma técnica altamente dependente da experiência do microscopista, exigindo atenção contínua e familiaridade com a morfologia dos bacilos álcool-ácido resistentes. O processo é demorado e sujeito à fadiga visual, além de vulnerável à variabilidade intra e interobservador, o que pode comprometer a sensibilidade e a padronização dos resultados. Essas limitações motivam o desenvolvimento de abordagens automatizadas.

Na literatura, observa-se um número crescente de investigações dedicadas à automação da detecção de bacilos utilizando redes neurais convolucionais (CNNs). Trabalhos como os de Serrão *et al.* (2020), Pérez López *et al.* (2017) e Kuok *et al.* (2019) demonstraram que arquiteturas profundas são capazes de identificar *Mycobacterium tuberculosis* com alta acurácia em imagens de baciloscopia de campo claro. Da mesma forma, Greeshma & Vishnukumar (2023) aplicaram U-Net para segmentação dos bacilos combinada a Floresta Aleatória (RF) para classificação, reforçando o potencial das técnicas de aprendizado profundo para mitigar limitações inerentes à análise manual.

Paralelamente às abordagens de aprendizado profundo, os métodos radiômicos têm ganhado destaque por sua capacidade de quantificar padrões texturais, estatísticos e estruturais presentes nas imagens, gerando descritores interpretáveis e potencialmente

complementares aos modelos baseados em CNNs. A radiômica parte do princípio de que alterações biológicas podem se refletir em variações mensuráveis na distribuição de intensidades e nas relações espaciais dos pixels, permitindo a extração sistemática de um grande número de características. Essa abordagem tem sido amplamente discutida por Lambin *et al.* (2017), e ferramentas consolidadas como o PyRadiomics (van Griethuysen *et al.*, 2017) já demonstraram eficácia em aplicações como predição de desfechos clínicos, diferenciação de padrões patológicos e análise de heterogeneidade tumoral (Aerts *et al.*, 2014; Yan *et al.*, 2021). No contexto da baciloscopia, a análise radiômica pode capturar heterogeneidades e informações sutis frequentemente imperceptíveis à inspeção humana, oferecendo uma perspectiva complementar aos modelos *end-to-end*.

Nesse cenário, o presente trabalho realiza uma análise comparativa entre diferentes arquiteturas de redes neurais e técnicas radiômicas para a detecção de bacilos em baciloscopias de campo claro. Utilizou-se um conjunto de dados composto por 31.484 recortes (*patches*) de imagens de microscopia, sendo 2.636 contendo bacilos (positivas) e 28.848 correspondentes ao fundo ou artefatos (negativas). Posteriormente, o conjunto foi balanceado para 2.635 imagens por classe. Dois *pipelines* radiômicos foram avaliados: o primeiro empregando máscaras segmentadas apenas para imagens positivas, combinadas a máscaras cheias para as negativas — nas quais a totalidade da área da imagem é considerada como Região de Interesse (ROI) para a extração de características; e o segundo utilizando máscaras cheias em todas as imagens. Em ambos os casos, as características extraídas pelo PyRadiomics passaram por filtragens sucessivas, incluindo remoção de variância zero, análise de correlação e ranqueamento por importância utilizando Random Forest. Em paralelo, uma CNN foi treinada diretamente nas imagens brutas, permitindo comparar uma abordagem ponta a ponta com modelos baseados em atributos radiômicos. Essa estratégia possibilita avaliar o impacto da segmentação, dos métodos de extração de características e da presença de potenciais vieses estruturais — como o vazamento de informação decorrente de atributos dependentes da máscara — oferecendo uma análise abrangente sobre o uso combinado de radiômica e redes neurais na detecção automatizada de bacilos.

## 1.1 Objetivo Geral

Avaliar e comparar o desempenho de diferentes abordagens baseadas em redes neurais e extração radiômica para a detecção de bacilos em imagens de baciloscopia de campo claro, analisando como distintas estratégias de definição de máscaras influenciam a qualidade das

características geradas e o desempenho final dos modelos.

## 1.2 Objetivos específicos

- Comparar o desempenho de diferentes arquiteturas de redes neurais, incluindo uma CNN treinada do zero e modelos MLP e Random Forest treinados a partir de características radiômicas.
- Avaliar a influência das estratégias de definição de máscara na qualidade das características radiômicas e no desempenho dos modelos, incluindo a identificação de potenciais fontes de viés ou vazamento de informação.

## 1.3 Organização do Trabalho

Este Trabalho de Conclusão de Curso está estruturado em seis capítulos, além dos elementos pré-textuais e pós-textuais.

- O Capítulo 1 – Introdução apresenta o contexto do estudo, a motivação, o problema de pesquisa, os objetivos geral e específicos e a organização do trabalho.
- O Capítulo 2 – Revisão da Literatura reúne os principais estudos relacionados à detecção automatizada de bacilos em baciloscopias de campo claro, abordando técnicas radiômicas, métodos de aprendizado profundo e outras abordagens relevantes presentes na literatura científica.
- O Capítulo 3 – Fundamentos Teóricos descreve os conceitos necessários para compreensão do estudo, incluindo aspectos relacionados à tuberculose, ao exame de baciloscopia de campo claro, às redes neurais convolucionais, aos métodos supervisionados, às técnicas radiômicas, aos procedimentos de seleção de características e às métricas de avaliação empregadas.
- O Capítulo 4 – Materiais e Métodos detalha o conjunto de dados utilizado, os processos de extração radiômica, as estratégias de definição das máscaras, o pipeline de seleção de características, as arquiteturas de redes neurais aplicadas e os procedimentos experimentais adotados.
- O Capítulo 5 – Resultados e Discussões apresenta os resultados obtidos com cada abordagem, comparando o desempenho dos modelos baseados em características radiômicas e da rede neural convolucional, além de discutir o impacto das decisões metodológicas.

- O Capítulo 6 – Conclusão sintetiza as conclusões do estudo, aponta limitações, expõe as contribuições da pesquisa e sugere direções para trabalhos futuros.

Ao final, são apresentadas notas e referências bibliográficas utilizadas ao longo do trabalho.

# Capítulo 2

## Revisão da Literatura

Este capítulo apresenta a revisão bibliográfica sobre detecção automatizada de bacilos em imagens de baciloscopia de campo claro, com foco em métodos de aprendizado de máquina, redes neurais convolucionais e técnicas radiômicas aplicadas ao diagnóstico da tuberculose. Foram selecionados artigos provenientes de bases como IEEE Xplore, PubMed e ScienceDirect, priorizando estudos que empregam modelos supervisionados na identificação de *Mycobacterium tuberculosis*, seja por meio de classificação direta, segmentação ou extração de características. A análise concentrou-se nos conjuntos de dados utilizados, nos métodos empregados e nas métricas de desempenho reportadas.

### **2.1 Automatic Bacillus Detection in Light Field Microscopy Images Using Convolutional Neural Networks and Mosaic Imaging Approach (Serrão *et al.*, 2020)**

Serrão *et al.* (2020) propõem a detecção automática de bacilos utilizando CNNs aplicadas a patches e mosaicos de baciloscopia de campo claro. A base de dados, fornecida pelo grupo de pesquisa da UFAM, contém inicialmente 13.977 *patches* de 40×40 pixels, corados pela técnica de Kinyoun. Após a remoção de amostras com artefatos, restaram 9.700 patches positivos e 25.000 negativos. Para cada *patch* positivo, foi disponibilizada uma máscara binária segmentando o bacilo; para negativos, a máscara é totalmente branca.

Além dos *patches*, foram gerados 5.000 mosaicos de 400×400 pixels, cada um composto por 100 *patches* (aprox. 50% positivos). Os *patches* foram divididos em 50% treino, 25% validação e 25% teste; os mosaicos em 60%/20%/20%.

Três arquiteturas de CNN (CNN1, CNN2 e CNN3) foram avaliadas, todas baseadas em blocos Conv 3×3 + *BatchNorm* + ReLU, *max pooling* 2×2 e convoluções transpostas 4×4 para *upsampling*. A CNN1 é a mais profunda (4 etapas de redução e 4 de expansão), a CNN2 intermediária e a CNN3 a mais compacta. O treinamento utilizou taxa de aprendizado 0,001, *batch size* 10, 20 épocas e parada após 6.000 iterações, com três otimizadores: SGDM, RMSProp e Adam.

A CNN1 apresentou o melhor desempenho: na validação, obteve 99,649% de

acurácia, 99,394% de precisão, 99,908% de sensibilidade, 99,388% de especificidade e *F1-score* de 99,650%. No teste, alcançou 99,665% de acurácia e *F1-score* de 99,667%. As CNN2 e CNN3 tiveram desempenho 1–4% inferior, mantendo sensibilidade acima de 99%.

O estudo demonstra que arquiteturas mais profundas produzem melhor detecção e menor ruído nas segmentações.

## **2.2 Detection of Mycobacterium tuberculosis in Ziehl–Neelsen Stained Tissue Images Using Zernike Moments and HMLP (Osman, Mashor & Laafar, 2010)**

Osman *et al.* (2010) apresentaram um método para detecção de bacilos em cortes histológicos corados por Ziehl–Neelsen (ZN) utilizando uma combinação de segmentação baseada em cor, momentos de Zernike e um classificador Perceptron Multicamadas Híbrido (HMLP). O estudo utilizou 15 lâminas histológicas, das quais foram obtidas aproximadamente 450–750 imagens (30 a 50 por lâmina), capturadas em 800×600 pixels, 24 bits, em microscópio Nikon Eclipse 80i (40×).

A metodologia inicia com um filtro inicial em espaço de cor C–Y, preservando tonalidades vermelhas características dos bacilos. Em seguida, a segmentação é realizada com *moving k-means* (k=2), utilizando o componente G (RGB) e o componente S (C–Y), permitindo lidar com variações de intensidade do vermelho causadas pela coloração manual. Após a clusterização, são aplicados pós-processamentos, incluindo mediana 5×5, *region growing* e remoção de regiões menores que 50 px ou maiores que 800 px.

Essas etapas resultaram em 607 objetos segmentados, subdivididos em 302 “*definite TB*” (bacilos isolados) e 305 “*possible TB*” (bacilos sobrepostos). Para caracterização, foram calculados momentos de Zernike até a 3ª ordem ( $S_1$ – $S_6$ ), devido à invariância a rotação, escala e translação.

O classificador empregado foi o HMLP, treinado com aproximadamente 200 amostras por classe. A melhor combinação de características ( $S_1 + S_2 + S_3$ ) obteve 98,07% de acurácia, 100% de sensibilidade e 96,19% de especificidade. O HMLP apresentou desempenho parecido ao MLP convencional, porém com menos neurônios ocultos e menos iterações, demonstrando maior eficiência.

### **2.3 Automatic Classification of Light Field Smear Microscopy Patches Using Convolutional Neural Networks for Identifying Mycobacterium tuberculosis (Pérez López *et al.*, 2017)**

Pérez López *et al.* (2017) propõem um método de classificação de bacilos em patches de baciloscopia obtidos por microscopia de campo claro com profundidade de foco estendida. O conjunto de dados utilizado, derivado do *TBimages*, contém 9.770 patches de 40×40 pixels, balanceados entre positivos (com um bacilo centralizado) e negativos (sem bacilos ou com pequenos fragmentos). As imagens originais foram obtidas a partir de 492 campos, combinando 11 planos focais, e todas as regiões contendo bacilos foram anotadas por especialistas. A versão aumentada do conjunto, com rotações de 90° e 180°, totalizou 29.310 *patches*, mantendo o balanceamento.

Três versões de entrada foram avaliadas: RGB, *grayscale* e R–G. Os autores testaram três arquiteturas de CNN (com 1, 2 e 3 camadas convolucionais), todas com ativação ReLU e camada de saída sigmoid. O treinamento utilizou 1.000 épocas, taxa de aprendizado inicial de 0,03 (reduzida até 0,0001), normalização dos patches, atualização por Nesterov *momentum* (0,9 a 0,999) e regularização via aumento de dados e *dropout*.

Os melhores resultados foram obtidos pelo modelo com três camadas convolucionais usando a entrada R–G, alcançando AUC de 99% e acurácia superior a 97%. Sem normalização, os modelos não aprendiam adequadamente (AUC  $\approx$  50%). Os autores destacam que pequenas variações cromáticas entre bacilos e fundo favorecem o uso de representações diferenciais (como R–G), e que modelos relativamente simples podem atingir alto desempenho em patches balanceados.

### **2.4 Identification of Tuberculosis Bacilli from Bright-Field Microscopic Sputum Smear Images Using U-Net and Random Forest (Greeshma & Vishnukumar, 2023)**

Este estudo propõe um *pipeline* em duas etapas para detecção de bacilos em baciloscopias de campo claro, combinando segmentação U-Net e classificação por *Random Forest*. O conjunto de dados utilizado contém 120 imagens RGB de baciloscopia coradas por Ziehl–Neelsen, capturadas com microscópio Zeiss Axioshop 40, aumento 100× e NA = 1.25. As imagens possuem dimensão de 2816 × 2112 pixels e incluem máscaras de ground-truth.

Na primeira etapa, as imagens são segmentadas usando U-Net, arquitetura

*encoder-decoder* composta por quatro blocos de contração e quatro de expansão, empregando *upsampling* para manter maior dimensão espacial. A U-Net foi adotada devido à sua capacidade de segmentar estruturas biológicas mesmo com conjuntos de dados reduzidos. Após a segmentação, regiões suspeitas são extraídas com base em contornos e área mínima de 590 pixels, valor definido empiricamente.

Cada região segmentada gera um *patch*, posteriormente classificado por um *Random Forest*, escolhido por sua robustez e melhor generalização quando comparado a regressão logística e *ridge classifier*. O classificador identifica *patches* contendo bacilos ou não-bacilos a partir de características visuais extraídas das regiões segmentadas.

Nos experimentos, o *Random Forest* apresentou o melhor desempenho entre os métodos avaliados, atingindo 93,98% de acurácia, precisão de 0,943, sensibilidade de 0,967 e *F1-score* de 0,946. A matriz de confusão reportada inclui 180 TPs, 14 FNs, 5 FPs e 456 TNs. Os autores ressaltam que a qualidade da segmentação influencia diretamente o desempenho final e destacam limitações do método em casos de bacilos sobrepostos ou aglomerados.

## **2.5 PyRadiomics: Computational Radiomics System to Decode the Radiographic Phenotype (van Griethuysen *et al.*, 2017)**

Van Griethuysen *et al.* (2017) introduziram o PyRadiomics, atualmente a principal ferramenta open-source para extração padronizada de características radiômicas. O objetivo do trabalho é estabelecer um pipeline reproduzível e validado para quantificar padrões de intensidade, textura e forma em imagens médicas, seguindo as recomendações do *Image Biomarker Standardization Initiative* (IBSI).

O sistema implementa um conjunto abrangente de descritores, incluindo estatísticas de primeira ordem, GLCM, GLRLM, GLSZM, GLDM e NGTDM, além de filtros como LoG e *wavelets*. A ferramenta opera sobre imagens e respectivas máscaras, permitindo análise bidimensional (a análise considera apenas os pixels vizinhos dentro de um mesmo plano horizontal e vertical (x, y)) ou tridimensional (a vizinhança de um ponto leva em conta pixels nas fatias superiores e inferiores (x, y, z)), e permitindo controlar parâmetros críticos como normalização de intensidade, discretização (*bin width*), interpolação e pré-processamento.

O artigo apresenta um *pipeline* completo composto por: (1) aquisição e segmentação da região de interesse; (2) pré-processamentos padronizados; (3) extração sistemática das características; e (4) análise estatística e modelagem preditiva. O PyRadiomics é validado em múltiplos conjuntos de dados clínicos heterogêneos, demonstrando baixo erro de

reprodutibilidade e consistência na extração das características entre diferentes softwares e configurações.

A contribuição central do trabalho é fornecer um *framework* robusto e transparente para estudos radiômicos, reduzindo variabilidade entre implementações e permitindo comparabilidade entre estudos. Desde sua publicação, o PyRadiomics tornou-se referência obrigatória em pesquisas envolvendo quantificação radiômica, sendo amplamente adotado em tarefas de classificação, segmentação, prognóstico e detecção de doenças.

## **2.6 Differentiating Nontuberculous Mycobacterium Pulmonary Disease from Pulmonary Tuberculosis Using CT Radiomics (Yan *et al.*, 2021)**

Yan *et al.* (2021) desenvolveram um modelo radiômico para diferenciar doença pulmonar por micobactérias não tuberculosas (NTM) de tuberculose pulmonar (PTB) por meio da análise de cavidades pulmonares em tomografias computadorizadas. O estudo retrospectivo envolveu 142 pacientes na coorte principal — 73 NTM e 69 PTB — provenientes de dois hospitais chineses (*Shandong Chest Hospital* e *Qilu Hospital*). Para validação externa, foram incluídos mais 40 pacientes (20 NTM e 20 PTB) de um terceiro centro. No total, foram segmentadas 379 cavidades em CT por dois radiologistas experientes, utilizando cada cavidade como ROI independente. A divisão do conjunto de dados considerou 80% para treino e 20% para validação interna.

Foram extraídas 1.409 características radiômicas por ROI, através da plataforma Huiying Radcloud — incluindo estatísticas de primeira ordem, texturas (GLCM, GLRLM, GLSZM, NGTDM), forma e filtros (como *wavelet*). A seleção de atributos ocorreu em múltiplas etapas: variância mínima, método *K-best* e regressão LASSO, reduzindo o conjunto para 29 características ótimas. Seis classificadores supervisionados foram avaliados: KNN, SVM, XGBoost, *Random Forest*, *Logistic Regression* (LR) e *Decision Tree*.

Nos resultados, todos os modelos obtiveram  $AUC > 0.97$  no treino, com sensibilidade e especificidade acima de 92%. Na validação interna, as seis abordagens mantiveram desempenho elevado ( $AUC > 0.84$ , sensibilidade  $> 0.79$ , especificidade  $> 0.79$ ). Na validação externa, os modelos também apresentaram  $AUC > 0.84$ , sendo o classificador *Logistic Regression* o melhor, alcançando precisão de 0.92, sensibilidade de 0.94 e *F1-score* de 0.93. Os autores destacam que padrões texturais das cavidades diferem significativamente entre NTM e PTB, e que a radiômica superou o desempenho dos radiologistas na distinção entre as duas patologias.

## 2.7 Discussão das Lacunas da Literatura

A revisão da literatura evidencia avanços significativos na detecção automatizada de bacilos e na aplicação de técnicas de análise de imagem ao diagnóstico da tuberculose. Contudo, observa-se que cada linha de pesquisa aborda apenas partes isoladas do problema, deixando lacunas importantes que justificam o presente estudo.

Primeiramente, os métodos baseados em CNNs para baciloscopia de campo claro (Serrão *et al.*, 2020; Pérez López *et al.*, 2017; Greeshma & Vishnukumar, 2023) concentram-se exclusivamente em classificação direta de patches ou em segmentação seguida de classificação, mas não exploram a extração de características radiômicas como alternativa ou complemento à aprendizagem profunda. Embora alcancem alto desempenho, esses modelos operam como caixas-pretas, limitando a interpretabilidade e dificultando comparações estruturadas entre diferentes formas de representação do sinal.

Em segundo lugar, estudos que aplicam radiômica à tuberculose (Yan *et al.*, 2021) abordam apenas imagens de tomografia computadorizada, com ênfase em cavidades e texturas pulmonares. Não existe, até o momento, aplicação consolidada de radiômica a imagens de baciloscopia de campo claro, nem trabalhos que comparem diretamente CNNs e pipelines radiômicos sobre o mesmo dataset microscópico.

Ainda não há estudos que utilizem máscaras segmentadas de bacilos como referência para extração radiômica, nem análises sobre o impacto de usar máscaras completas versus segmentadas — aspecto crítico devido ao risco de vazamento de informação, especialmente em atributos dependentes do tamanho da máscara.

Por fim, nenhum dos trabalhos revisados implementa uma comparação sistemática entre abordagens ponto a ponto (CNNs) e modelos supervisionados baseados em características radiômicas, o que reforça a relevância deste estudo. A seguir é apresentado na Tabela 1 um resumo da revisão da literatura.

Tabela 1: Tabela de revisão da literatura

Estudo	Dataset	Técnica / Metodologia	Arquitetura / Modelo
Serrão <i>et al.</i> (2020)	9.700 <i>patches</i> positivos e 25.000 negativos (40×40); 5.000 mosaicos 400×400; corante Kinyoun	Classificação de <i>patches</i> e mosaicos	Três CNNs (CNN1, CNN2) com conv + BN + ReLU, e deconvoluções
Osman <i>et al.</i> (2010)	450–750 imagens histológicas ZN (800×600); 607 objetos segmentados	Segmentação por cor + Zernike Moments + HMLP	HMLP ( <i>Hybrid Multilayer Perceptron</i> )
Pérez López <i>et al.</i> (2017)	9.770 <i>patches</i> 40×40 (balanceados); versão aumentada 29.310 <i>patches</i>	Classificação de <i>patches</i> (RGB, grayscale, R-G)	CNN com 1, 2 e 3 camadas convolucionais
Greeshma & Vishnukumar (2023)	120 imagens RGB (2816×2112); máscaras GT; corante ZN	U-Net para segmentação + RF para classificação de <i>patches</i>	U-Net (4 <i>encoders</i> + 4 <i>decoders</i> ) + <i>Random Forest</i>
van Griethuysen <i>et al.</i> (2017)	Múltiplos conjuntos de dados clínicos (diversas modalidades)	Extração padronizada de características radiômicas (IBSI)	PyRadiomics (1st order, GLRLM, GLSZM, GLDM, e filtros)
Yan <i>et al.</i> (2021)	73 NTM + 69 PTB (coorte principal); 40 pacientes validação externa; 379 cavidades segmentadas	Radiômica + seleção ANOVA, K-best e LASSO	6 classificadores: KNN, S, XGBoost, RF, LR, DT

# Capítulo 3

## Fundamentos Teórico

O presente capítulo reúne os principais conceitos que sustentam as abordagens desenvolvidas neste estudo. Nos últimos anos, o uso de métodos computacionais no apoio ao diagnóstico da tuberculose tem evoluído de forma expressiva, impulsionado pela digitalização dos exames microscópicos e pela consolidação de técnicas de aprendizado supervisionado capazes de identificar padrões visuais complexos em dados biomédicos (Serrão *et al.*, 2020; Greeshma & Vishnukumar, 2023).

Neste trabalho, investiga-se a detecção de *Mycobacterium tuberculosis* em imagens de baciloscopia de campo claro por meio de diferentes estratégias de modelagem: uma rede neural convolucional aplicada diretamente às imagens, além de modelos supervisionados — Floresta Aleatória e Perceptron Multicamadas — treinados a partir de características radiômicas extraídas com o PyRadiomics. A adoção simultânea dessas abordagens permite comparar o desempenho entre métodos baseados em representação automática (CNN) e métodos guiados por descritores manuais (*radiomics*), destacando vantagens e limitações de cada paradigma no contexto da baciloscopia.

Para fornecer o embasamento conceitual necessário à interpretação dos resultados, este capítulo apresenta uma visão geral da tuberculose e do exame de baciloscopia, discute os fundamentos das redes neurais convolucionais, revisa os princípios dos classificadores supervisionados empregados nos experimentos e descreve os conceitos essenciais da radiômica, incluindo extração e seleção de características. Por fim, são introduzidas as métricas de avaliação utilizadas para mensurar o desempenho dos modelos, seguindo recomendações amplamente aceitas na literatura de aprendizagem de máquina aplicada à saúde.

### 3.1 Tuberculose

A tuberculose (TB) é uma doença infecciosa causada pelo complexo *Mycobacterium tuberculosis*, permanecendo como um dos principais desafios globais de saúde pública. De acordo com o relatório mais recente da Organização Mundial da Saúde, mais de 10 milhões de novos casos são registrados anualmente, com aproximadamente 1,3 milhão de mortes,

configurando-se entre as doenças infecciosas com maior impacto socioeconômico no mundo (WHO, 2023). A transmissão ocorre predominantemente pela via aérea, a partir da inalação de aerossóis contendo bacilos expelidos por indivíduos infectados, especialmente na forma pulmonar, que representa cerca de 85% dos casos notificados (Pai *et al.*, 2016).

O diagnóstico da TB pulmonar tradicionalmente envolve um conjunto de métodos clínicos, laboratoriais e de imagem. Entre eles, a baciloscopia de campo claro, com coloração de Ziehl–Neelsen ou Kinyoun, permanece amplamente utilizada em países de baixa e média renda devido ao baixo custo, rapidez e simplicidade operacional (Balcells *et al.*, 2019). Nesse exame, bacilos álcool-ácido resistentes aparecem como estruturas alongadas e avermelhadas contrastando com o fundo azul ou lilás da lâmina. Apesar de ser um método consolidado, a baciloscopia apresenta limitações importantes, sobretudo a variabilidade interobservador e a queda de sensibilidade em casos paucibacilares ou em amostras com artefatos (Steingart *et al.*, 2006).

Com o avanço tecnológico, diferentes métodos complementares foram introduzidos, como a cultura em meios sólidos ou líquidos e os testes moleculares rápidos, a exemplo do GeneXpert MTB/RIF. Embora possuam sensibilidade superior, esses métodos demandam maior infraestrutura e custos operacionais elevados, o que dificulta sua implementação universal, especialmente em regiões remotas ou com limitações de recursos (Nahid *et al.*, 2022). Assim, a baciloscopia continua sendo o primeiro exame utilizado em grande parte dos sistemas de saúde, reforçando a necessidade de aprimorar sua acurácia por meio de ferramentas computacionais.

Nesse contexto, técnicas de processamento de imagens e aprendizado de máquina têm sido exploradas como alternativas promissoras para mitigar limitações do diagnóstico tradicional. Estudos demonstram que a automação da detecção de bacilos reduz substancialmente erros de contagem, aumenta a reprodutibilidade e pode acelerar o fluxo laboratorial, especialmente quando integrada a microscopia digital ou sistemas de varredura automatizada (Costa Filho *et al.*, 2017). Redes neurais convolucionais (CNNs), em particular, têm apresentado desempenho superior na classificação de patches e na segmentação de bacilos, aproximando-se de especialistas humanos.

Apesar dos avanços em aprendizado profundo, abordagens tradicionais de engenharia manual de características — como a radiômica — também ganham relevância no estudo da TB. A radiômica tem se mostrado eficiente na descrição quantitativa de padrões morfológicos e texturais em imagens clínicas, permitindo análises objetivas de estruturas biológicas (van

Griethuysen *et al.*, 2017). Embora ainda pouco explorada diretamente em baciloscopias, ela oferece potencial para caracterizar regiões suspeitas, compreender padrões de textura relacionados ao bacilo e comparar desempenhos entre modelos baseados em imagens cruas e modelos baseados em atributos extraídos.

## 3.2 Baciloscopia de Campo Claro

A baciloscopia de campo claro permanece como um dos métodos mais difundidos para o diagnóstico da tuberculose pulmonar, especialmente em países de baixa e média renda. Trata-se de uma técnica microscópica que identifica bacilos álcool-ácido resistentes (BAAR) a partir de esfregaços de escarro corados, tradicionalmente pela técnica de Ziehl–Neelsen. Por ser simples, rápida e de baixo custo, integra o conjunto de exames recomendados pela Organização Mundial da Saúde há décadas, sendo amplamente utilizada como porta de entrada para a confirmação inicial da doença (WHO, 2023; Steingart *et al.*, 2006).

O princípio da técnica baseia-se na resistência dos bacilos ao álcool-ácido após a coloração com fucsina fenicada, característica que permite distingui-los visualmente como estruturas alongadas e avermelhadas sobre um fundo azul ou roxo, dependendo do método de contra-coloração. A morfologia típica inclui formato bastonete, extremidades arredondadas e comprimento variando entre 2 e 4  $\mu\text{m}$ , embora artefatos, detritos celulares e até estruturas celulares degradadas podem, ocasionalmente, mimetizar bacilos verdadeiros, dificultando a distinção visual em lâminas pouco contrastadas.. Essa semelhança morfológica representa um desafio importante, contribuindo para a variabilidade interobservador.

Apesar de sua ampla adoção, a baciloscopia apresenta limitações notáveis. Sua sensibilidade é relativamente baixa, sobretudo em casos paucibacilares, pacientes vivendo com HIV ou indivíduos em fases iniciais da infecção, com estimativas variando entre 50% e 60% quando comparada à cultura como padrão-ouro (Balcells *et al.*, 2019). Além disso, fatores como qualidade da amostra, preparo da lâmina, variações de coloração e fadiga do observador influenciam significativamente o desempenho diagnóstico, especialmente em regiões com alta demanda e infraestrutura limitada (Chakravorty *et al.*, 2017).

Com a digitalização da microscopia e a crescente adoção de câmeras acopladas a microscópios convencionais, a baciloscopia tem sido apropriada pelo campo da análise computacional de imagens. Diversos estudos têm demonstrado que a varredura automatizada e os algoritmos de detecção auxiliada por computador reduzem erros humanos e aumentam a reprodutibilidade, abrindo caminho para sistemas que combinam algoritmos clássicos de

segmentação, aprendizado de máquina e redes neurais profundas (Costa Filho *et al.*, 2017). Além de aumentar a eficiência do processo, essas abordagens apresentam potencial para padronizar critérios diagnósticos em diferentes contextos laboratoriais.

A estrutura visual das lâminas de baciloscopia também favorece a aplicação de técnicas baseadas em radiômica, uma vez que os bacilos diferem nitidamente do fundo em termos de textura, intensidade e padrões morfológicos.

Embora a radiômica seja amplamente aplicada em exames tomográficos e de ressonância magnética, sua aplicação em microscopia óptica começa a emergir como uma estratégia promissora, capaz de extrair descritores quantitativos que auxiliam na identificação de bacilos em ambientes complexos e com ruído visual (van Griethuysen *et al.*, 2017). Dessa forma, o estudo radiômico da baciloscopia fortalece o avanço de modelos interpretáveis e comparáveis a métodos tradicionais.

### 3.3 Redes Neurais Convolucionais (CNNs)

As redes neurais convolucionais (*Convolutional Neural Networks* – CNNs) representam um dos pilares do aprendizado profundo aplicado à análise de imagens. Propostas inicialmente por LeCun *et al.* (1989) e popularizadas com o sucesso do modelo AlexNet no ImageNet em 2012 (Krizhevsky, Sutskever & Hinton, 2012), essas arquiteturas foram projetadas para explorar padrões espaciais, como bordas, texturas e formas, de maneira hierárquica. Sua capacidade de aprender automaticamente representações discriminativas tem impulsionado avanços significativos em aplicações biomédicas, incluindo diagnóstico assistido por imagem, detecção de lesões e reconhecimento de estruturas celulares.

No contexto da baciloscopia de campo claro, as CNNs são particularmente relevantes porque conseguem identificar sutis variações morfológicas e cromáticas presentes nos bacilos da *Mycobacterium tuberculosis*. Diferentemente de métodos tradicionais baseados em engenharia manual de características, as CNNs extraem diretamente dos pixels padrões capazes de discriminar bacilos de artefatos, ruído ou fundo microscópico. Essa capacidade é especialmente valiosa em imagens microscópicas, onde contrastes são baixos e a variabilidade visual elevada.

A arquitetura de uma CNN é composta por camadas especializadas que realizam convoluções, agregação espacial, normalização e aprendizado não linear. Esses blocos são organizados de forma progressiva, permitindo que camadas iniciais capturem padrões

simples, enquanto camadas profundas modelam estruturas complexas. O treinamento é baseado em backpropagation e otimização estocástica, técnicas consolidadas na literatura (Goodfellow, Bengio & Courville, 2016), que ajustam os pesos da rede de forma a minimizar o erro de classificação.

### 3.3.1 Camadas Convolucionais

As camadas convolucionais constituem o núcleo das CNNs. Cada neurônio aplica um filtro (*kernel*) que percorre a imagem realizando multiplicações ponto a ponto, gerando mapas de ativação que destacam padrões locais, como bordas, pontos ou texturas. Esse mecanismo explora a estação espacial da imagem, reduzindo o número de parâmetros em comparação com redes totalmente conectadas (LeCun *et al.*, 1998).

Além disso, os filtros são aprendidos durante o treinamento, o que permite que a rede extraia automaticamente as características mais relevantes para a tarefa.

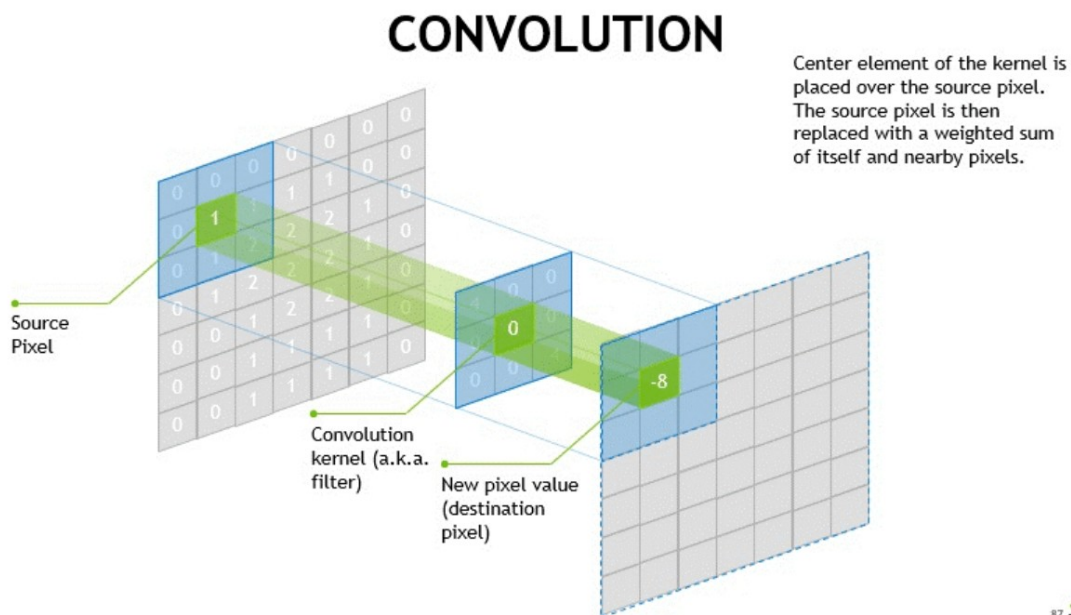


Figura 1: Funcionamento da camada convolutiva. Fonte: Nvidia (2025).

### 3.3.2 Camadas de Pooling

As camadas de pooling realizam operações de subamostragem, reduzindo a dimensão espacial dos mapas de ativação. O tipo mais comum é o *max pooling*, que seleciona o maior valor em regiões locais da imagem. Essa redução ajuda a controlar o sobreajuste, aumenta a

invariância a pequenas translações e diminui o custo computacional (Scherer, Müller & Behnke, 2010).

Embora *pooling* seja tradicional, arquiteturas modernas utilizam também *strided convolutions* como alternativa para redução espacial.

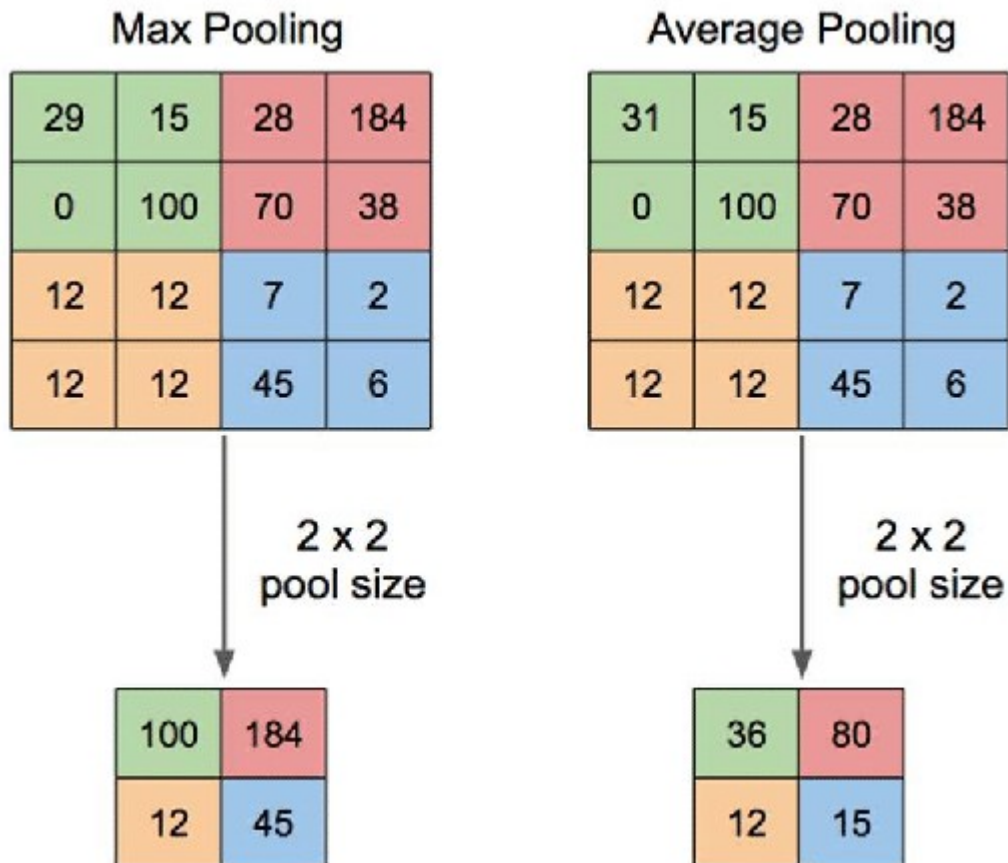


Figura 2: Ilustração do funcionamento da camada de *Max Pooling* e *Average Pooling*.  
Fonte: Muhamad Yani *et al* (2019).

### 3.3.3 Funções de Ativação

As funções de ativação introduzem não linearidade ao modelo, permitindo aprender representações mais complexas do que simples combinações lineares. A função mais utilizada é a ReLU (*Rectified Linear Unit*), popularizada por Nair & Hinton (2010), que acelera a convergência e simplifica o cálculo dos gradientes.

Variantes como Leaky ReLU e ELU podem ser empregadas para mitigar o problema de “neurônios mortos”, ampliando a estabilidade do treinamento.

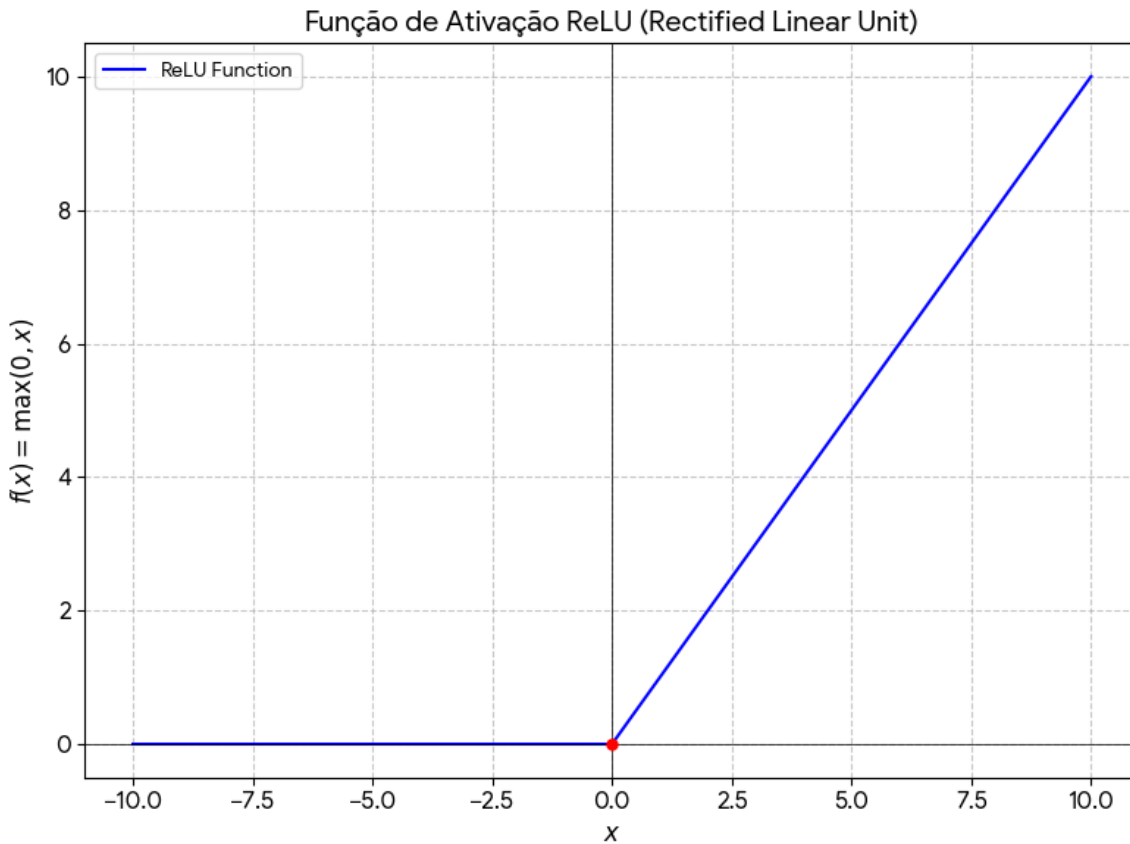


Figura 3: Gráfico da função de ativação ReLU.

### 3.3.4 Normalização

Para melhorar a estabilidade e acelerar o treinamento, técnicas de normalização como *Batch Normalization* (BN) são amplamente utilizadas. Introduzida por Ioffe e Szegedy (2015), essa técnica normaliza a ativação de cada camada com base no *mini-batch*, reduzindo problemas como *internal covariate shift*. Em muitas CNNs modernas, o uso de BN tornou-se padrão, permitindo treinar redes mais profundas com maior eficiência.

### 3.3.5 Regularização e Dropout

A regularização tem papel fundamental para evitar sobreajuste em modelos com grande número de parâmetros. Uma das técnicas mais comuns é o *Dropout*, introduzido por Srivastava *et al.* (2014), que desativa aleatoriamente neurônios durante o treinamento, forçando a rede a aprender representações redundantes e mais robustas.

## Funcionamento do Dropout em Redes Neurais

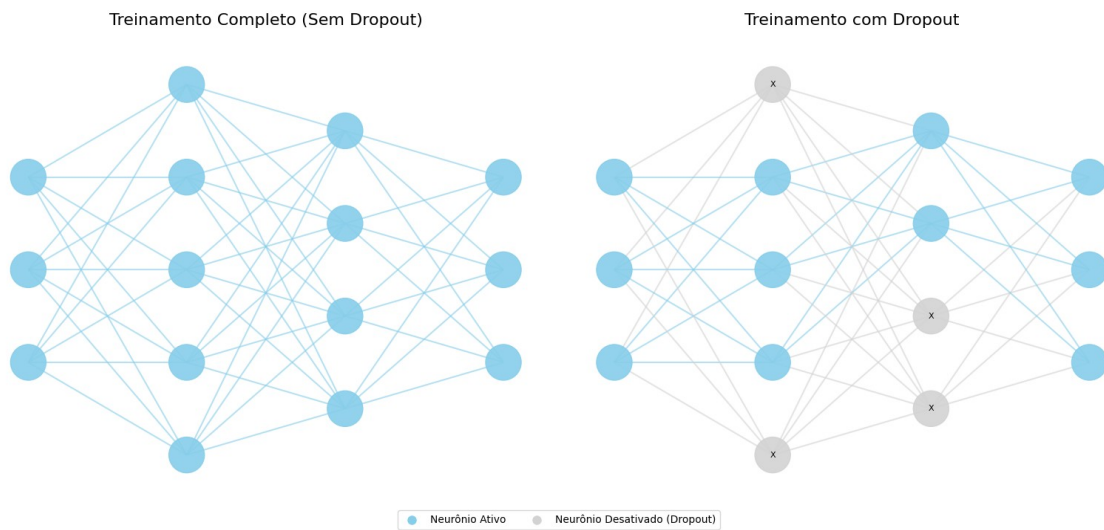


Figura 4: Método *Dropout*, proposto por Srivastava *et al.* (2014).

Outras formas incluem regularização L2, aumento de dados e parada antecipada, frequentemente aplicadas em classificação de imagens médicas.

### 3.3.6 Otimização e Aprendizado

O aprendizado em CNNs é realizado por otimização iterativa dos pesos por meio de algoritmos como SGD, SGDM ou Adam (Kingma & Ba, 2015). Esses métodos ajustam gradualmente os parâmetros para minimizar uma função de perda, geralmente baseada em entropia cruzada para problemas de classificação binária ou multiclasse.

A escolha do otimizador influencia diretamente a estabilidade do treinamento, a velocidade de convergência e o desempenho final.

## 3.4 Métodos Supervisionados

Os métodos supervisionados constituem a base do aprendizado de máquina aplicado à classificação de padrões, sobretudo em contextos onde há rótulos disponíveis para cada amostra do conjunto de dados. Nessas abordagens, o modelo aprende a associação entre atributos de entrada e uma variável-alvo, ajustando parâmetros internos de forma a minimizar o erro nas previsões. Em aplicações biomédicas, tais métodos são amplamente empregados em tarefas de diagnóstico, triagem e detecção de anomalias, devido à sua capacidade de generalização e à interpretabilidade variável entre diferentes algoritmos (Kononenko, 2001).

No contexto deste trabalho, modelos supervisionados foram aplicados às características radiômicas extraídas das imagens de baciloscopia de campo claro. Tais características, geradas pelo PyRadiomics, representam propriedades estatísticas e texturais capazes de captar padrões estruturais relevantes nas imagens. Para explorar diferentes perspectivas de modelagem, duas abordagens supervisionadas foram consideradas: Floresta Aleatória, um método baseado em aprendizado em conjunto de árvores de decisão, e Perceptron Multicamadas (MLP), uma rede neural totalmente conectada adequada para dados tabulares de alta dimensionalidade. Cada um desses modelos apresenta vantagens particulares — o RF destaca-se pela robustez e interpretabilidade, enquanto o MLP possui maior capacidade de modelar relações complexas e não lineares entre as variáveis.

### 3.4.1 Floresta Aleatória (RF)

O RF, introduzido por Breiman (2001), é um algoritmo de aprendizado em conjunto que combina múltiplas árvores de decisão para produzir classificações mais estáveis e generalizáveis. Cada árvore é treinada com um subconjunto aleatório das amostras (*bootstrap*) e utiliza subconjuntos aleatórios das variáveis em cada divisão, processo conhecido como *feature bagging*. Essa diversidade reduz a variância do modelo, tornando-o menos suscetível ao sobreajuste — um problema comum em árvores individuais.

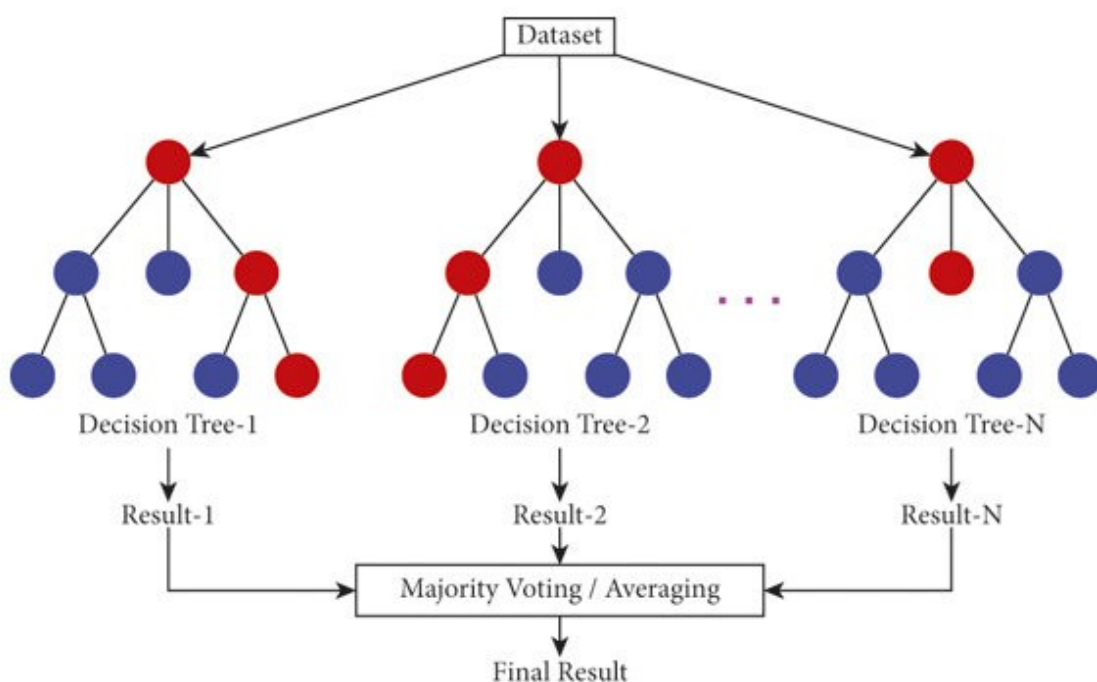


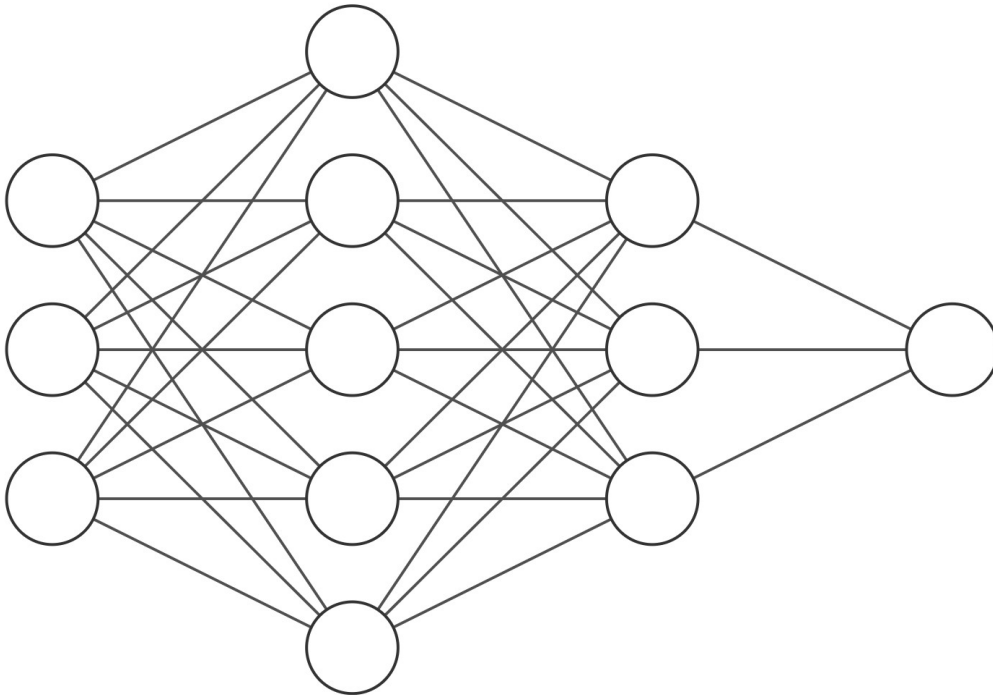
Figura 5: Topologia RF. Fonte: Researchgate (2025).

Uma das vantagens centrais do RF é sua capacidade de lidar com dados de alta dimensionalidade, característica frequente em estudos radiômicos, nos quais centenas ou milhares de atributos podem ser gerados a partir de um único ROI. Além disso, o algoritmo estima medidas internas de importância das variáveis, como importância de Gini e importância de permutação, permitindo interpretar quais características contribuem mais para a decisão do modelo (Geurts, Ernst & Wehenkel, 2006). Essa interpretabilidade é particularmente valiosa em aplicações clínicas, onde a transparência do processo de decisão pode ser determinante para a adoção do sistema.

Devido à sua robustez, tolerância a ruídos e bom desempenho em conjuntos pequenos ou moderados, o RF é amplamente empregado em tarefas biomédicas de classificação e diagnóstico, incluindo análises radiômicas em oncologia, doenças pulmonares e patologias infecciosas (Parmar *et al.*, 2015).

### 3.4.2 Perceptron Multicamadas (MLP)

O MLP é uma rede neural artificial composta por camadas totalmente conectadas, nas quais cada neurônio realiza uma combinação linear dos atributos de entrada seguida por uma função de ativação não linear. Essa arquitetura foi formalmente estabelecida por Rumelhart, Hinton e Williams (1986), com o desenvolvimento do algoritmo de retropropagação, que permite ajustar os pesos da rede de maneira eficiente.



Input Layer  $\in \mathbb{R}^3$     Hidden Layer  $\in \mathbb{R}^5$     Hidden Layer  $\in \mathbb{R}^3$     Output Layer  $\in \mathbb{R}^1$

Figura 6: Topologia MLP.

No caso de dados radiômicos, o MLP é particularmente adequado por sua capacidade de modelar relações não lineares complexas entre os atributos, capturando interações entre estatísticas de primeira ordem, texturas e características derivadas de transformações filtradas. Uma rede típica contém uma camada de entrada, uma ou mais camadas ocultas e uma camada de saída com ativação sigmoide ou softmax, conforme a natureza do problema. O processo de treinamento utiliza algoritmos de otimização baseados em gradiente, como Adam (Kingma & Ba, 2015), que aceleram a convergência mesmo em espaços de parâmetros de alta dimensionalidade.

Embora o MLP tenha menor interpretabilidade intrínseca quando comparado ao Random Forest, apresenta a vantagem de modelar relações altamente não lineares e de se ajustar de maneira flexível à distribuição dos dados. Em aplicações biomédicas, redes MLP têm sido amplamente empregadas na análise de preditores clínicos, sinais fisiológicos e atributos extraídos de imagens médicas, frequentemente alcançando desempenho competitivo, especialmente em cenários de alta dimensionalidade.

## 3.5 Radiômica

A radiômica consiste na extração quantitativa de atributos de intensidade, textura e forma a partir de imagens médicas, com o objetivo de revelar padrões não perceptíveis visualmente. Seu marco inicial é atribuído a Lambin *et al.* (2012), que demonstraram que características matemáticas podem auxiliar no diagnóstico e prognóstico em diferentes doenças. Posteriormente, Aerts *et al.* (2014) consolidaram o campo ao demonstrar que os descritores radiômicos são capazes de capturar heterogeneidade tumoral de maneira reprodutível.

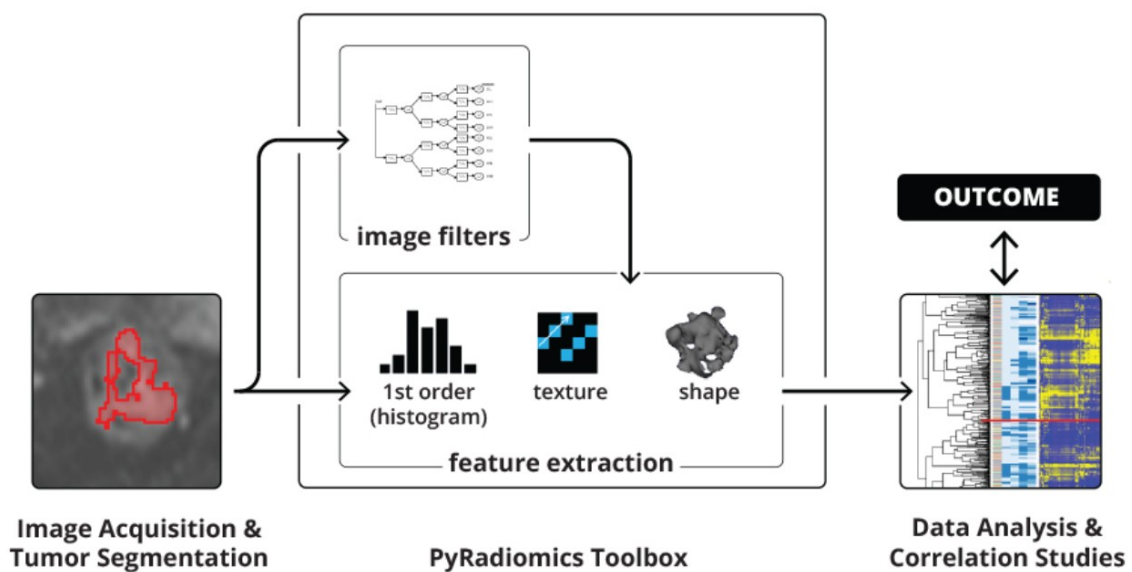


Figura 7: PyRadiomics pipeline. Fonte: PyRadiomics (2017).

### 3.5.1 Segmentação e definição da ROI

A etapa de segmentação é considerada crítica, pois todas as características dependem exclusivamente dos pixels contidos na região de interesse (ROI). Máscaras inconsistentes podem introduzir padrões artificiais relacionados ao tamanho ou à forma da ROI, comprometendo a validade do modelo (Zwanenburg *et al.*, 2020). Em imagens microscópicas, essa sensibilidade se intensifica, tornando obrigatório o controle rigoroso da segmentação.

## 3.5.2 Pré-processamentos: normalização, discretização e filtros

O pré-processamento radiômico inclui padronização de intensidade, interpolação e discretização. O *Image Biomarker Standardization Initiative* (IBSI) recomenda discretização por *fixed bin width* para garantir comparabilidade entre amostras (Zwanenburg *et al.*, 2020). Filtros como LoG e *wavelets* ajudam a destacar detalhes em diferentes escalas (Parekh & Jacobs, 2019), aumentando a diversidade de atributos disponíveis para os modelos.

## 3.5.3 Categorias de características radiômicas

As características extraídas por ferramentas como o PyRadiomics (van Griethuysen *et al.*, 2017) incluem:

- Primeira ordem: estatísticas básicas da distribuição de intensidades;
- GLCM, GLRLM, GLSZM, GLDM, NGTDM: texturas que capturam padrões espaciais e repetitividade;
- Wavelet/LoG: atributos filtrados em múltiplas escalas.

Essas categorias são amplamente utilizadas em estudos clínicos e apresentam capacidade reconhecida de capturar heterogeneidade estrutural em diferentes modalidades de imagem (Aerts *et al.*, 2014).

## 3.5.4 Alta dimensionalidade e seleção de características

A radiômica frequentemente gera centenas a milhares de atributos por ROI, o que introduz multicolinearidade e risco de sobreajuste. Assim, métodos de seleção de características — como variância mínima, correlação, LASSO e importância por RF — são essenciais para obter modelos generalizáveis (Vallières *et al.*, 2018).

## 3.5.5 Aplicação neste estudo

Neste trabalho, a radiômica é aplicada às imagens de baciloscopia para quantificar padrões visuais internos aos patches. Esse conjunto de atributos alimenta modelos supervisionados como RF e MLP, permitindo comparar diretamente o desempenho desses classificadores com o de uma CNN treinada nas imagens brutas. A seção também discute

como inconsistências na máscara podem gerar vazamento de informação — especialmente em atributos relacionados ao tamanho da ROI — reforçando a necessidade de controle metodológico rigoroso.

## 3.6 Seleção de Características

A seleção de características é uma etapa crucial em pipelines radiômicos, especialmente devido ao elevado número de atributos gerados (centenas ou milhares por ROI). Seu objetivo é reduzir dimensionalidade, remover redundâncias, evitar sobreajuste e melhorar tanto a interpretabilidade quanto o desempenho dos modelos supervisionados. Esse processo é amplamente discutido na literatura de aprendizado de máquina aplicado à saúde (Guyon & Elisseeff, 2003; Kursu & Rudnicki, 2010).

Diversas estratégias foram propostas para identificar subconjuntos relevantes de características, e sua escolha afeta diretamente a capacidade de generalização do modelo. Em radiômica, é comum combinar métodos estatísticos simples com algoritmos baseados em modelagem preditiva, como recomendado por Zwanenburg *et al.* (2020) e Parekh & Jacobs (2019).

A seguir, descrevem-se as abordagens mais relevantes, incluindo aquelas utilizadas neste estudo.

### 3.6.1 Remoção de Variância Zero

Características com variância zero apresentam o mesmo valor para todas as amostras, não contribuindo para distinção entre classes. A remoção é direta e amplamente recomendada como primeira etapa de filtragem (Guyon & Elisseeff, 2003).

Formalmente, uma feature  $x$  é removida se:

$$\text{Var}(x) = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 = 0$$

Essa etapa também elimina descritores artificiais introduzidos por máscaras inconsistentes, como atributos derivados do tamanho da ROI.

### 3.6.2 Filtragem por Correlação

A presença de multicolinearidade é comum em radiômica, especialmente entre texturas derivadas do mesmo filtro. Duas características altamente correlacionadas tendem a capturar a mesma informação, prejudicando modelos sensíveis a redundâncias (como MLPs).

Usa-se normalmente o coeficiente de Pearson:

$$r_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \cdot \sum_{i=1}^n (y_i - \bar{y})^2}}$$

Características com correlação acima de um limiar (frequentemente  $|r| > 0,90$ ) são removidas, mantendo-se a mais estável ou mais relevante.

Referência clássica: Rodgers & Nicewander (1988).

### 3.6.3 Seleção Baseada em Importância de Árvores (*Random Forest Importance*)

Modelos baseados em árvores, como o RF, podem estimar a importância de cada variável durante o treinamento, permitindo reduzir dimensionalidade de forma supervisionada (Breiman, 2001; Vallières *et al.*, 2018). Entre as medidas mais utilizadas estão:

- Impureza de Gini:

$$\text{Imp}(f_k) = \sum_{t \in T_k} \Delta G_t$$

onde:

$T_k$  é o conjunto de nós em que a feature  $f_k$  é usada,

$\Delta G_t$  é a redução da impureza de Gini no nó  $t$ .

- Importância de Gini (Mean Decrease in Impurity — MDI):

$$G = 1 - \sum_{c=1}^C p_c^2$$

onde  $p_c$  é a proporção de amostras da classe  $c$  no nó.

- Importância de Permutação (*Mean Decrease in Accuracy* — MDA):

$$\text{Imp}(x_j) = \frac{1}{T} \sum_{t=1}^T \left( \text{Acc}_{\text{base}}^{(t)} - \text{Acc}_{\text{perm}(j)}^{(t)} \right)$$

onde:

$T$  é o número de árvores,

$\text{Acc}_i$  é a acurácia original da árvore  $i$ ,

$\text{Acc}_i^{(k)}$  é a acurácia após embaralhar a característica  $f_k$ .

A terceira é mais robusta, pois avalia a queda na performance quando uma feature é embaralhada.

Essa estratégia foi usada neste trabalho para gerar o ranking das características e em estudos radiômicos recentes para reduzir dimensionalidade (Vallières *et al.*, 2018).

### 3.6.4 Operador de Seleção e Menor Contração Absoluta (LASSO)

O LASSO é um dos métodos mais utilizados em radiômica clínica (Tibshirani, 1996), pois realiza simultaneamente regularização e seleção de características. Sua formulação é:

$$\min_{\beta} \left( \|y - X\beta\|_2^2 + \lambda \|\beta\|_1 \right)$$

O termo de penalização  $\|\beta\|_1$  força alguns coeficientes a se tornarem exatamente zero, eliminando automaticamente características irrelevantes.

## 3.7 Métricas de Avaliação

A avaliação de modelos de aprendizado supervisionado desempenha papel central na verificação da capacidade de generalização e na comparação entre abordagens distintas. Em problemas de classificação binária — como a identificação de bacilos em baciloscopias de campo claro — são amplamente utilizadas métricas derivadas da matriz de confusão, incluindo acurácia, precisão, sensibilidade, especificidade e *F1-score*. Além dessas, curvas ROC e a métrica AUC são empregadas para analisar o desempenho global do classificador

em diferentes limiares de decisão (Fawcett, 2006). A seguir, descrevem-se as principais métricas adotadas neste estudo.

### 3.7.1 Matriz de Confusão

A matriz de confusão organiza o desempenho do modelo segundo quatro categorias (Powers, 2011):

- Verdadeiros Positivos (TP) — predições corretas da classe positiva;
- Verdadeiros Negativos (TN) — predições corretas da classe negativa;
- Falso Positivos (FP) — erros onde a classe negativa foi classificada como positiva;
- Falso Negativos (FN) — erros onde a classe positiva foi classificada como negativa.

A matriz pode ser representada como:

$$\begin{bmatrix} TP & FP \\ FN & TN \end{bmatrix}$$

### 3.7.2 Acurácia

A acurácia expressa a proporção de classificações corretas sobre o total de amostras:

$$\text{Acurácia} = \frac{TP + TN}{TP + TN + FP + FN}$$

Apesar de amplamente utilizada, é considerada pouco informativa em bases desbalanceadas (Saito & Rehmsmeier, 2015).

### 3.7.3 Precisão

A precisão quantifica quantas das amostras classificadas como positivas são realmente positivas:

$$\text{Precisão} = \frac{TP}{TP + FP}$$

É especialmente relevante em contextos onde o custo de um falso positivo é elevado.

### 3.7.4 Sensibilidade

O sensibilidade mede a capacidade do modelo de identificar corretamente todas as amostras positivas:

$$\text{Recall} = \frac{TP}{TP + FN}$$

Em detecção de bacilos, esta métrica é crítica, pois falsos negativos implicam falha em identificar caso suspeito.

### 3.7.5 *F1-score*

O *F1-score* combina precisão e sensibilidade na forma de média harmônica (Chinchor, 1992):

$$F1 = 2 \cdot \frac{\text{Precisão} \cdot \text{Recall}}{\text{Precisão} + \text{Recall}}$$

É apropriado quando há necessidade de balancear as duas métricas.

### 3.7.6 Especificidade

Útil em contextos diagnósticos, a especificidade descreve a capacidade de identificar corretamente casos negativos:

$$\text{Especificidade} = \frac{TN}{TN + FP}$$

### 3.7.7 Curva ROC e AUC

A Receiver Operating Characteristic (ROC) representa a relação entre taxa de verdadeiros positivos (TPR) e taxa de falsos positivos (FPR):

$$TPR = \frac{TP}{TP + FN}, \quad FPR = \frac{FP}{FP + TN}$$

A área sob a curva, AUC, fornece uma medida global de desempenho independentemente do limiar:

$$AUC = \int_0^1 TPR(FPR) d(FPR)$$

Modelos com AUC próxima de 1 apresentam excelente separabilidade entre as classes (Hanley & McNeil, 1982).

# Capítulo 4

## Materiais e Métodos

Este capítulo apresenta, de forma detalhada, os materiais e procedimentos metodológicos adotados para o desenvolvimento dos experimentos realizados neste estudo. Assim, o capítulo descreve integralmente o conjunto de dados utilizado, as estratégias de construção das máscaras aplicadas às imagens, o processo de extração e seleção das características radiômicas, bem como as arquiteturas e configurações dos modelos supervisionados empregados, as métricas de avaliação e o ambiente computacional utilizado.

### 4.1 Materiais

A presente seção descreve o conjunto de dados e ambiente computacional utilizado para execução dos experimentos.

#### 4.1.1 Conjunto de Dados

O conjunto de dados utilizado neste estudo, denominado DDS1, foi disponibilizado pelo *Pattern Recognition and Optimization Research Group* da Universidade Federal do Amazonas (UFAM), conforme documentação associada ao banco, aprovado pelo Comitê de Ética em Pesquisa do Instituto Nacional de Pesquisas da Amazônia (INPA), protocolo 186/08. Este mesmo conjunto de dados foi empregado por Serrão et al. (2020) em sua proposta de detecção automática de bacilos utilizando redes neurais convolucionais (CNNs).

As imagens são oriundas de baciloscopia de campo claro do escarro corado pelo método de Kinyoun. As imagens originais foram adquiridas com uma câmera Canon PowerShot A640 (sensor CCD de 10 MP) acoplada a um microscópio Zeiss Axioskop 40 (aumento 100×, AN 1.25). Cada campo microscópico possui dimensão de 1388 × 1040 pixels, 24 bits de profundidade, gerado a partir de fusão multifocal (Extended Depth of Field – EDF) composta por 11 planos de foco.

O conjunto de dados DDS1 consiste em *patches* de 40×40 pixels, em formato BMP, extraídos das imagens EDF. Ao todo, contém 31.484 *patches*, sendo 28.848 negativos e 2.636 positivos.

Para cada *patch* positivo, o conjunto disponibiliza uma máscara binária manual como apresentados na Figura 8, na qual o bacilo aparece em branco (valor 255) e o fundo em preto (valor 0). Esses arquivos constituem o padrão ouro para imagens positivas.

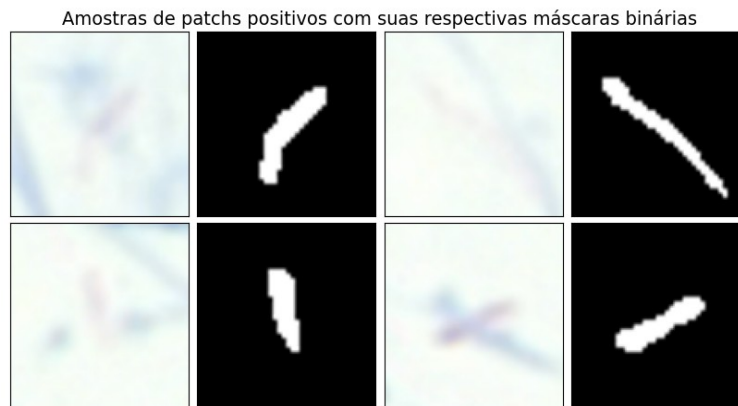


Figura 8: Amostra do dataset positivo e suas respectivas máscaras segmentadas com intensidade dos pixels já invertida.

Nos patches negativos (Figura 8), o padrão ouro correspondente é uma máscara inteiramente branca, indicando ausência de bacilos.

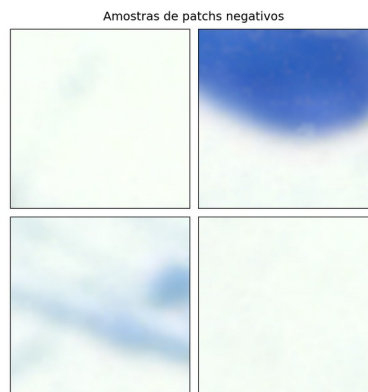


Figura 9: Amostra do dataset negativo.

A estrutura de diretórios separa imagens positivas, negativas e suas máscaras, permitindo manipulação independente dos arquivos e facilitando tanto *pipelines* radiômicos quanto de aprendizado profundo.

## 4.1.2 Ambiente Computacional

A extração radiômica e o treinamento dos modelos supervisionados foram realizados em um ambiente computacional baseado em um computador pessoal com o seguinte hardware: Processador: AMD Ryzen 7 7700, memória RAM: 32 GB, GPU: NVIDIA GeForce RTX 4070 Super, sistema operacional: Windows 11 (64 bits).

A implementação dos experimentos foi conduzida em Python 3.10.18 para extração de características com PyRadiomics, e Python 3.12.3 para treino e avaliação das arquiteturas.

Bibliotecas usadas: PyRadiomics 3.0.1, TensorFlow 2.19.0, scikit-learn 1.7.2, NumPy 2.1.3, OpenCV 4.11.0.86, Matplotlib 3.10.5 e Seaborn 0.13.2 para visualizações.

## 4.2 Métodos

Esta seção apresenta os métodos empregados no pré-processamento, geração da base de dados, construção de modelos e formas de avaliação. Na Figura 10 é ilustrado o fluxograma seguido durante a metodologia.

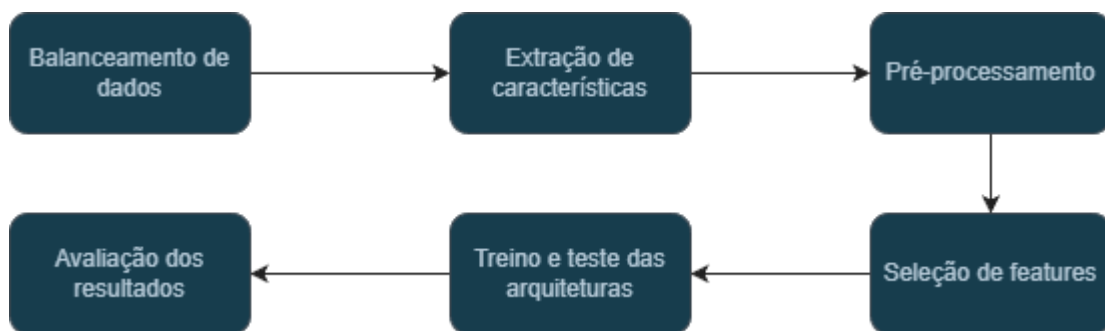


Figura 10: Fluxograma da metodologia

### 4.2.1 Balanceamento do Conjunto de Dados

Além do conjunto completo, foi extraído um subconjunto balanceado, composto por 2.635 patches positivos e 2.635 negativos, selecionados aleatoriamente. Esse subconjunto foi utilizado exclusivamente para os modelos supervisionados.

### 4.2.2 Estratégias de Construção das Máscaras

Para imagens negativas, onde não há bacilos segmentáveis, empregou-se uma máscara cheia  $40 \times 40$ , indicando que toda a imagem deveria ser analisada — abordagem comum em estudos radiômicos sem ROIs explícitas.

Além disso, um segundo conjunto radiômico foi produzido aplicando máscara cheia também às imagens positivas, permitindo avaliar os modelos considerando a textura global da imagem.

### 4.2.3 Extração Radiômica

A extração das características radiômicas foi realizada utilizando a biblioteca PyRadiomics, amplamente empregada em estudos de análise quantitativa de imagem médica graças à sua aderência às normas da *Image Biomarker Standardisation Initiative* (IBSI) (Zwanenburg *et al.*, 2020).

O extrator foi configurado de forma controlada, habilitando exclusivamente as classes de características utilizadas na literatura radiômica tradicional: *first-order statistics*, GLCM, GLRLM, GLSZM, GLDM e NGTDM. Todas as demais categorias — incluindo *shape features* e *diagnostics* — permaneceram desabilitadas, assegurando consistência entre os dois conjuntos radiômicos gerados. Ao todo, 116 atributos foram inicialmente produzidos para cada imagem do conjunto de dados balanceados.

A extração foi realizada duas vezes, correspondendo às duas estratégias de mascaramento descritas anteriormente. No primeiro conjunto (“máscara segmentada”), apenas a região correspondente ao bacilo foi considerada nas imagens positivas, enquanto nas negativas toda a imagem foi analisada. No segundo conjunto (“máscara cheia”), o PyRadiomics processou integralmente todas as imagens, independentemente do rótulo.

Os atributos extraídos são apresentados na Tabela 2, essas famílias de atributos descrevem desde o brilho global da região analisada até padrões texturais de alta ordem que frequentemente diferenciam objetos microscópicos e ruídos de fundo.

Tabela 2: Atributos extraídos

<b>Família de Atributos</b>	<b>Exemplos de atributos</b>
Medidas de intensidade global	<i>Entropy, Mean, Total Energy</i>
Textura baseada em coocorrência (GLCM)	<i>Idmn, Imc1</i>
Padrões de repetição de níveis de cinza (GLRLM)	<i>RunLengthNonUniformity, ShortRunEmphasis</i>
Distribuição espacial de zonas homogêneas (GLSZM)	<i>SmallAreaEmphasis, ZoneVariance</i>
Dependência entre pixels (GLDM)	<i>LargeDependenceHighGrayLevelEmphasis</i>
Métricas perceptuais de textura (NGTDM)	<i>Busyness, Coarseness, Contrast</i>

A radiômica foi processada de forma padronizada em todas as imagens do conjunto balanceado, resultando em dois subconjuntos finais com 5.270 instâncias cada — um para

máscaras segmentadas e outro para máscaras cheias. Esses subconjuntos serviram como entrada para a etapa subsequente de seleção e redução de características.

#### 4.2.4 Seleção e Tratamento de Características

A primeira etapa consistiu na remoção de colunas não numéricas, incluindo metadados gerados automaticamente pelo PyRadiomics, bem como atributos diagnósticos internos utilizados apenas para depuração. Esse procedimento assegurou que apenas variáveis estritamente quantitativas fossem levadas às fases posteriores. Para o conjunto segmentado restaram 98 atributos, enquanto para o conjunto com máscara cheia permaneceram 96 atributos, valores coerentes com a habilitação exclusiva das seis famílias principais de características.

Na sequência, foram eliminadas todas as variáveis com variância zero e posteriormente, aplicou-se uma filtragem baseada na correlação linear entre pares de atributos. De acordo com as diretrizes da literatura radiômica clínica e da IBSI, correlações superiores a 0,90 são consideradas indesejáveis, por refletirem redundância sem ganho informativo (Parmar *et al.*, 2015). Assim, de cada par altamente correlacionado, manteve-se apenas uma variável representativa. Após esse procedimento, o conjunto segmentado foi reduzido para 43 atributos, enquanto o conjunto com máscaras cheias resultou em 27 atributos.

Para ranquear os atributos remanescentes, empregou-se um modelo Random Forest treinado sobre cada dataset. A importância das variáveis foi calculada com base na redução média de impureza, estratégia amplamente utilizada em trabalhos de seleção de características pela sua estabilidade e robustez a relações não lineares (Breiman, 2001). Em seguida, foram selecionadas as 20 características mais importantes de cada conjunto, resultando em dois subconjuntos finais destinados ao treinamento do RF e do MLP.

Os dois conjuntos resultantes — um com 20 características derivado de máscaras segmentadas e outro com 20 características derivado de máscaras cheias — foram então utilizados como entrada para os modelos supervisionados avaliados nos experimentos.

#### 4.2.5 Modelos Supervisionados Utilizados

- Rede Neural Convolutacional (CNN)

A arquitetura convolucional utilizada neste trabalho foi construída especificamente para operar sobre os patches  $40 \times 40 \times 3$  do dataset balanceado normalizadas para o intervalo  $[0,1]$ . O modelo segue uma estrutura clássica de extração hierárquica de padrões, composta por duas camadas convolutivas seguidas de *pooling*, um bloco denso e camada de saída sigmoideal.

O primeiro bloco convolutivo emprega 32 filtros de  $3 \times 3$  com ativação ReLU, seguido por *max pooling*  $2 \times 2$  para reduzir a dimensionalidade espacial preservando as características relevantes. O segundo bloco aplica 64 filtros de  $3 \times 3$ , também seguido por *max pooling*. Após a etapa convolutiva, os mapas de ativação são achatados e conectados a uma camada densa de 128 neurônios com ativação ReLU, acompanhada de *dropout* de 0,5 para reduzir sobreajuste. A camada final, com um único neurônio e ativação sigmoideal, realiza a classificação binária. A Figura 11 apresenta a arquitetura proposta:

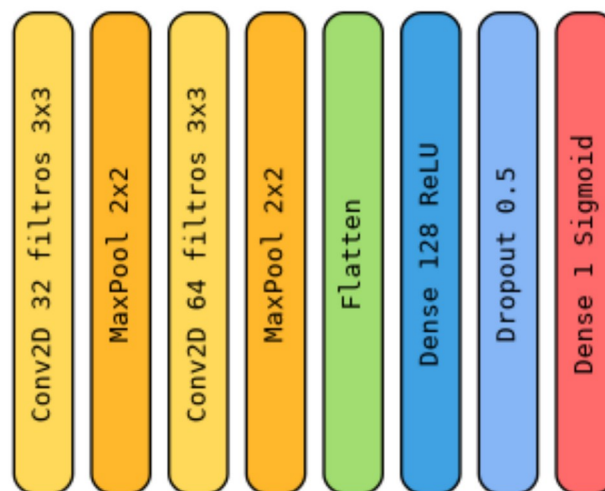


Figura 11: Arquitetura da rede CNN utilizada

O treinamento foi conduzido com o otimizador Adam, função de perda *binary cross-entropy*, *batch size* de 32 e 50 épocas. O conjunto foi dividido em treino, validação e teste nas proporções 70/15/15, com estratificação dos rótulos. A CNN obteve desempenho elevado, alcançando aproximadamente 0,99 de precisão, sensibilidade e *F1-score*, demonstrando forte capacidade discriminativa mesmo em imagens de pequeno porte.

- Floresta Aleatória (RF)

O modelo RF foi empregado tanto para a seleção de atributos quanto como classificador final dos conjuntos radiômicos. Para o treinamento, utilizou-se 100 árvores,

critério de impureza Gini e profundidade livre, permitindo ao modelo identificar relações não lineares entre as características extraídas pelo PyRadiomics.

O treinamento foi realizado sobre os dois datasets reduzidos a 20 atributos:

- conjunto derivado de máscaras segmentadas
- conjunto derivado de máscaras cheias

A divisão dos dados em 80% para treino e 20% para teste, com estratificação, assegurou equilíbrio entre as classes. No conjunto com máscaras segmentadas, o RF atingiu métricas perfeitas (1.00). No conjunto com máscaras cheias, alcançou acurácia de 0,89, com sensibilidade superior para a classe positiva. Esses resultados refletem a diferença estruturante entre os dois tipos de máscara.

- Perceptron Multicamadas (MLP)

O MLP foi construído com duas camadas densas intermediárias (64 e 32 neurônios, respectivamente), ambas com ativação ReLU, e uma camada de saída softmax para classificação binária após codificação one-hot. Antes do treino, as características foram normalizadas por meio de *StandardScaler*, uma prática recomendada para modelos baseados em gradiente descendente.

O otimizador Adam foi utilizado com perda *categorical cross-entropy*, *batch size* de 16 e 30 épocas de treinamento. Assim como o RF, o MLP foi treinado separadamente nos dois conjuntos radiômicos reduzidos para 20 atributos. No conjunto segmentado, o modelo atingiu desempenho perfeito (1.00) em todas as métricas. Já no conjunto com máscaras cheias, obteve acurácia de 0,88, semelhante ao RF.

## 4.2.6 Métricas de Avaliação

A avaliação dos modelos foi realizada utilizando as principais métricas derivadas da matriz de confusão — acurácia, precisão, sensibilidade, especificidade e *F1-score*.

A acurácia fornece uma visão geral do desempenho global do classificador. A precisão e a sensibilidade são particularmente importantes em aplicações médicas: a primeira reduz falsos positivos, enquanto a segunda minimiza falsos negativos, aspecto crítico na detecção de bacilos. A especificidade complementa a sensibilidade ao mensurar o quanto o modelo reconhece corretamente amostras negativas. O *F1-score*, por sua vez, resume o equilíbrio entre precisão e sensibilidade e é amplamente recomendado em cenários biomédicos (Sokolova & Lapalme, 2009; Chicco & Jurman, 2020).

## Capítulo 5

# Resultados e Discussões

Este capítulo apresenta os resultados obtidos pelos três grupos de modelos avaliados — CNN com conjunto de imagens e RF e MLP com subconjunto gerado de características radiômicos com máscaras segmentadas e com máscara cheia. Também é discutida a influência do vazamento estrutural observado no conjunto segmentado e suas implicações para a interpretação dos resultados.

### 5.1 Resultados da Rede Neural Convolutacional (CNN)

A CNN desenvolvida apresentou excelente desempenho na tarefa de classificação dos patches de baciloscopia, alcançando acurácia de 0,99 e perda de 0,03 no conjunto de teste como mostrado nas curvas de convergência (Figuras 12 e 13). As métricas por classe — precisão, sensibilidade e *F1-score* — permaneceram igualmente em 0,99, indicando que o modelo foi capaz de distinguir corretamente tanto patches contendo bacilos quanto patches negativos, com mínima ocorrência de erros.

A matriz de confusão (Figura 14) confirma este comportamento homogêneo entre as classes, demonstrando ausência de viés e estabilidade no aprendizado.

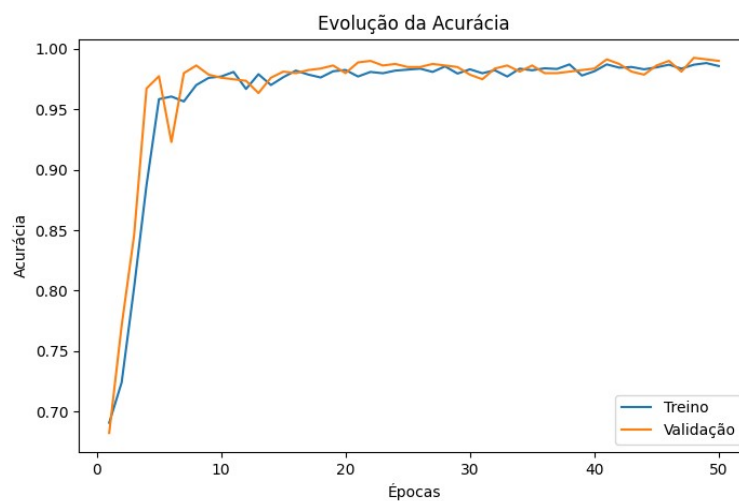


Figura 12: Curva de convergência de acurácia da rede CNN

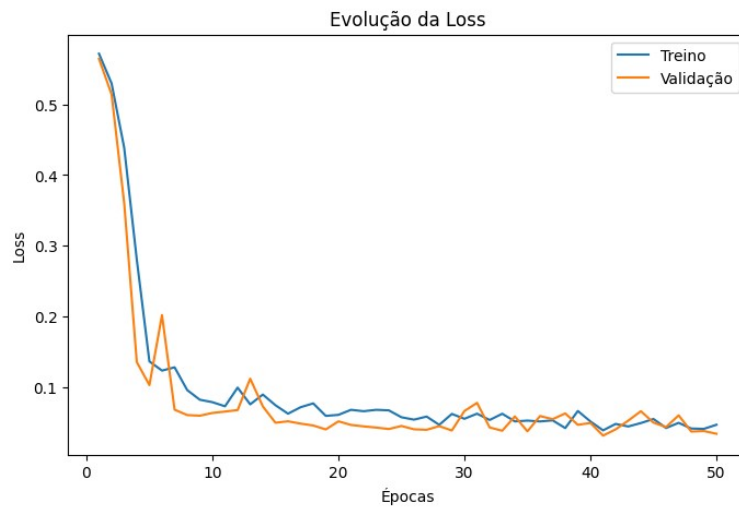


Figura 13: Curva de convergência de perda da rede CNN

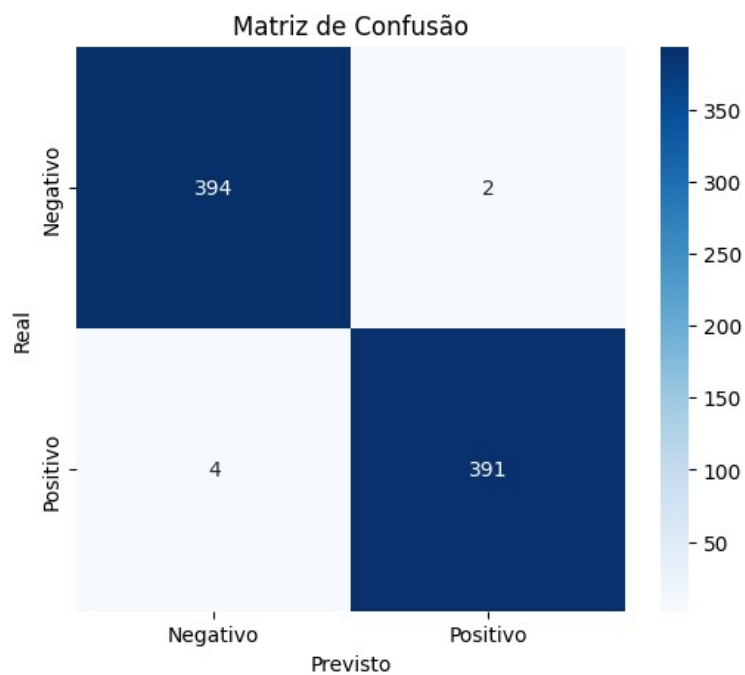


Figura 14: Matriz de confusão dos resultados com a rede CNN

Esse desempenho elevado é compatível com observações da literatura, que destaca a forte separabilidade visual entre bacilos e fundo microscópico — especialmente devido ao formato alongado, contraste e coloração característicos (Pérez-López *et al.*, 2017). Padrões morfológicos bem definidos favorecem o aprendizado hierárquico das CNNs, permitindo ao modelo extrair automaticamente texturas e bordas discriminativas.

## 5.2 Resultados dos Modelos Radiômicos com Máscaras Segmentadas

Nos experimentos com máscaras segmentadas, tanto o Random Forest quanto o MLP apresentaram desempenho perfeito nas métricas avaliadas (acurácia, precisão, sensibilidade e  $F1\text{-score} = 1.00$ ) sobre o conjunto de teste (527 positivos e 527 negativos).

Entretanto, esse resultado não reflete exclusivamente a capacidade discriminativa das características radiômicas. A análise das características selecionadas revela a presença da variável *diagnostics\_Mask-original\_VoxelNum*, que permaneceu entre as 20 mais importantes após o processo de seleção como mostrado na Figura 15 e na Tabela 3.

Tabela 3: Descrição das 10 melhores características selecionadas para dataset usando máscaras segmentadas

Nome da Feature (PyRadiomics)	Categoria / Matriz	Descrição Técnica
<i>diagnostics_Mask-original_VoxelNum</i>	Diagnóstico (máscara)	Número de pixels incluídos na máscara (tamanho da ROI).
<i>original_glrlm_GrayLevelNonUniformity</i>	GLRLM	Mede a variação dos níveis de cinza em todas as <i>runs</i> da imagem.
<i>original_glrlm_RunVariance</i>	GLRLM	Mede a variância do comprimento das <i>runs</i> (sequências de tons iguais).
<i>original_glrlm_LongRunHighGrayLevelEmphasis</i>	GLRLM	Realça <i>runs</i> longas com altos níveis de cinza.
<i>original_glszm_LargeAreaHighGrayLevelEmphasis</i>	GLSZM	Mede zonas grandes e de alta intensidade.
<i>original_glrlm_LongRunEmphasis</i>	GLRLM	Mede a predominância de <i>runs</i> longas, independentemente da intensidade.
<i>original_glszm_ZonePercentage</i>	GLSZM	Proporção de zonas em relação ao total de pixels.
<i>original_glrlm_ShortRunLowGrayLevelEmphasis</i>	GLRLM	Ênfase em <i>runs</i> curtas e escuras.
<i>original_glszm_ZoneVariance</i>	GLSZM	Variância do tamanho das zonas.
<i>original_glcm_Idmn</i>	GLCM	Mede a homogeneidade local da textura, com base na similaridade entre pixels vizinhos.

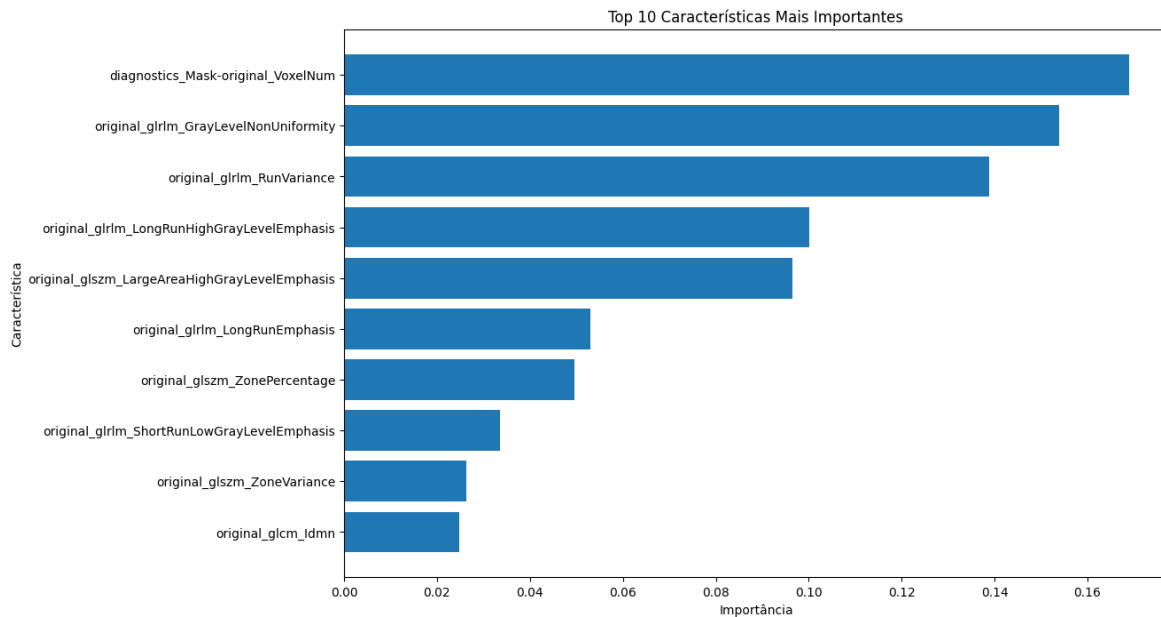


Figura 15: Rank das 10 melhores características por importância em % com dataset usando máscaras segmentadas.

Essa variável corresponde ao tamanho da região segmentada. Nas imagens positivas, a máscara cobre apenas o bacilo, variando de poucas dezenas a algumas centenas de pixels; nos patches negativos, a máscara é completamente cheia ( $40 \times 40 = 1600$  pixels). Assim, o tamanho da ROI atua como um marcador quase direto da classe, garantindo separação linear entre positivos e negativos.

Esse tipo de vazamento estrutural viola as recomendações da Image Biomarker Standardisation Initiative (IBSI), que orienta que máscaras devem ser comparáveis entre amostras para evitar inclusão de informações irrelevantes ao conteúdo da imagem (Zwanenburg *et al.*, 2020).

### 5.3 Resultados dos Modelos Radiômicos com Máscara Cheia

Quando as máscaras são uniformes e idênticas para todas as amostras ( $40 \times 40$  pixels), o vazamento estrutural é eliminado. Nesse cenário, os modelos passam a depender exclusivamente das informações texturais intrínsecas às imagens. O RF alcançou acurácia de 0.89, com precisão de 0.95 para negativos e 0.85 para positivos, sensibilidade de 0.83 e 0.96, e *F1-scores* de 0.89 e 0.90. O MLP exibiu valores semelhantes: acurácia de 0.88, precisão de 0.94/0.83, sensibilidade de 0.81/0.95 e *F1-scores* de 0.87/0.89 para negativos e positivos, respectivamente. As demais métricas indicam desempenho equilibrado entre as classes, com pequenas diferenças atribuídas à variabilidade textural mais evidente nos patches positivos.

Aqui, como todas as máscaras possuem exatamente o mesmo tamanho, não existe qualquer feature estrutural que revele a classe automaticamente como mostrado na Figura 16 e descritos na Tabela 4.

Essas características são compatíveis com estudos radiômicos que mostram importância de texturas de coocorrência, repetitividade e heterogeneidade para discriminar estruturas microscópicas (Parmar *et al.*, 2015; Vallières *et al.*, 2018).

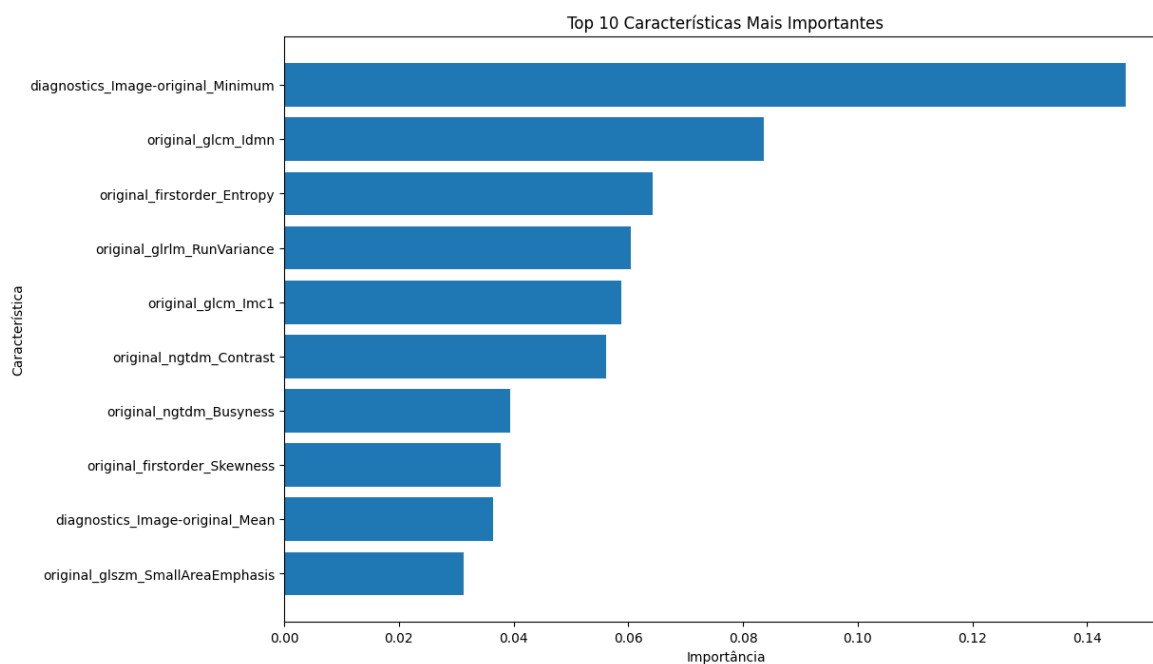


Figura 16: Rank das 10 melhores características por importância em % com dataset usando máscaras segmentadas.

Tabela 4: Descrição das 10 melhores características selecionadas para dataset usando máscaras cheias

Nome da Feature (PyRadiomics)	Categoria/ Matriz	Descrição Técnica
-------------------------------	----------------------	-------------------

diagnostics_Image-original_Minimum	Diagnóstico (imagem)	Menor intensidade presente na imagem delimitada pela máscara.
original_glcm_Idmn	GLCM	Mede a homogeneidade normalizada entre pixels vizinhos.
original_firstorder_Entropy	Primeira Ordem	Quantifica a desordem da distribuição de intensidades.
original_glrIm_RunVariance	GLRLM	Mede a variabilidade nos comprimentos das runs.
original_glcm_Imc1	GLCM	Índice de Informação Mútua 1; quantifica dependência estrutural da textura.
original_ngtdm_Contrast	NGTDM	Contraste local entre um pixel e a média dos vizinhos.
original_ngtdm_Busyness	NGTDM	Mede o quanto a intensidade muda rapidamente na vizinhança.
original_firstorder_Skewness	Primeira Ordem	Mede a assimetria da distribuição de intensidades.
diagnostics_Image-original_Mean	Diagnóstico (imagem)	Média da intensidade dos pixels dentro da máscara.
original_glszm_SmallAreaEmphasis	GLSZM	Realça zonas pequenas com intensidades semelhantes.

## 5.4 Discussão Comparativa entre as Abordagens

A Tabela 5 resume os resultados comparativos entre CNN, radiômica segmentada e radiômica com máscara cheia.

A CNN apresenta o melhor desempenho real, sem dependência de segmentação ou vazamento. Já o conjunto radiômico segmentado apresentou desempenho artificialmente perfeito devido ao tamanho desigual das máscaras, violando princípios fundamentais de consistência de ROI (Zwanenburg *et al.*, 2020).

Quando o vazamento é removido, como no conjunto de máscaras cheias, o desempenho cai para 0,88–0,89, representando a real capacidade da radiômica para essa aplicação.

Tabela 5: Métricas obtidas para os modelos avaliados

Abordagem	Acurácia	Precisão (Neg / Pos)	Sensibilidade (Neg / Pos)	F1-score (Neg / Pos)
CNN	0.99	0.99 / 0.99	0.99 / 0.99	0.99 / 0.99
RF – Segmentado	1.00	1.00 / 1.00	1.00 / 1.00	1.00 / 1.00

MLP – Segmentado	1.00	1.00 / 1.00	1.00 / 1.00	1.00 / 1.00
RF – Máscara cheia	0.89	0.95 / 0.85	0.83 / 0.96	0.89 / 0.90
MLP – Máscara cheia	0.88	0.94 / 0.83	0.81 / 0.95	0.87 / 0.89

---

# Capítulo 6

## Conclusão

Este trabalho avaliou três abordagens para a detecção de bacilos em esfregaços de baciloscopia de campo claro: radiômica com máscaras segmentadas, radiômica com máscaras cheias e uma rede neural convolucional treinada diretamente nas imagens. Os resultados mostraram diferenças claras entre as estratégias e revelaram limitações importantes relacionadas ao uso de máscaras no pipeline radiômico.

Em síntese, o estudo evidenciou que redes neurais convolucionais constituem uma solução robusta e prática para a detecção automática de bacilos, enquanto pipelines radiômicos requerem controle rigoroso das máscaras para evitar vieses e garantir validade experimental. Ainda assim, os resultados mostram que é plenamente viável utilizar características radiômicas extraídas de patches microscópicos nesse formato, desde que o processo de segmentação seja consistente entre classes. Pequenas inconsistências estruturais nas máscaras podem comprometer significativamente a análise, mas quando corretamente controladas, a radiômica se apresenta como uma abordagem complementar promissora para estudos de detecção em baciloscopia de campo claro.

Como trabalhos futuros, sugere-se investigar pipelines híbridos que combinem radiômica e deep learning, bem como avaliar os modelos em cenários clínicos reais para verificar sua aplicabilidade prática.

# Notas

Durante a elaboração deste trabalho, o autor utilizou a ferramenta Gemini 1.5 Flash para o refinamento gramatical de trechos da introdução, auxílio na padronização de citações bibliográficas e esclarecimento de conceitos técnicos relacionados ao processamento de imagens. Após a geração do conteúdo, o autor realizou a revisão crítica e edição necessária, assegurando a precisão técnica dos dados e assumindo total responsabilidade pelo texto final.

# Referências Bibliográficas

- [1] Aerts, H. J. W. L., *et al.* Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nature Communications*, v. 5, 4006, 2014. DOI: <https://doi.org/10.1038/ncomms5006>
- [2] Balcells, M. E., *et al.* Accuracy of light microscopy for diagnosing tuberculosis. *Clinical Microbiology and Infection*, v. 25, n. 11, p. 1404–1409, 2019.
- [3] Breiman, L. Random forests. *Machine Learning*, v. 45, p. 5–32, 2001.
- [4] Chakravorty, S., *et al.* The role of smear microscopy in tuberculosis diagnosis. *European Respiratory Journal*, 2017.
- [5] Chicco, D.; Jurman, G. The advantages of the MCC over F1 score and accuracy. *BMC Genomics*, v. 21, 2020.
- [6] Chinchor, N. MUC-4 Evaluation Metrics. In: *Message Understanding Conference – MUC-4*. 1992.
- [7] Costa, M. G. F.; Costa Filho, C. F. F. Automated tuberculosis bacillus detection using SVM. In: *XX SIBGRAPI*. IEEE, 2008.
- [8] Fawcett, T. An introduction to ROC analysis. *Pattern Recognition Letters*, v. 27, p. 861–874, 2006.
- [9] Geurts, P.; Ernst, D.; Wehenkel, L. Extremely randomized trees. *Machine Learning*, v. 63, p. 3–42, 2006.
- [10] Goodfellow, I.; Bengio, Y.; Courville, A. *Deep Learning*. MIT Press, 2016.
- [11] Guyon, I.; Elisseeff, A. An introduction to variable and feature selection. *JMLR*, v. 3, p. 1157–1182, 2003.
- [12] Hanley, J.; McNeil, B. The meaning and use of the AUC. *Radiology*, v. 143, p. 29–36, 1982.
- [13] Ioffe, S.; Szegedy, C. Batch Normalization. *ICML*, 2015.
- [14] Kingma, D.; Ba, J. Adam: A method for stochastic optimization. *ICLR*, 2015.

- [15] Kononenko, I. Machine learning for medical diagnosis. *Artificial Intelligence in Medicine*, v. 23, p. 89–109, 2001.
- [16] Krizhevsky, A.; Sutskever, I.; Hinton, G. ImageNet classification with deep CNNs. *NIPS*, 2012.
- [17] Kursa, M.; Rudnicki, W. Feature selection with Boruta. *J. Stat. Software*, v. 36, 2010.
- [18] Lambin, P., *et al.* Radiomics: extracting more information... *Eur. J. Cancer*, v. 48, p. 441–446, 2012.
- [19] Lambin, P., *et al.* Radiomics: the bridge to personalized medicine. *Nat. Rev. Clin. Oncol.*, v. 14, p. 749–762, 2017.
- [20] LeCun, Y., *et al.* Backpropagation applied to handwritten zip codes. *Neural Computation*, v. 1, p. 541–551, 1989.
- [21] LeCun, Y., *et al.* Gradient-based learning applied to document recognition. *Proc. IEEE*, 1998.
- [22] Nahid, P., *et al.* Update on tuberculosis diagnostics. *AJRCCM*, v. 205, p. 1255–1273, 2022.
- [23] Nair, V.; Hinton, G. ReLUs improve RBMs. *ICML*, 2010.
- [24] Parekh, V.; Jacobs, M. Radiomics: methodology and clinical application. *Journal of Nuclear Medicine*, v. 60, 2019.
- [25] Parmar, C., *et al.* Robust radiomics feature quantification. *Radiology*, v. 275, p. 770–782, 2015.
- [26] Pérez López, Y.; Costa Filho, C. F. F.; Aguilera, L.; Costa, M. Automatic classification of light field smear microscopy patches using CNNs. *CHILECON*, 2017.
- [27] Powers, D. M. W. Evaluation: from precision, recall and F to ROC, MCC... *JMLT*, v. 2, 2011.
- [28] Rodgers, J.; Nicewander, W. Thirteen ways to look at the correlation coefficient. *The American Statistician*, v. 42, 1988.
- [29] Saito, T.; Rehmsmeier, M. Precision-recall vs ROC. *PLOS ONE*, 2015.

- [30] Scherer, D.; Müller, A.; Behnke, S. Evaluation of pooling operations. ICANN, 2010.
- [31] Serrão, E., *et al.* Machine learning in tuberculosis diagnosis: a survey. *Artificial Intelligence in Medicine*, 2020.
- [32] Shamout, F., *et al.* A review of deep learning in medical applications. *npj Digital Medicine*, v. 4, 2021.
- [33] Singh, D.; Kumar, M. Tuberculosis bacilli detection using Zernike moments and MLP. *Journal of Medical Systems*, 2020.
- [34] Sokolova, M.; Lapalme, G. Performance measures for classification. *Information Processing & Management*, 2009.
- [35] Srivastava, N., *et al.* Dropout: a simple way to prevent overfitting. *JMLR*, 2014.
- [36] Steingart, K., *et al.* Sputum smear microscopy for TB diagnosis. *The Lancet Infectious Diseases*, v. 6, 2006.
- [37] Tibshirani, R. Regression shrinkage and selection via LASSO. *JRSS B*, v. 58, 1996.
- [38] van Griethuysen, J., *et al.* Computational radiomics system. *Cancer Research*, v. 77, 2017.
- [39] Vallières, M., *et al.* Radiomics strategies for cancer risk assessment. *Scientific Reports*, v. 8, 2018.
- [40] WHO. Global tuberculosis report 2023. World Health Organization, 2023.
- [41] Zwanenburg, A., *et al.* The Image Biomarker Standardisation Initiative. *Radiology*, v. 295, 2020.
- [42] Grupo de Pesquisa em Reconhecimento de Padrões e Otimização — UFAM. Disponível em: <http://dgp.cnpq.br/dgp/espelhogrupo/4186>. Acesso em: 21 out. 2025.
- [43] Predicting Sugarcane Yield via the Use of an Improved Least Squares Support Vector Machine and Water Cycle Optimization Model - Scientific Figure on ResearchGate.
- [44] Nvidia. (2024). Convolutional Neural Network (CNN). Glossary – NVIDIA. Disponível em: <https://www.nvidia.com/en-in/glossary/convolutional-neural-network/>. [acessado em 27 Nov 2025].

[45] py-Radiomics. Open-source radiomics library written in python. Disponível em: <https://radiomics.github.io/pyradiomics.html>. [acessado em 27 Nov 2025].

[46] Automated Prediction of Good Dictionary EXamples (GDEX): A Comprehensive Experiment with Distant Supervision, Machine Learning, and Word Embedding-Based Deep Learning Techniques - Scientific Figure on ResearchGate. Available from: [https://www.researchgate.net/figure/Illustration-of-random-forest-trees\\_fig4\\_354354484](https://www.researchgate.net/figure/Illustration-of-random-forest-trees_fig4_354354484). [acessado em 27 Nov 2025].