



DAVI HORNER HOE DE CASTRO

**MODELOS DE DETECÇÃO DE TUBERCULOSE ATIVA VIA
IMAGENS DE RAIO X UTILIZANDO CURVAS PRINCIPAIS**

LAVRAS – MG

2025

DAVI HORNER HOE DE CASTRO

**MODELOS DE DETECÇÃO DE TUBERCULOSE ATIVA VIA IMAGENS DE RAO X
UTILIZANDO CURVAS PRINCIPAIS**

Dissertação apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação de Mestrado em Engenharia de Sistemas e Automação, para obtenção do título de Mestre.

Prof. DSc. Danton Diego Ferreira
Orientador

**LAVRAS – MG
2025**

**Ficha catalográfica elaborada pela Coordenadoria de Processos Técnicos da Biblioteca
Universitária da UFLA, com dados informados pelo(a) próprio(a) autor(a).**

Castro, Davi Horner Hoe de

Modelos de Detecção de Tuberculose ativa via imagens de raio X utilizando curvas principais / Davi Horner Hoe de Castro.
– Lavras : UFLA, 2025.

73p. : il.

Dissertação(mestrado)–Universidade Federal de Lavras,
2025.

Orientador: Prof. DSc. Danton Diego Ferreira.

Bibliografia.

1. Tuberculose. 2. Inteligência Artificial . 3. Curva Principal. 4. K-Segmentos. 5. Análise de Imagens. 6. Ética. 7. IA Sustentável. I. Ferreira, Danton Diego. II. Universidade Federal de Lavras. III. Título

CDD 808.066

DAVI HORNER HOE DE CASTRO

**MODELOS DE DETECÇÃO DE TUBERCULOSE ATIVA VIA IMAGENS DE RAO X
UTILIZANDO CURVAS PRINCIPAIS**

Dissertação apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação de Mestrado em Engenharia de Sistemas e Automação, para obtenção do título de Mestre.

APROVADA em 29 de maio de 2025.

Prof. DSc. Danton Diego Ferreira	UFLA
Prof. DSc. Bruno Henrique Groenner Barbosa	UFLA
Profa. DSc. Wilian Soares Lacerda	UFLA
Prof. DSc. Adriano Olimpo Tonelli	IFMG

Prof. DSc. Danton Diego Ferreira
Orientador

**LAVRAS – MG
2025**

AGRADECIMENTOS

Eu gostaria de agradecer ao meu orientador Danton, a UFLA, ao grupo AIA, a todos os meus professores do Mestrado e aos órgãos de fomento (CAPES, CNPq e FAPEMIG) por me apoiar e proporcionar essa oportunidade para aprimorar meus estudos.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001

RESUMO

A Tuberculose (TB) permanece como uma das doenças infecciosas mais letais do mundo, sobretudo em regiões de baixa e média renda, onde o acesso a sistemas de saúde eficientes ainda é limitado. Quando não diagnosticada precocemente, a doença pode evoluir para quadros graves, levando ao óbito ou gerando sequelas significativas no sistema respiratório, o que reforça a necessidade de aprimoramento contínuo das ferramentas de triagem e diagnóstico, especialmente diante do aumento no volume de exames e da demanda por métodos rápidos, robustos e acessíveis. Nesta dissertação é proposto e avaliado um método de inteligência computacional para detecção de tuberculose por meio da análise de imagens de Raio X do tórax, cujo diferencial é seu caráter parcimonioso, privilegiando simplicidade, reduzido número de parâmetros e baixa complexidade computacional, alinhando-se às diretrizes contemporâneas da IA Sustentável (*Green AI*), que ressaltam a importância de sistemas energeticamente eficientes e com baixa pegada do carbono. O método integra a técnica de Curvas Principais (CP), utilizada como mecanismo especialista para extrair estruturas geométricas relevantes das imagens. O método principal usa a Análise de Componentes Principais (PCA), responsável pela redução de dimensionalidade e preservação da variabilidade do conjunto de dados (CPE+PCA), mas foi testado um método utilizando SVD para redução de dimensões (CPE+SVD), um onde só se utilizou as curvas principais (CPE-Solo), e por fim um onde se utilizou uma Resnet-50 para extração de características em conjunto das curvas principais e Análise de Componentes Principais (CPE+Resnet+PCA). Para a avaliação, empregou-se um banco de dados amplamente utilizado na literatura, o TBX11K, composto por três classes: pacientes com tuberculose (800 imagens), com outras doenças pulmonares (3.800 imagens) e indivíduos saudáveis (3.800 imagens), permitindo simular um cenário realista de triagem clínica. Os resultados indicaram acurácia de 0,87 e uma precisão para TB de 0.69, tempo médio de processamento de 0,020 segundos por imagem e pegada de carbono de 0.010 kg de CO₂-eq e um gasto de energia de 0.065 kWh, valores competitivos frente a métodos mais complexos baseados em aprendizado profundo, sobretudo considerando o baixo custo computacional e a elevada eficiência energética. Esses achados demonstram que a abordagem proposta oferece uma alternativa promissora para aplicações em contextos com recursos limitados e reforçam a viabilidade de desenvolver soluções de diagnóstico assistido que sejam simultaneamente precisas, eficientes e ambientalmente responsáveis, contribuindo para o avanço das aplicações de inteligência artificial na saúde.

Palavras-chave: Tuberculose; Inteligência Artificial; Curva Principal; K-Segmentos, Análise de Imagens, Ética, IA Sustentável

ABSTRACT

Tuberculosis (TB) remains one of the most lethal infectious diseases worldwide, particularly in low- and middle-income regions where access to efficient healthcare systems is still limited. When not diagnosed early, the disease can progress to severe stages, leading to death or causing significant long-term respiratory sequelae, which reinforces the need for continuous improvement of screening and diagnostic tools, especially given the increasing volume of examinations and the demand for methods that are fast, robust, and accessible. This dissertation proposes and evaluates a computational intelligence method for tuberculosis detection using chest X-ray image analysis, whose main distinguishing feature is its parsimonious nature, emphasizing simplicity, a reduced number of parameters, and low computational complexity, aligning with contemporary Green AI principles that highlight the importance of energy-efficient systems with minimal environmental impact. The principal method integrates the Principal Curves (PC) technique, employed as a specialist mechanism to extract relevant geometric structures from images, with Principal Component Analysis (PCA), responsible for dimensionality reduction while preserving essential variability in the dataset (CPE+PCA), but was also tested a method using SVD for reduction of dimensions (CPE+SVD), one where it was only used principal curves (CPE-Solo), and finally one where was utilized a Resnet-50 for feature extraction together with principal curves and Principal Component Analysis (CPE+Resnet+PCA). For evaluation, a widely referenced database was used, the TBX11K, containing three classes: tuberculosis patients (800 images), patients with other pulmonary diseases (3,800 images), and healthy individuals (3,800 images), enabling the simulation of a realistic clinical screening scenario. The results indicated an accuracy of 0.87, a precision of 0.69 for TB, an average processing time of 0.020 seconds per image, and a carbon footprint of 0.010 kg of CO₂-eq and an energy expenditure of 0.065 kWh, values that are competitive when compared to more complex deep learning-based methods, particularly given the model's low computational cost and high energy efficiency. These findings demonstrate that the proposed approach offers a promising alternative for applications in resource-constrained settings and reinforce the feasibility of developing assisted diagnostic solutions that are simultaneously accurate, efficient, and environmentally responsible, contributing to the advancement of artificial intelligence applications in healthcare.

Keywords: Tuberculosis; Artificial intelligence; Main Curve; K-Segments, Image Analysis, Ethics, Green AI

INDICADORES DE IMPACTO

A pesquisa desenvolvida apresenta impactos diretos e potenciais alinhados ao Objetivo de Desenvolvimento Sustentável (ODS 3 — Saúde e Bem-Estar), estabelecido pela Organização das Nações Unidas (ONU), ao propor uma abordagem computacional voltada à detecção automática da tuberculose a partir de imagens médicas, com maior eficiência no uso de recursos computacionais. No âmbito da Política Nacional de Extensão, os impactos decorrentes do estudo enquadram-se predominantemente nas áreas temáticas de Saúde e Meio Ambiente, considerando a redução da demanda computacional e do consumo energético associada aos métodos propostos.

Sob a perspectiva tecnológica e social, a pesquisa resulta no desenvolvimento de uma abordagem que reduz de forma significativa os custos computacionais usualmente associados a modelos de aprendizado profundo. Os resultados preliminares demonstram tempo de treinamento aproximadamente quatro vezes menor em relação a abordagens baseadas em redes neurais profundas, bem como tempo de inferência por imagem até cerca de dez vezes inferior, caracterizando impacto tecnológico direto. Esses ganhos ampliam a viabilidade de aplicação da solução em ambientes hospitalares com infraestrutura computacional limitada.

Os benefícios decorrentes da maior eficiência computacional apresentam especial relevância para unidades de saúde localizadas em territórios socialmente vulneráveis, nos quais há restrições de recursos financeiros e tecnológicos. Nesses contextos, a pesquisa possui impacto social potencial sobre grupos populacionais com maior suscetibilidade à tuberculose, ao contribuir para o fortalecimento de estratégias de diagnóstico precoce e para o suporte à tomada de decisão clínica.

Adicionalmente, a menor exigência de capacidade computacional e o consequente menor consumo energético configuram impactos ambientais indiretos, alinhados aos princípios de sustentabilidade e uso eficiente de recursos, em consonância com diretrizes contemporâneas de tecnologias verdes aplicadas à área da saúde. Dessa forma, os resultados obtidos evidenciam o potencial da pesquisa para gerar benefícios tecnológicos, sociais e ambientais, contribuindo para práticas mais sustentáveis no contexto da saúde pública.

IMPACT INDICATORS

The research developed presents direct and potential impacts aligned with the Sustainable Development Goal (SDG 3 — Good Health and Well-Being), established by the United Nations (UN), by proposing a computational approach aimed at the automatic detection of tuberculosis from medical images, with greater efficiency in the use of computational resources. Within the scope of the National Extension Policy, the impacts arising from the study are predominantly classified within the thematic areas of Health and Environment, considering the reduction in computational demand and energy consumption associated with the proposed methods.

From a technological and social perspective, the research results in the development of an approach that significantly reduces the computational costs typically associated with deep learning models. Preliminary results demonstrate a training time approximately four times shorter compared to approaches based on deep neural networks, as well as an image inference time up to approximately ten times lower, characterizing a direct technological impact. These gains increase the feasibility of applying the solution in hospital environments with limited computational infrastructure.

The benefits derived from increased computational efficiency are particularly relevant for healthcare units located in socially vulnerable territories, where financial and technological resources are constrained. In these contexts, the research exhibits potential social impact on population groups with greater susceptibility to tuberculosis, by contributing to the strengthening of early diagnostic strategies and supporting clinical decision-making processes.

Additionally, the reduced requirement for computational capacity and the resulting lower energy consumption constitute indirect environmental impacts, aligned with the principles of sustainability and efficient resource use, in accordance with contemporary guidelines on green technologies applied to healthcare. Thus, the results obtained highlight the potential of the research to generate technological, social, and environmental benefits, contributing to more sustainable practices within the context of public health.

LISTA DE FIGURAS

Figura 1.1 – Mortes causadas por Tuberculose entre 2010 e 2022.	17
Figura 1.2 – Exemplos de radiografias, imagens 1 e 2 - Pulmão com Tuberculose, 3 - Pulmão com outra doença, 4 - Pulmão Saudável.	20
Figura 2.1 – Representação geométrica da Análise de Componentes Principais (PCA) . .	23
Figura 2.2 – PCA obtido para a base de dados.	24
Figura 2.3 – Criação de uma Curva Principal	27
Figura 2.4 – Exemplo de Curva Principal construída com 7 segmentos (K=7).	27
Figura 2.5 – Fluxograma do algoritmo K-seg para obtenção da Curva Principal.	28
Figura 2.6 – Exemplo do uso de Curvas Principais como classificador.	30
Figura 2.7 – Exemplo de uma VGG16.	31
Figura 2.8 – Representação ilustrativa da arquitetura ResNet-34.	32
Figura 2.9 – Representação da Arquitetura da Densenet	34
Figura 2.10 – Representação da Arquitetura da ViT16	35
Figura 3.1 – Diagrama dos Métodos propostos	45
Figura 4.1 – Matriz de confusão referente a validação do melhor fold para o classificador VGG16	52
Figura 4.2 – Matriz de confusão referente a validação do melhor fold para o classificador Resnet-50	53
Figura 4.3 – Matriz de confusão referente a validação do melhor fold para o classificador Densenet-121	53
Figura 4.4 – Matriz de confusão referente a validação do melhor fold para o classificador ViT16	54
Figura 4.5 – Matriz de confusão referente a validação do melhor fold para o classificador CPE+PCA	54
Figura 4.6 – Matriz de confusão referente a validação do melhor fold para o classificador CPE+SVD	55
Figura 4.7 – Matriz de confusão referente a validação do melhor fold para o classificador CPE-Solo	55

Figura 4.8 – Matriz de confusão referente a validação do melhor fold para o classificador

 CPE+Resnet+PCA 56

Figura 4.9 – Distribuição das acurácias dos métodos de classificação. 62

Figura 4.10 – Distribuição das precisões para TB dos métodos de classificação. 63

LISTA DE TABELAS

Tabela 2.1 – Resumo das principais arquiteturas analisadas.	37
Tabela 3.1 – Parâmetros utilizados na Abordagem I (CPE+PCA).	46
Tabela 3.2 – Parâmetros utilizados na Abordagem II (CPE+SVD).	47
Tabela 3.3 – Parâmetros utilizados na Abordagem III (CPE-Solo).	47
Tabela 3.4 – Parâmetros utilizados na Abordagem IV (CPE+Resnet+PCA).	48
Tabela 3.5 – Hiperparâmetros utilizados nos experimentos.	50
Tabela 4.1 – Tempo total de treinamento dos métodos comparados	57
Tabela 4.2 – Tempo médio de análise por imagem (média $\pm$ desvio padrão), com 4 casas decimais.	57
Tabela 4.3 – Tabela do gasto de energia (kwh)	58
Tabela 4.4 – Emissões de COO ₂ e por modelo avaliado	59
Tabela 4.5 – Resultados da Acurácia (média $\pm$ desvio padrão) dos métodos	60
Tabela 4.6 – Resultados da Precisão TB(média $\pm$ desvio padrão) dos métodos	61
Tabela 4.7 – Resultados da Análise Estatística: Teste de Friedman e Pós-hoc Nemenyi para Acurácia e Precisão TB.	61
Tabela 4.8 – Matriz de p-valores do pós-teste Nemenyi para Acurácia.	61
Tabela 4.9 – Matriz de p-valores do pós-teste Nemenyi para Precisão TB.	62
Tabela 4.10 – Comparação das abordagens propostas com outros métodos	64
Tabela 4.11 – Comparação das abordagens propostas com outros métodos da literatura	66

LISTA DE SIGLAS

CP	Curvas Principais
CPE	Curvas Principais Especialistas
PCA	Análise de Componentes Principais
SVD	Decomposição em Valores Singulares
TB	Tuberculose

SUMÁRIO

1	INTRODUÇÃO	16
1.1	Objetivos	18
1.2	Organização do Trabalho	19
1.3	Desafios e Motivação	19
2	REVISÃO BIBLIOGRÁFICA	21
2.1	Tuberculose	21
2.2	Análise de Componentes Principais (PCA)	23
2.3	Decomposição em Valores Singulares (SVD)	25
2.4	Cálculo de Curvas Principais pelo Método K-seg	25
2.4.1	Introdução	25
2.4.2	Definição formal	26
2.4.3	Motivação para a aproximação por segmentos	26
2.4.4	Descrição do método K-seg	27
2.4.5	Observações Importantes	29
2.4.6	Curvas Principais aplicadas à classificação	29
2.4.7	Limitações	30
2.5	Revisão das arquiteturas comparativas utilizadas	30
2.5.1	VGG16	31
2.5.1.1	Características Arquiteturais	31
2.5.1.2	Pontos Fortes e Limitações	32
2.5.2	Resnet50	32
2.5.2.1	Conexões Residuais	33
2.5.2.2	Arquitetura em Blocos Tipo <i>Bottleneck</i>	33
2.5.2.3	Pontos Fortes e Limitações	33
2.5.3	Densenet-121	34
2.5.3.1	Conectividade Densa	34
2.5.3.2	Estrutura Arquitetural	34
2.5.3.3	Pontos Fortes e Limitações	35
2.5.4	Vision Transformer (ViT)	35
2.5.4.1	Formação dos <i>Patches</i>	36
2.5.4.2	Codificador Transformer	36

2.5.4.3	Pontos Fortes e Limitações	36
2.5.5	Considerações Comparativas	37
2.6	IA de Confiança e IA Sustentável	37
2.6.1	Contextualização	37
2.6.2	IA de Confiança	38
2.6.3	IA Sustentável	39
3	Metodologia	41
3.1	Base de dados	41
3.2	Contextualização Metodológica	42
3.3	Protocolo Experimental	42
3.4	Configurações do equipamento utilizado	43
3.5	Pré Processamento	43
3.5.1	Procedimentos Comuns	44
3.5.2	Pré-processamento Específico dos Métodos Propostos	44
3.5.3	Pré-processamento nas Arquiteturas Comparativas	45
3.6	Método CPE+PCA: Curvas Principais Especialistas com PCA	46
3.7	Método CPE+SVD: Curvas Principais Especialistas com SVD	46
3.8	Método CPE-Solo: Curvas Principais Especialistas	47
3.9	Método CPE+Resnet+PCA: Curvas Principais Especialistas com extração de características usando Resnet-50 com PCA	47
3.10	Métodos comparativos	48
3.10.1	VGG16	48
3.10.2	Resnet-50	48
3.10.3	Densenet-121	49
3.10.4	ViT	49
3.10.5	Hiperparâmetros dos modelos comparativos	49
3.11	Seleção do Melhor Modelo	49
3.12	Avaliação Final	50
3.13	Monitoramento de Emissões e Custo Computacional	50
4	Resultados e Discussão	52
4.1	Matriz de confusão	52
4.2	Tempo de processamento e custo computacional	56

4.2.1	Tempo de inferência	56
4.2.2	Complexidade computacional	58
4.3	Consumo energético estimado e Pegada de carbono (COO ₂ eq)	58
4.4	Análise do Teste Estatístico	59
4.5	Discussão dos resultados	61
5	CONCLUSÃO	67
	REFERÊNCIAS	69

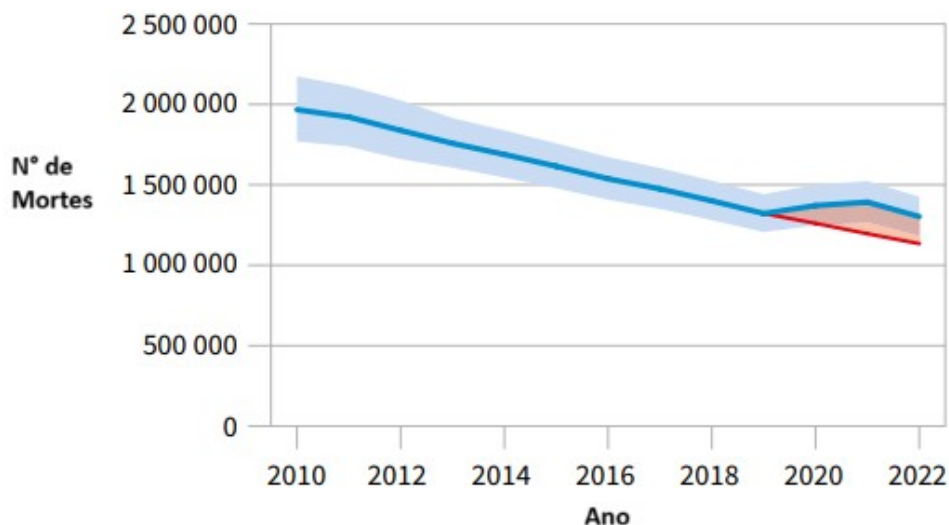
1 INTRODUÇÃO

A Tuberculose (TB) matou mais de um bilhão de pessoas nos últimos dois séculos e permanece como a doença infecciosa mais letal, provocada pela bactéria *Mycobacterium tuberculosis* (Chapagain; Yaqoob, 2019). Apesar do progresso contínuo em métodos diagnósticos, políticas de controle e desenvolvimento de terapias, a doença ainda causa milhões de infecções e um número significativo de óbitos anualmente. Mesmo após a cura, muitos pacientes continuam a apresentar sequelas significativas, como a redução permanente da capacidade respiratória; essas sequelas afetam mais da metade das pessoas diagnosticadas com a forma pulmonar da doença, sua manifestação mais comum, embora a infecção possa atingir outros órgãos, como rins, intestinos e olhos (Zaman, 2010). Assim, o diagnóstico precoce da tuberculose é essencial tanto para o tratamento individual quanto para o controle epidemiológico (Brasil. Ministério da Saúde. Secretaria de Vigilância em Saúde. Departamento de Vigilância das Doenças Transmissíveis, 2019).

Segundo o relatório global da Organização Mundial da Saúde (OMS) de 2023 (OMS, 2024), estima-se que 10,8 milhões de pessoas tenham contraído TB em 2023, embora apenas 8,2 milhões tenham sido oficialmente diagnosticadas. A Figura 1.1 apresenta a evolução das mortes causadas pela doença ao longo dos últimos anos. A linha azul representa as mortes estimadas, com a área sombreada indicando um intervalo de confiança de 95%; já a linha vermelha apresenta uma estimativa contrafactual na qual a pandemia não teria ocorrido. Os dados mostram que, mesmo em 2023, os esforços de controle ainda não retornaram ao nível pré-pandemia, devido, sobretudo, à redução do investimento e à sobrecarga dos sistemas de saúde (Pai Tereza Kasaeva, 2022). O relatório destaca que, em 2022, houve um número recorde de 7,5 milhões de diagnósticos, que foi superado pelos 8,2 milhões em 2023, mostrando quão perturbadora a pandemia de COVID foi para os serviços de saúde relacionados à TB e como ainda afeta pessoas em todo o mundo. Outro destaque é que, em 2023, as mortes por TB ultrapassaram as mortes por COVID-19 e por VIH/SIDA, refletindo a persistente gravidade da doença (Trajman *et al.*, 2025).

Além dos impactos clínicos, mudanças recentes no cenário global têm influenciado diretamente o desenvolvimento de tecnologias de diagnóstico. Em particular, a pandemia da COVID-19 acelerou globalmente a adoção de tecnologias digitais, ampliando o uso de automação e inteligência artificial em diversos setores, especialmente na saúde (Peterson Terence To-

Figura 1.1 – Mortes causadas por Tuberculose entre 2010 e 2022.



Fonte: Adaptado de Organização Mundial de Saúde (OMS, 2024).

land, 2020). Entre os exemplos anteriores à pandemia, destaca-se o sistema Watson, da IBM, que integra ferramentas de reconhecimento de padrões, automação de processos e apoio à decisão clínica. A vertente Watson Health (Watson, 2022) foi desenvolvida especificamente para o domínio médico, com o propósito de analisar imagens, prever resultados, identificar efeitos colaterais e apoiar diagnósticos.

A pandemia impulsionou investimentos em automação, particularmente em IA para imagens médicas (Peterson Terence Toland, 2020). A *HealthTech Magazine* (Steger, 2020) destaca iniciativas como o algoritmo Cimatec-XCOV19 (Furtado *et al.*, 2022), desenvolvido para apoiar o diagnóstico de COVID-19 em cenários sem a disponibilidade de testes moleculares. Avanços em arquiteturas profundas, incluindo CNNs e Transformers, têm ampliado a capacidade de extração de características e alcançado desempenhos notáveis em classificação e segmentação (Cantrell, 2018; Alzubaidi *et al.*, 2021). Contudo, esses modelos são computacionalmente caros, demandam hardware especializado e geram um consumo energético elevado, dificultando sua adoção em ambientes clínicos com infraestrutura limitada.

Embora a radiografia de tórax seja um exame amplamente utilizado para a triagem da TB, estudos mostram que, devido à presença de sinais sutis e muitas vezes imperceptíveis ao olho humano, a acurácia de radiologistas em relação ao exame de referência situa-se em torno de 68,7% (Geric *et al.*, 2023). Além disso, o exame confirmatório mais preciso é caro, demorado e requer infraestrutura especializada, dificultando seu acesso em regiões de baixa renda.

Essas limitações reforçam a relevância de ferramentas automatizadas, como os sistemas de diagnóstico assistido por computador (CAD), já recomendados pela OMS para apoio à detecção da tuberculose (Geric *et al.*, 2023).

Entretanto, o desenvolvimento de modelos de aprendizado profundo depende de bases de dados extensas e de alta qualidade, algo ainda raro na área de TB. A disponibilização de radiografias é limitada por políticas de privacidade, custos de aquisição e domínio privado de muitos repositórios (Liu *et al.*, 2023). Em muitos países, a realização de radiografias enfrenta desafios técnicos e econômicos adicionais (Geric *et al.*, 2023), o que contribui para a escassez de bancos de dados públicos robustos. Mesmo nos bancos disponíveis, a heterogeneidade dos equipamentos e a variabilidade da qualidade das imagens prejudicam o treinamento de modelos generalizáveis.

1.1 Objetivos

Este trabalho tem como objetivo avaliar e propor uma nova abordagem para a classificação de radiografias de tórax na detecção de tuberculose pulmonar, utilizando o conjunto TBX11K. A proposta baseia-se no uso de Curvas Principais (CP) por meio do algoritmo de k-segmentos (Verbeek; Vlassis; Kröse, 2002) como técnica de extração e seleção de características em dados de alta dimensionalidade. São investigadas duas estratégias: (i) CPE+PCA: modelagem de cada classe por uma Curva Principal Especialista (CPE), combinando PCA com o algoritmo de k-segmentos; (ii) CPE+SVD: uma variação que substitui a PCA por um método alternativo de redução de dimensionalidade, o SVD; (iii) CPE-Solo: uma variação que não utiliza redução de dimensão, utilizando apenas as CPE; e análises iniciais de (iv) CPE+Resnet+PCA: um método que utiliza uma Resnet-50 para extração de características.

Os métodos propostos são comparados com arquiteturas profundas consolidadas (VGG16, ResNet-50, DenseNet-121 e ViT), considerando métricas de desempenho, custo energético, tempo de inferência, estabilidade estatística e pegada de carbono. Busca-se, assim, analisar a viabilidade das Curvas Principais como uma alternativa eficiente e alinhada aos princípios de IA Sustentável.

1.2 Organização do Trabalho

O Capítulo 2 apresenta uma revisão bibliográfica sobre a tuberculose, as Curvas Principais, a Análise de Componentes Principais, a Decomposição em Valores Singulares (SVD), além de discutir aspectos a respeito da IA de Confiança e da IA Sustentável

O Capítulo 3 apresenta uma descrição detalhada do banco de dados, do pré-processamento, da implementação do método CPE, bem como dos métodos utilizados para comparação: VGG16, Resnet-50, ViT e DenseNet-121. E protocolos de reprodutibilidade.

O Capítulo 4 apresenta e analisa os resultados obtidos.

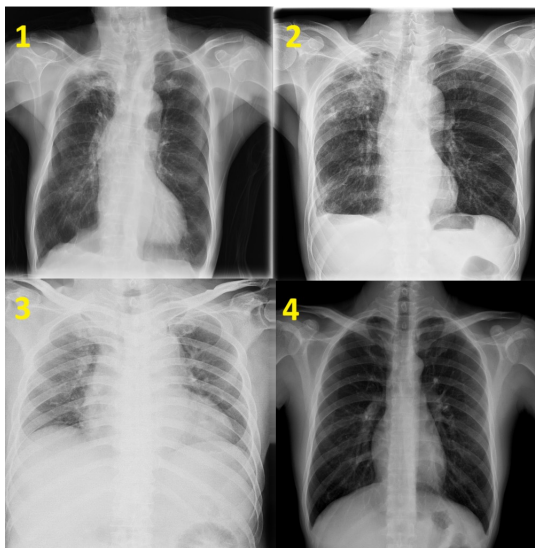
O Capítulo 5 apresenta conclusões e sugestões para trabalhos futuros.

1.3 Desafios e Motivação

O diagnóstico automatizado da tuberculose enfrenta desafios que combinam limitações computacionais, restrições de acesso a dados e necessidades práticas de aplicabilidade em cenários reais. Métodos de aprendizado profundo frequentemente exigem grande capacidade de processamento, apresentam baixa interpretabilidade e demandam longos tempos de treinamento. Exemplos encontrados na literatura para problemas de tuberculose utilizam uma combinação de ferramentas, mas as mais constantes são VGG16, Ensemble com CNN e Resnet-50 (Pramana *et al.*, 2024; Pham; Nguyen; Nguyen, 2023; Dey *et al.*, 2022; Liu *et al.*, 2023). Essas características dificultam sua implementação em hospitais com orçamento reduzido, dispositivos embarcados ou sistemas móveis de triagem, que são essenciais em áreas remotas.

Uma limitação das imagens utilizadas é a alta variabilidade dessas imagens, pois as radiografias do banco de dados TBX11K, para combater o tamanho inferior dos outros bancos de dados presentes na literatura, utilizam imagens de 4 bancos de dados diferentes, sendo estes: DA, DB, Montgomery e Shenzhen. Isso gera inconsistências de contraste, resolução e artefatos, tornando difícil para os modelos generalizarem adequadamente. Como se pode ver na Figura 1.2, uma amostra das 3 classes existentes no banco de dados mostra uma grande diferença em contraste, formato e posicionamento dos pacientes entre elas. Além disso, existe um desequilíbrio entre as proporções de imagens saudáveis, doentes e específicas para TB, o que exige uma validação robusta, como o *Stratified K-Fold* (Szeghalmy; Fazekas, 2023).

Figura 1.2 – Exemplos de radiografias, imagens 1 e 2 - Pulmão com Tuberculose, 3 - Pulmão com outra doença, 4 - Pulmão Saudável.



Fonte: Adaptado de TBX11K (Liu *et al.*, 2023).

Além das limitações técnicas, há preocupações éticas relacionadas à justiça experimental, à transparência dos modelos e à reprodutibilidade dos resultados, temas cada vez mais discutidos no contexto da IA de Confiança promovida pela União Europeia (High-Level Expert Group on Artificial Intelligence, European Commission, 2019). Em paralelo, cresce a consciência sobre os impactos ambientais do treinamento de modelos de grande escala, destacando a necessidade de soluções mais eficientes e alinhadas aos princípios da IA Sustentável.

Do ponto de vista humano, doenças respiratórias representam um problema que afeta milhões de pessoas globalmente e, para muitas famílias, incluindo a do autor desta Dissertação, constituem uma vivência próxima, marcada por episódios de complicações pulmonares. Essa perspectiva reforça a relevância social do desenvolvimento de técnicas acessíveis e energeticamente eficientes, capazes de apoiar a triagem e o diagnóstico em contextos reais.

Assim, este trabalho nasce da convergência entre desafios técnicos, preocupações éticas e motivação pessoal, propondo o uso de Curvas Principais como uma alternativa leve, transparente e de baixo custo computacional, visando contribuir para diagnósticos mais acessíveis e sustentáveis.

Apesar de todos esses problemas, os materiais incluídos neste trabalho, figuras, imagens e banco de dados, foram utilizados em conformidade com suas respectivas licenças de uso, e nenhum dos conteúdos empregados contém informações pessoais identificáveis, preservando, assim, as diretrizes éticas para o uso de imagens em pesquisas científicas.

2 REVISÃO BIBLIOGRÁFICA

Este capítulo apresenta os principais conceitos teóricos que fundamentam o desenvolvimento deste trabalho. Inicialmente, discute-se a tuberculose pulmonar (TB), sua relevância e os desafios no diagnóstico por imagem. Em seguida, são introduzidos os conceitos de Curvas Principais (Trevor; Werner, 1989) e o algoritmo K-segmentos (Verbeek; Vlassis; Kröse, 2002), utilizado para sua construção.

Também são abordados os fundamentos da Análise de Componentes Principais (PCA), uma técnica amplamente utilizada para a redução de dimensionalidade em dados de alta dimensão que foi utilizada no pré-processamento.

2.1 Tuberculose

A Tuberculose é uma doença infecciosa transmitida pela bactéria *Mycobacterium tuberculosis*. Trata-se de um problema de saúde pública global que afeta a humanidade há milênios, com registros que remontam a mais de 9.000 anos. Por ser transmitida pelo ar, a tuberculose atinge com maior frequência os pulmões, embora também possa acometer o cérebro, o trato intestinal, os rins, a coluna ou praticamente qualquer tecido ou órgão, com sequelas para a vida toda (Trajman *et al.*, 2025).

Devido ao modo de transmissão aérea, a TB impacta desproporcionalmente populações que vivem em condições precárias de moradia e alta densidade populacional, como abrigos, prisões ou comunidades indígenas. A doença manifesta-se de forma ainda mais severa em indivíduos imunossuprimidos, especialmente em pessoas que vivem com HIV/AIDS (SAÚDE, 2024).

Em função de sua elevada taxa de mortalidade, a TB foi reconhecida, em 1993, como a primeira doença infecciosa a ser declarada emergência global pela Organização Mundial da Saúde (OMS). A infecção pode permanecer latente, sem manifestação clínica, ou evoluir para a forma ativa da doença, que pode ou não ser contagiosa e pode apresentar ou não sintomas. Segundo os Centros de Controle e Prevenção de Doenças (CDC). As pessoas infectadas podem atingir 23% da população mundial (CDC, 2024a), e nos dois primeiros anos após a infecção, o risco de progressão para a doença ativa é significativamente elevado (Brasil. Ministério da

Saúde. Secretaria de Vigilância em Saúde. Departamento de Vigilância das Doenças Transmissíveis, 2019), o que reforça a importância de estratégias preventivas eficazes.

Caso não seja diagnosticada e tratada precocemente, a tuberculose pode levar à morte ou deixar sequelas graves. A adoção, na década de 1980, de um tratamento padronizado e precoce levou a uma redução significativa no número de casos, especialmente em países desenvolvidos. Contudo, em países em desenvolvimento ou subdesenvolvidos, essa redução tem ocorrido de forma mais lenta. Fatores como clima, infraestrutura precária, alta densidade populacional e aspectos socioculturais contribuem para a persistência da doença. Tais elementos explicam, em parte, a dificuldade de controle da TB em países como o Brasil, onde a doença continua a ser responsável por milhões de mortes anuais (Cortez *et al.*, 2021).

A OMS estima que seria necessário um investimento anual de aproximadamente 2 bilhões de dólares para erradicar a tuberculose até 2030. No entanto, o valor máximo arrecadado foi de apenas 0,9 bilhão de dólares em 2020, um montante significativamente inferior ao necessário. Essa cifra contrasta fortemente com os dados da plataforma de financiamento Devex (Cornish, 2021), segundo os quais foram destinados cerca de 21,7 trilhões de dólares para o combate à COVID-19. Essa discrepância não se deve exclusivamente ao número de mortes causadas por cada doença, mas também à distribuição geográfica e socioeconômica das populações afetadas. A COVID-19 teve impacto global, afetando países desenvolvidos e em desenvolvimento de forma ampla, enquanto a TB se concentra principalmente nas regiões mais pobres e vulneráveis do planeta.

Em 2020, durante a pandemia de COVID-19, 1,8 milhão de pessoas morreram em decorrência da COVID-19, enquanto a tuberculose foi responsável por 1,5 milhão de mortes (Maciel *et al.*, 2022). Esses números indicam uma proximidade alarmante entre os impactos das duas doenças, embora a resposta global à tuberculose tenha sido substancialmente inferior quando se leva em conta o caráter endêmico global e de longa duração da doença.

Essa disparidade evidencia o quanto a tuberculose continua sendo uma doença negligenciada, apesar de sua letalidade. Diante desse cenário, torna-se urgente o incentivo à pesquisa para o desenvolvimento de novas formas de diagnóstico, tratamento, medicamentos e vacinas. Um estudo realizado em hospitais demonstrou uma taxa de acerto de 68,7% em comparação ao exame de referência (Liu *et al.*, 2023). O CDC recomenda a utilização do teste de sangue QuantiFERON-TB Gold Plus (QFT-Plus), cujo resultado leva cerca de seis dias, ou o teste cutâneo (PPD), com resultado em até três dias (CDC, 2024b). Essa espera pelo resultado já é

prejudicial, pois o tratamento deve acontecer o mais rápido possível. Em situações graves, é possível detectar danos pulmonares por meio de exames de imagem, evidenciando a necessidade do uso de ferramentas auxiliares para triagem e diagnóstico em pacientes que não têm acesso ao exame de referência.

2.2 Análise de Componentes Principais (PCA)

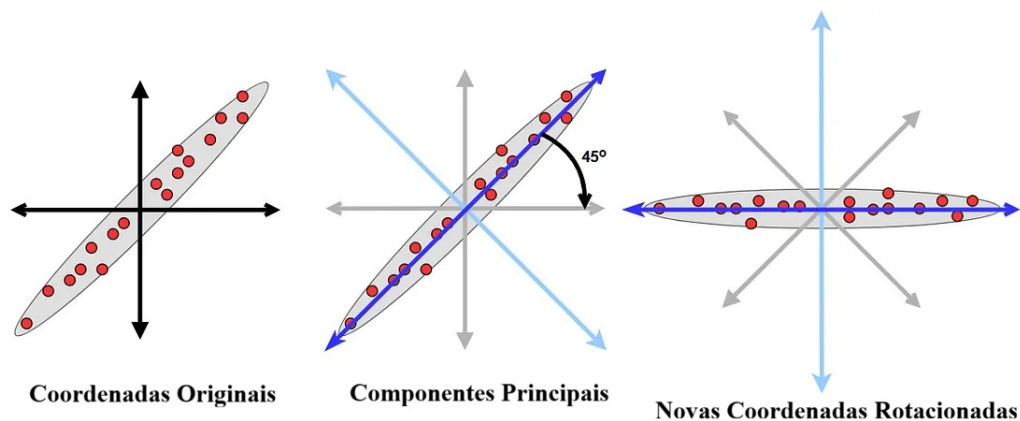
A ideia central da Análise de Componentes Principais (PCA, *Principal Component Analysis*) (Hyvärinen; Karhunen; Oja, 2001) é uma técnica estatística amplamente utilizada para a redução de dimensionalidade em conjuntos de dados multivariados. Inicialmente proposta por Karl Pearson em 1901 (Pearson, 1901), a PCA busca transformar os dados originais em um novo sistema de coordenadas formado por componentes ortogonais (componentes principais), que retêm a maior parte da variância dos dados.

Formalmente, dado um conjunto de dados centrado $X \in \mathbb{R}^{n \times d}$, onde n é o número de amostras e d o número de dimensões, o PCA resolve o seguinte problema de otimização 2.1:

$$\text{maximize } \text{Var}(X\mathbf{w}), \quad \text{sujeito a } \|\mathbf{w}\| = 1, \quad (2.1)$$

onde $\mathbf{w}$ é um vetor de projeção. A solução envolve a decomposição espectral da matriz de covariância de $\mathbf{X}$. Uma representação desse processo pode ser observada na Figura 2.1

Figura 2.1 – Representação geométrica da Análise de Componentes Principais (PCA)



Fonte: Adaptado de Günay (2023)

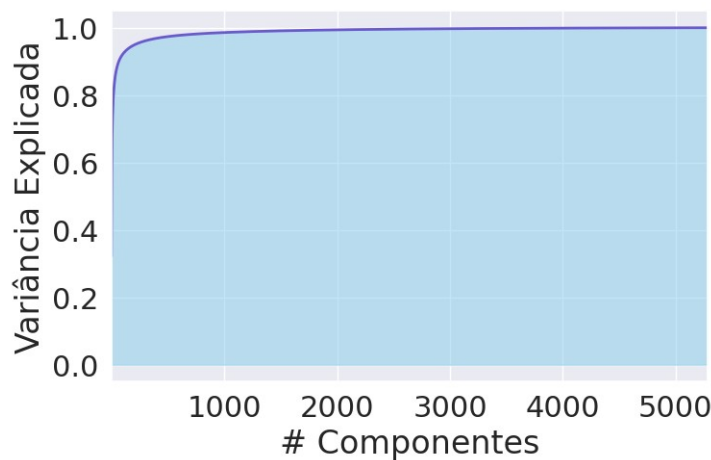
O uso do PCA traz diversas vantagens:

- Reduz a complexidade computacional ao eliminar redundâncias nos dados;

- Facilita a visualização de dados de alta dimensão;
- Atenua o risco de overfitting em modelos de aprendizado supervisionado;
- Remove ruídos e correlações espúrias entre variáveis.

A Figura 2.2 apresenta a curva da variância explicada acumulada em função do número de componentes principais extraídos. Observa-se que os primeiros componentes concentram a maior parte da variabilidade presente nos dados, evidenciada pelo crescimento acentuado da curva em seus estágios iniciais. À medida que novos componentes são incorporados, o incremento marginal na variância explicada torna-se progressivamente menor, indicando que as direções subsequentes contribuem de forma menos significativa para a representação da estrutura dos dados. Esse comportamento fundamenta a aplicação da PCA como técnica de redução de dimensionalidade, permitindo a preservação da maior parte da informação com um número reduzido de componentes.

Figura 2.2 – PCA obtido para a base de dados.



Fonte: Do Autor

Na área de visão computacional, e especialmente em imagens médicas, o PCA é comumente utilizado como etapa de pré-processamento para extrair representações mais compactas e relevantes dos dados visuais. No contexto deste trabalho, o PCA é aplicado para a extração de características, com o objetivo de selecionar as componentes mais informativas antes da aplicação das Curvas Principais.

Essa combinação visa preservar a estrutura geométrica dos dados relevantes para a classificação, ao mesmo tempo em que reduz a dimensionalidade e o custo computacional do processo de modelagem.

2.3 Decomposição em Valores Singulares (SVD)

Outra forma de redução de dimensionalidade é a Decomposição em Valores Singulares (SVD, *Singular Value Decomposition*). O SVD é uma técnica da álgebra linear usada para recuperação de informações, redução dimensional e remoção de ruídos, o que é particularmente vantajoso em problemas com alta dimensionalidade, como as imagens médicas utilizadas neste trabalho (Mariano, 2023).

Dado o conjunto de dados centrado $X \in \mathbb{R}^{n \times d}$, o SVD o fatoriza na forma 2.2:

$$X = U\Sigma V^T, \quad (2.2)$$

onde U e V são matrizes ortonormais, e Σ é uma matriz diagonal contendo os valores singulares σ_i . Esses valores singulares estão diretamente relacionados à variância explicada em cada componente principal. A relação entre SVD e PCA surge do fato de que os autovetores da matriz de covariância podem ser obtidos pela própria matriz V , enquanto a variância associada a cada componente é proporcional a σ_i^2 .

Como o SVD possui essas similaridades com o PCA, isso faz com que também produza o mesmo conjunto de componentes principais quando aplicado à centralização dos dados. O gráfico apresentado na Figura 2.2, que relaciona o número de componentes e a variância acumulada, seria essencialmente idêntico se obtido através do SVD, visto que ambos compartilham os mesmos autovalores e autovetores subjacentes.

No contexto deste trabalho, o entendimento do SVD reforça a explicação teórica por trás da seleção de componentes principais. E será usado em mais um método para demonstrar o uso de CPEs.

2.4 Cálculo de Curvas Principais pelo Método K-seg

2.4.1 Introdução

Curvas Principais (*Principal Curves*) são generalizações não lineares da Análise de Componentes Principais (PCA), cuja finalidade é representar a trajetória central de uma nuvem de pontos em um espaço de características, mesmo quando essa trajetória apresenta curvatu-

ras pronunciadas (Trevor; Werner, 1989). Em contraste com a PCA, que fornece subespaços lineares, as Curvas Principais descrevem uma curva suave que passa “pelo meio” dos dados, oferecendo uma representação mais adequada quando a estrutura intrínseca é não linear.

A implementação computacional das Curvas Principais utilizada neste trabalho baseia-se na biblioteca desenvolvida por Borges (2024), que implementa o algoritmo K-seg proposto por Verbeek, Vlassis & Kröse (2002), permitindo a estimação eficiente de Curvas Principais por meio de uma cadeia de segmentos lineares (poligonal), reduzindo a complexidade de estimação e oferecendo estabilidade numérica e convergência prática. Nesta seção, descreve-se o conceito, a formulação matemática, o algoritmo e suas aplicações para classificação em imagens médicas (Peng *et al.*, 2018; Fernandez, 2005).

2.4.2 Definição formal

Seja $\mathbf{x} \in \mathbb{R}^n$ uma observação do conjunto de dados. Uma Curva Principal $\gamma(t)$, $t \in \mathcal{T} \subset \mathbb{R}$, satisfaz a propriedade de *autoconsistência* 2.3:

$$\gamma(t) = \mathbb{E}[\mathbf{x} \mid \text{Proj}_{\gamma}(\mathbf{x}) = \gamma(t)], \quad (2.3)$$

onde $\text{Proj}_{\gamma}(\mathbf{x})$ denota a projeção ortogonal de $\mathbf{x}$ sobre a curva γ (Trevor; Werner, 1989). Essa condição implica que a curva é definida de forma que, em cada ponto $\gamma(t)$, a média das observações projetadas naquele ponto coincide com o próprio ponto da curva.

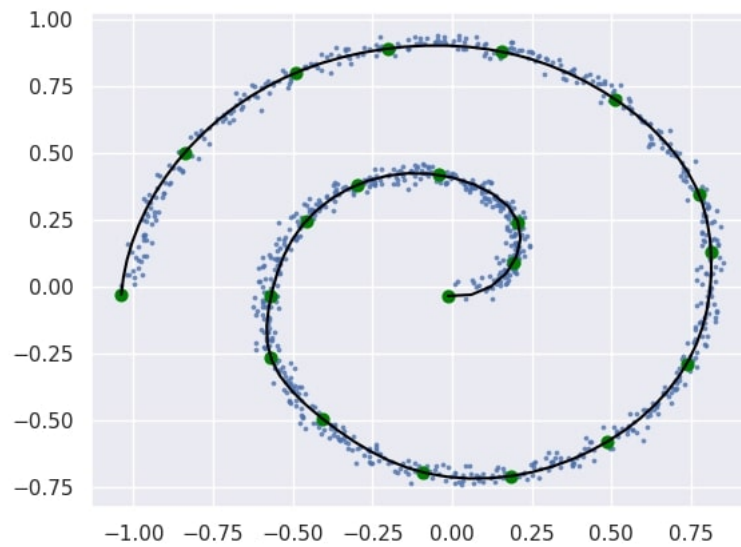
A Figura 2.3 ilustra a criação de uma Curva Principal.

2.4.3 Motivação para a aproximação por segmentos

A estimação de curvas suaves imposta por *splines* ou funções paramétricas pode ser computacionalmente custosa e sensível a mínimos locais. A aproximação em segmentos lineares (K-seg) apresenta vantagens práticas:

- implementação simples e interpretável;
- menor suscetibilidade a mínimos locais em relação a ajustes contínuos;
- custo computacional reduzido e convergência prática observada em aplicações.

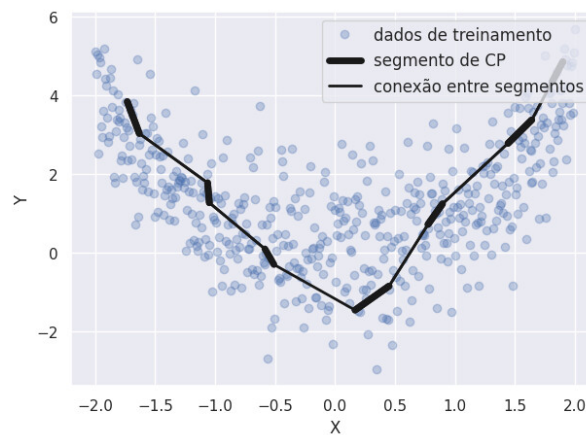
Figura 2.3 – Criação de uma Curva Principal



Fonte: Do Autor

A Figura 2.4 ilustra uma Curva Principal aproximada por 7 segmentos, aplicada a um conjunto sintético bidimensional.

Figura 2.4 – Exemplo de Curva Principal construída com 7 segmentos ($K=7$).

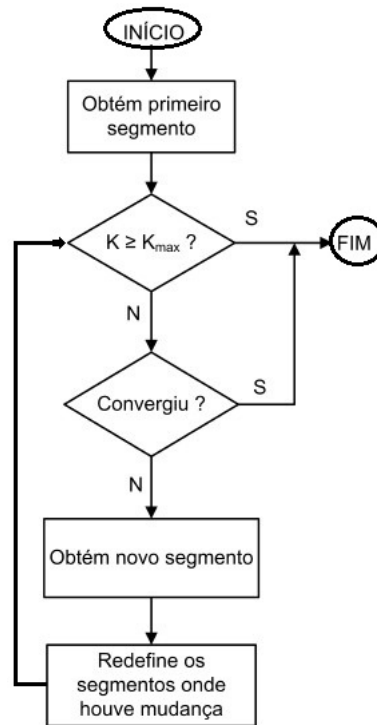


Fonte: Do Autor

2.4.4 Descrição do método K-seg

O método K-seg constrói uma curva poligonal composta por K segmentos conectados. A seguir, descreve-se o fluxo geral do algoritmo e suas etapas principais; um fluxograma do processo é apresentado na Figura 2.5.

Figura 2.5 – Fluxograma do algoritmo K-seg para obtenção da Curva Principal.



Fonte: Do Autor

1. **Inicialização.** Obtém-se o primeiro segmento da curva principal (correspondente a $K = 1$), tipicamente inicializado a partir da primeira componente principal (PCA) dos dados.
2. **Verificação do número máximo de segmentos.** Caso $K \geq K_{max}$, o algoritmo é encerrado; caso contrário, prossegue-se para a etapa iterativa.
3. **Projeção e parametrização.** Para cada observação $\mathbf{x}_i$, calcula-se sua projeção ortogonal sobre a curva atual, determinando o segmento correspondente e sua posição paramétrica ao longo da curva poligonal.
4. **Verificação de convergência.** Avalia-se se houve alteração significativa na configuração dos segmentos. Caso o critério de convergência seja satisfeito, o algoritmo é encerrado.
5. **Inserção de novo segmento.** Caso não haja convergência, incrementa-se o número de segmentos ($K \leftarrow K + 1$), obtendo-se um novo segmento na região de maior erro de aproximação.
6. **Redefinição local dos segmentos.** Apenas os segmentos cuja associação de pontos foi modificada são reajustados, por meio de mínimos quadrados ortogonais, atualizando a curva poligonal.

7. **Iteração.** Retorna-se à etapa de verificação do número máximo de segmentos.

2.4.5 Observações Importantes

Alguns pontos importantes para se avaliar cuidadosamente:

- O ajuste ortogonal de segmentos pode ser implementado por métodos de regressão total *least squares* ou por minimização da soma das distâncias ortogonais ao segmento.
- Para evitar agrupamentos degenerados, recomenda-se impor um tamanho mínimo de amostra por segmento (ex.: ≥ 3 observações) durante a etapa de segmentação.
- A inicialização pode influenciar a velocidade de convergência; usar a PCA como estimativa inicial costuma ser suficiente na prática (Verbeek; Vlassis; Kröse, 2002).

A seleção do número de segmentos K é crítica: um K pequeno pode subestimar a curvatura; um K grande pode causar sobreajuste. Estratégias usuais incluem:

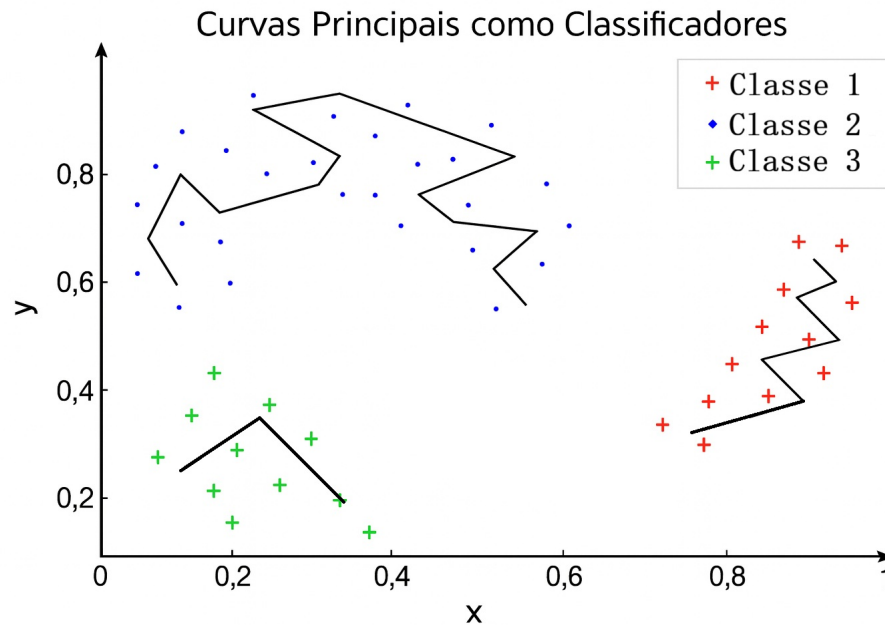
- validação cruzada em termos de erro de projeção médio;
- critérios de informação penalizados (AIC/BIC) adaptados ao contexto de segmentos;
- análise visual e heurística baseada na complexidade esperada da curva (domínio do problema).

2.4.6 Curvas Principais aplicadas à classificação

A Figura 2.6 ilustra o uso de Curvas Principais como classificadores em uma abordagem supervisionada. Para cada classe, ajusta-se uma curva principal que segue a estrutura geométrica dos dados dessa classe. Um novo ponto é classificado de acordo com a menor distância entre esse ponto e cada uma das curvas principais. Assim, as curvas funcionam como modelos representativos das classes, permitindo capturar distribuições não lineares e orientações complexas.

Curvas Principais (e sua versão segmentada K-seg) já foram aplicadas em tarefas de contorno e caracterização anatômica, bem como na extração de características para classificação de imagens médicas (Peng *et al.*, 2018; Fernandez, 2005). As principais vantagens são a

Figura 2.6 – Exemplo do uso de Curvas Principais como classificador.



Fonte: Do Autor

interpretabilidade, o baixo custo computacional relativo e a boa performance em problemas com estruturas unidimensionais ou quase-unidimensionais embutidas em espaços de alta dimensão.

2.4.7 Limitações

O método das curvas principais possui algumas limitações; essas limitações incluem:

- Sensibilidade à escolha de K e à densidade amostral ao longo da curva.
- Incapacidade de impor suavidade contínua intrínseca (a menos que se aplique pós-processamento de suavização).
- Restrição a estruturas aproximadamente unidimensionais; dados com múltiplas cadeias ou topologia complexa podem exigir abordagens mais sofisticadas.

2.5 Revisão das arquiteturas comparativas utilizadas

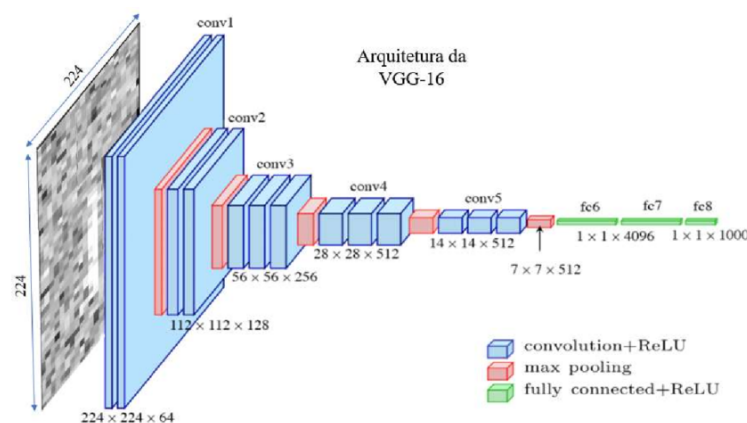
Nesta seção, são apresentadas as arquiteturas de referência amplamente utilizadas na literatura em visão computacional e em aplicações com radiografias de tórax. Incluem-se mo-

delos convolucionais clássicos (VGG16), arquiteturas com conexões otimizadas (ResNet-50 e DenseNet-121) e modelos baseados em mecanismos de atenção (Vision Transformer). Todas essas redes possuem implementações consolidadas e, de modo geral, empregam pesos pré-treinados no *ImageNet*, o que favorece a adaptação a tarefas médicas, mesmo com bases menores.

2.5.1 VGG16

A VGG16, proposta por (Simonyan; Zisserman, 2014), tornou-se um marco na evolução das CNNs ao demonstrar que o uso de múltiplas convoluções pequenas (3×3), empilhadas em profundidade, é altamente eficiente para a extração de características. A arquitetura obteve resultados de destaque no *ImageNet*, influenciando diversas redes subsequentes. A Figura 2.7 mostra um exemplo de uma arquitetura padrão da VGG16 (Falqueto; Paes; Passaro, 2023).

Figura 2.7 – Exemplo de uma VGG16.



Fonte: Adaptado de Falqueto, Paes & Passaro (2023)

2.5.1.1 Características Arquiteturais

A VGG16 tem uma estrutura com as seguintes características:

- 16 camadas com pesos treináveis (13 convolucionais + 3 totalmente conectadas).
- Convoluções 3×3 com *stride* 1 e *padding* 1.
- Redução espacial por max-pooling.

- Número muito elevado de parâmetros (≈ 138 milhões), tornando o modelo pesado para fins computacionais.

2.5.1.2 Pontos Fortes e Limitações

Vantagens:

- Arquitetura simples e de fácil interpretação.
- Influenciou diversas redes subsequentes.

Limitações:

- Custo computacional elevado e de memória.
- Suscetível ao problema de gradiente de desaparecimento em profundidades maiores.

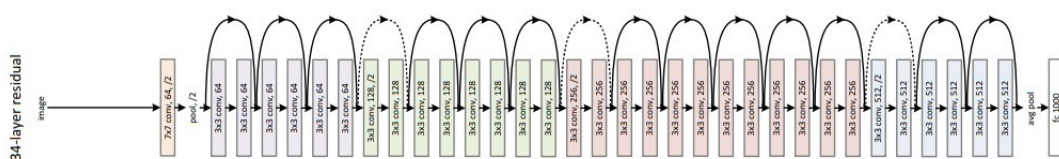
Apesar de suas limitações, a VGG16 continua sendo um baseline relevante devido à sua robustez e importância histórica; ainda é usada como base para alguns métodos (Liu *et al.*, 2016).

2.5.2 Resnet50

A família ResNet, proposta por (He *et al.*, 2016), revolucionou o campo ao introduzir o conceito de aprendizado residual, permitindo o treinamento de redes profundas de forma mais estável. A ResNet-50, composta por 50 camadas, tornou-se uma das arquiteturas amplamente utilizadas em tarefas de visão computacional devido à sua elevada capacidade representacional.

A Figura 2.8 apresenta uma representação ilustrativa da arquitetura ResNet-34, utilizada para fins didáticos na explicação dos blocos residuais. Embora estruturalmente semelhante, a ResNet-50 diferencia-se por empregar blocos do tipo *bottleneck* e por sua maior profundidade.

Figura 2.8 – Representação ilustrativa da arquitetura ResNet-34.



Fonte: He *et al.* (2016)

2.5.2.1 Conexões Residuais

A principal inovação é a conexão de atalho, que adiciona a entrada diretamente à saída do bloco 2.4:

$$y = F(x) + x \quad (2.4)$$

Essa estratégia facilita o fluxo de gradientes e evita o problema de degradação, tornando possível treinar redes com centenas de camadas.

2.5.2.2 Arquitetura em Blocos Tipo *Bottleneck*

A ResNet-50 utiliza blocos residuais com três convoluções:

- $1 \times 1 \rightarrow$ redução de dimensionalidade.
- $3 \times 3 \rightarrow$ processamento espacial.
- $1 \times 1 \rightarrow$ restauração da dimensionalidade original.

Esse arranjo reduz o custo computacional, mantendo desempenho elevado.

2.5.2.3 Pontos Fortes e Limitações

Vantagens:

- Treinamento estável, mesmo em redes profundas.
- Boa relação entre desempenho e número de parâmetros (≈ 25 milhões).
- Flexibilidade e generalização amplamente comprovadas.
- Largamente utilizada em aplicações médicas devido à robustez do aprendizado residual (Liu *et al.*, 2023).

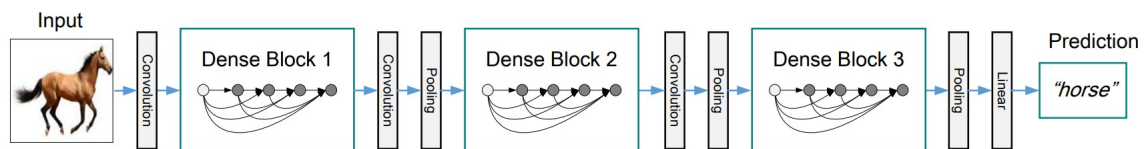
Limitações:

- Ainda apresenta um custo computacional elevado.
- Mantém o viés local das convoluções.

2.5.3 Densenet-121

A DenseNet, introduzida por Huang *et al.* (2017), evolui o conceito de conectividade ao introduzir conexões densas. Na DenseNet-121, cada camada recebe como entrada todas as saídas das camadas anteriores dentro do mesmo bloco, promovendo uma forte reutilização de características. A Figura 2.9 mostra sua arquitetura.

Figura 2.9 – Representação da Arquitetura da Densenet



Fonte: Huang *et al.* (2017)

2.5.3.1 Conectividade Densa

Cada camada recebe como entrada a concatenação das saídas de todas as camadas anteriores do mesmo bloco 2.5:

$$x_l = H_l([x_0, x_1, \dots, x_{l-1}]) \quad (2.5)$$

Essa estratégia promove a forte reutilização de características e reduz redundâncias.

2.5.3.2 Estrutura Arquitetural

A DenseNet-121 é composta por:

- 4 blocos densos.
- Camadas de transição (convolução 1×1 + *average pooling* 2×2).
- *Growth rate* típico de 32, indicando quantos mapas de características cada camada adiciona à concatenação.

2.5.3.3 Pontos Fortes e Limitações

Vantagens:

- Número reduzido de parâmetros (Resnet-50 -> 25 milhões; VGG16 -> 100 milhões; DenseNet-121 -> $\approx$ 8 milhões).
- Fortalece o fluxo de gradientes e reduz a redundância.
- Excelente desempenho mesmo com conjuntos de dados menores, importante em aplicações médicas (Ronneberger; Fischer; Brox, 2015).

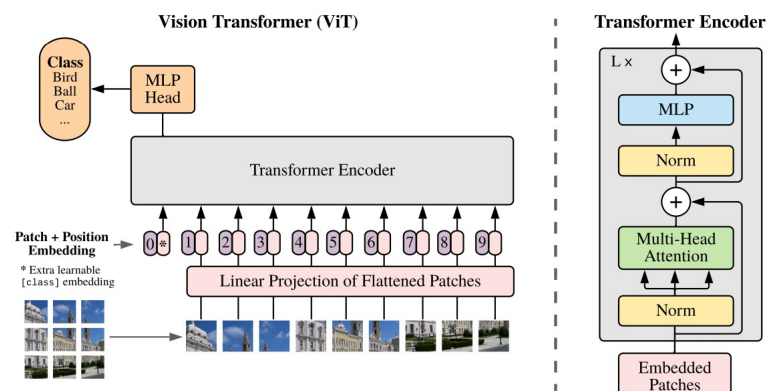
Limitações:

- A concatenação crescente aumenta a exigência de memória.
- Inferência mais lenta para imagens de alta resolução.

2.5.4 Vision Transformer (ViT)

O *Vision Transformer* (ViT), introduzido por (Dosovitskiy *et al.*, 2020), representa uma mudança de paradigma ao substituir completamente as convoluções por mecanismos de atenção. Adaptado do Transformer originalmente aplicado em linguagem natural, o ViT opera sobre blocos (*patches*) da imagem. A Figura 2.10 mostra sua arquitetura.

Figura 2.10 – Representação da Arquitetura da ViT16



Fonte: Dosovitskiy *et al.* (2020)

2.5.4.1 Formação dos *Patches*

A imagem é dividida em *patches* fixos. Cada patch é achatado e projetado em um vetor de *embedding* 2.6:

$$z_0 = [x_{p1}E; x_{p2}E; \dots] \quad (2.6)$$

Em seguida, somam-se codificações posicionais que permitem ao modelo preservar a estrutura espacial.

2.5.4.2 Codificador Transformer

Cada bloco é composto por:

- Mecanismo de *Multi-Head Self-Attention*.
- Normalização de camada.
- Rede MLP.
- Conexões residuais.
- A autoatenção permite a interação global entre todos os *patches* desde as camadas iniciais.

2.5.4.3 Pontos Fortes e Limitações

Vantagens:

- Campo receptivo global imediato.
- Escala muito bem com grandes bases de dados.
- Desempenho competitivo em diversas tarefas visuais.

Limitações:

- Exige bases grandes ou fortes, além de regularização para evitar *overfitting*.
- Custo computacional elevado, especialmente em imagens grandes.

- Menor viés indutivo do que CNNs, o que pode ser desfavorável em bancos de dados médicos menores, como TBX11K.

2.5.5 Considerações Comparativas

As arquiteturas aqui revisadas representam diferentes níveis de complexidade e indutividade: CNNs tradicionais com filtros locais (VGG16), redes com otimização de gradientes (ResNet e DenseNet) e modelos baseados em atenção global (ViT). Essa diversidade é amplamente explorada na literatura, permitindo comparar como distintas estruturas aprendem padrões relevantes em radiografias de tórax e servem como base sólida para análises posteriores. Um resumo das características de forma simplificada está presente na Tabela 2.1.

Tabela 2.1 – Resumo das principais arquiteturas analisadas.

Modelo	Ano	Tipo	Nº de Parâmetros	Inovação Central
VGG16	2014	CNN	~138M	Profundidade com kernels 3×3
ResNet-50	2016	CNN	~25M	Conexões residuais
DenseNet-121	2017	CNN	~8M	Conectividade densa / Reutilização total
ViT	2020	Transformer	Variável	Atenção global baseada em patches

2.6 IA de Confiança e IA Sustentável

Esta seção tem como objetivo fomentar o debate sobre os limites conceituais da chamada IA de confiança e da IA Sustentável (*Green IA*), duas vertentes frequentemente negligenciadas em função da ênfase excessiva nos resultados alcançados, em detrimento dos processos e custos envolvidos.

2.6.1 Contextualização

Com o crescimento acelerado do uso da inteligência artificial (IA) e sua crescente capacidade de resolver problemas complexos em diversas áreas, intensifica-se a discussão sobre a necessidade de sua regulamentação. Um indicativo dessa preocupação foi a coletiva de imprensa com o Papa Leão XIV, na qual ele comparou a atual Revolução que a IA está causando no

mundo com a Revolução Industrial, evocando a encíclica *Rerum Novarum* do Papa Leão XIII, que tratava das condições de trabalho após a industrialização (Catholic News Service, 2025).

Entretanto, o rápido avanço da tecnologia frequentemente ultrapassa o ritmo do debate regulatório. Exemplos como o ChatGPT e a *AI Art* ilustram como novas aplicações surgem antes mesmo de uma compreensão ampla de seus impactos. Observa-se também que, em vários países, incluindo o Brasil, por exemplo, a PL2338/2023 de autoria do Senador Rodrigo Pacheco (Pacheco, 2023), iniciativas de regulamentação da IA têm sido propostas sem um debate profundo e fundamentado.

Nesse cenário, o trabalho de Ryan (2020) alerta para os problemas conceituais na elaboração do documento do grupo GPAN IA da União Europeia, em especial na escolha de termos ambíguos como “confiança”. Conforme argumenta Plínio Corrêa de Oliveira (Oliveira, 1965), termos com alta elasticidade semântica podem ser deturpados, comprometendo a clareza normativa; por isso, é importante definir e utilizar termos precisos quando são utilizados em documentos importantes, como normas e legislações.

O termo IA Sustentável, ou *Green AI*, refere-se a uma abordagem da inteligência artificial que prioriza a eficiência energética e a sustentabilidade ambiental no desenvolvimento e na aplicação de modelos de aprendizado de máquina. Essa perspectiva surge como uma resposta crítica ao crescimento exponencial do consumo de recursos computacionais, especialmente no treinamento de modelos de grande escala, que demandam elevadas quantidades de energia elétrica e, conseqüentemente, acarretam uma significativa pegada de carbono. A proposta do *Green AI* é promover práticas mais conscientes e sustentáveis, incentivando o desenvolvimento de soluções que equilibrem o desempenho técnico com o impacto ambiental, por meio da otimização de algoritmos, uso de hardware mais eficiente e maior transparência nos custos ambientais das pesquisas em IA.

2.6.2 IA de Confiança

Um dos problemas mais sérios relacionados à IA é a tendência à antropomorfização: à medida que a IA responde aos comandos de forma cada vez mais humana, e dado que o termo *inteligência* sugere atributos cognitivos, muitos usuários acabam atribuindo a ela capacidades humanas, inclusive valores morais. Isso se torna problemático quando se propõe estabelecer relações de *confiança* com sistemas artificiais.

Por exemplo, como apontado por Ryan (2020), a União Europeia criou a comissão *European Commission's High-level Expert Group on AI* (HLEG), onde adotou o princípio de que deve-se estabelecer um relacionamento de confiança (*trust*) com a IA. No entanto, a confiança é uma das relações interpessoais mais sérias e complexas entre os seres humanos. Atribuir esse mesmo nível de relação a um sistema computacional é uma afirmação arriscada e imprecisa.

Segundo Ketzer (2015), na filosofia moral, distingue-se entre dois tipos de confiança: 1) a confiança (*trust*), que se caracteriza por ser uma relação interpessoal mais profunda, a qual envolve boa-fé, vulnerabilidade e um possível sentimento de traição quando é quebrada; 2) a fiabilidade (*reliability*), que se refere à previsibilidade do funcionamento de objetos ou sistemas. Um exemplo: espera-se que um termômetro funcione corretamente (fiabilidade); se falhar, há um sentimento de decepção, mas não um sentimento de traição. Já nas relações humanas baseadas na confiança, quando ocorre uma traição, isso acarreta sofrimento emocional.

Como destaca Baier (1986), a questão essencial é: *em quem eu poderia confiar, de que modo e por quê?*. Isso pressupõe agentes morais humanos. Portanto, a relação que se deve ter com sistemas de IA é de fiabilidade, não de confiança. A antropomorfização da IA pode servir para isentar os verdadeiros agentes morais — os desenvolvedores e usuários — de suas responsabilidades.

Consequentemente, é mais adequado empregar o termo *IA fiável*. Isso torna possível responsabilizar os envolvidos e evita mal-entendidos em um campo tão dinâmico e complexo quanto o da inteligência artificial.

2.6.3 IA Sustentável

O avanço acelerado da inteligência artificial (IA), especialmente com o surgimento de modelos de larga escala, como os de aprendizado profundo, tem trazido benefícios significativos para diversos setores da sociedade. No entanto, esse progresso tecnológico também tem implicações ambientais relevantes, sobretudo no que diz respeito à sua pegada de carbono. Nesse contexto, surge o conceito de IA Sustentável ou *Green AI*, que propõe práticas destinadas a mitigar esses impactos (Schwartz *et al.*, 2020).

A pegada de carbono associada à IA decorre principalmente do elevado consumo de energia necessário para o treinamento e operação de modelos complexos. Modelos de linguagem de grande porte, por exemplo, demandam vastas quantidades de dados e recursos com-

putacionais para serem treinados, muitas vezes envolvendo milhares de unidades de processamento gráfico (GPUs) operando por dias ou semanas. Esse consumo energético, dependendo da matriz elétrica da região, resulta em emissões significativas de gases de efeito estufa (GEE); um dos principais causadores desse consumo exacerbado é a busca constante pelo resultado *estado-da-arte*, frequentemente privilegiada em competições que valorizam apenas a acurácia, negligenciando critérios como eficiência energética, entre outros (Schwartz *et al.*, 2020).

Um estudo realizado em 2019 pela Universidade de Massachusetts Amherst estimou que o treinamento de um único modelo de linguagem de grande porte pode emitir mais de 284 toneladas de CO₂ equivalente à aproximadamente cinco vezes a pegada de carbono de um carro durante toda a sua vida útil. Além do treinamento inicial, o uso cotidiano desses modelos em escala global, por meio de serviços em nuvem, assistentes virtuais e aplicações comerciais, também contribui para as emissões, embora em menor escala por operação individual (Schneider *et al.*, 2025).

A escolha da infraestrutura tecnológica, a eficiência energética dos data centers e a matriz energética utilizada (renovável ou fóssil) são fatores críticos que influenciam diretamente a magnitude dessa pegada. Grandes empresas de tecnologia têm investido em soluções mais sustentáveis, como a otimização de algoritmos, a utilização de hardware especializado e a adoção de fontes de energia limpa. No entanto, o crescimento exponencial da demanda por serviços baseados em IA pode anular esses ganhos de eficiência, resultando em impactos ambientais crescentes (Patterson *et al.*, 2021).

Na área médica, esse problema é mais exacerbado devido a alta pegada de carbono nos testes de imagem, como radiografias e ressonâncias (HealthManagement.org, 2023; Carvalho; Lourenço *et al.*, 2023). Como apresentado por Kocak *et al.* (2025), as emissões do treinamento e da inferência de um modelo de grande escala podem chegar a mais de 1000 vezes o gasto de um treinamento.

Portanto, é essencial que o desenvolvimento e a aplicação da inteligência artificial estejam alinhados com políticas e práticas sustentáveis. Isso inclui o estabelecimento de métricas transparentes sobre o consumo energético, incentivos ao uso de energia renovável e o desenvolvimento de modelos mais leves e eficientes, sem comprometer a eficácia das soluções.

3 METODOLOGIA

Este capítulo descreve de forma detalhada o protocolo experimental, o método que foi proposto, as arquiteturas utilizadas na avaliação, os procedimentos de treinamento, validação e teste, e o monitoramento de impacto ambiental realizado. O objetivo foi garantir uma comparação justa, reproduzível e consistente entre modelos de diferentes famílias arquiteturais, em linha com os princípios de eficiência, transparência e sustentabilidade defendidos por abordagens de IA de Confiança e a IA Sustentável, conforme discutido em trabalhos recentes (Ahmed, 2025; Pramana *et al.*, 2024; Morís, 2024; Kotei; Thirunavukarasu, 2023).

3.1 Base de dados

O banco de dados utilizado foi o TBX11K¹ (Liu *et al.*, 2023), que totaliza 11200 imagens de radiografias do tórax. Tal número indica uma quantidade razoável para o treinamento de modelos de aprendizado profundo. Um aspecto positivo deste banco de dados que impactou sua escolha é que as imagens já passam por um pré-processamento para torná-las padronizadas, o que permite o treinamento de detectores sofisticados. Essa padronização inclui a utilização de uma anotação de caixa delimitadora, que obriga todas as imagens a terem as mesmas dimensões, facilitando a leitura e o tratamento das imagens no método em discussão, assim como os métodos utilizados para comparação.

Esse banco de dados é dividido em 3 classes distintas: 800 imagens de pacientes com TB, 3800 imagens de pacientes com outras doenças e 3800 imagens de pacientes saudáveis. Pode observar na imagem 1.2 um exemplo das imagens presentes no banco de dados.

O banco de dados é amplamente adotado em estudos recentes envolvendo CNNs, Transformers e modelos de detecção (Bista *et al.*, 2023; Morís, 2024; Huy; Lin, 2023; El-Ghany *et al.*, 2024).

¹ Fonte do banco de dados: <https://mmcheng.net/tb/>; O banco de dados utilizado nos experimentos pertence ao Laboratório de Computação de Mídia da Universidade de Nankai (Nankai University Media Computing Lab) e é distribuído sob a Creative Commons Attribution 4.0 International License (CC BY 4.0) O texto completo da licença pode ser consultado em: <https://creativecommons.org/licenses/by/4.0/legalcode>.

3.2 Contextualização Metodológica

Neste cenário, este trabalho investiga a técnica das Curvas Principais (CP) (Trevor; Werner, 1989; Verbeek; Vlassis; Kröse, 2002) como alternativa ao uso de modelos profundos. As CPs são curvas autoconsistentes cujos pontos são definidos pela média das projeções dos dados, funcionando como uma generalização não linear da Análise de Componentes Principais (PCA). A técnica permite a representação de dados de alta dimensionalidade por meio de estruturas unidimensionais eficientes, com baixo custo computacional durante a fase de operação.

As CPs já demonstraram bons resultados em aplicações de reconhecimento de padrões, tais como a detecção de distúrbios elétricos em sistemas de potência (Ferreira *et al.*, 2015) e o contorno de pulmões em imagens médicas (Peng *et al.*, 2018). Recentes adaptações da técnica têm permitido sua aplicação em classificação (Borges, 2024; Moraes *et al.*, 2020), o que motiva sua utilização neste trabalho como abordagem de extração e modelagem de características em imagens de raio-X.

Neste contexto, o presente trabalho propõe, avalia e discute métodos baseados em Curvas Principais Especialistas (CPE), gerando uma CP para cada classe como uma abordagem alternativa mais parcimoniosa. Para avaliar os métodos propostos, utilizou-se o conjunto de dados TBX11K, que se destaca por sua escala, diversidade e complexidade. Esse banco contém imagens heterogêneas provenientes de diferentes fontes, com variados níveis de qualidade clínica, o que permite testar modelos em um ambiente realista. Tanto as arquiteturas profundas quanto o método CPE foram comparados sob múltiplas métricas: acurácia, precisão, sensibilidade, F1-score, tempo de inferência e custo energético medido com CodeCarbon (Courty *et al.*, 2024). A experimentação foi conduzida com rigor metodológico, incorporando validação com *Stratified K-Fold*.

3.3 Protocolo Experimental

Na tentativa de assegurar a imparcialidade da comparação entre todos os métodos avaliados, incluindo as arquiteturas utilizadas para comparação, VGG16, ResNet-50, DenseNet-121 e Vision Transformer, foi adotado um protocolo experimental padronizado. Todos os modelos foram treinados utilizando validação cruzada estratificada em 5 *folds*, garantindo equilíbrio entre as classes em cada partição e permitindo, dessa forma, avaliar a estabilidade das métricas

e reduzir a variância associada à divisão do conjunto de dados. Esse procedimento segue boas práticas recentes em estudos de classificação de tuberculose via raios-X, que enfatizam a robustez frente ao desbalanceamento e à diversidade de padrões radiológicos (Pramana *et al.*, 2024; Kotei; Thirunavukarasu, 2023).

Esse procedimento é consistente com recomendações recentes de avaliação em radiografias de tórax, especialmente em tarefas de baixa disponibilidade de dados anotados (Musinguzi; Katumba; Murindanyi, 2025).

No presente trabalho, o K-seg é empregado como etapa de extração de características e classificação em espaço reduzido, aproveitando sua eficiência computacional e interpretabilidade. Existem implementações deste método em linguagens como Python e Matlab; a biblioteca Python utilizada neste trabalho é a *ocpc-py* (Borges, 2024).

Esta pesquisa utiliza a biblioteca *codecarbon* (Courty *et al.*, 2024). Esta biblioteca foi criada para analisar códigos e hardwares e calcular uma aproximação da pegada de carbono e do gasto de energia. Como essa biblioteca utiliza a localização para calcular a pegada de carbono, dependendo da matriz energética da região, essa pegada de carbono varia; por esse motivo, decidiu-se utilizar também a métrica do gasto de energia.

3.4 Configurações do equipamento utilizado

Foi utilizado um computador do laboratório do grupo AIA (Artificial Intelligence and Automation) com as seguintes configurações:

- Sistema Operacional: Windows 11;
- Processador: i7-14700 2.1 GHz;
- Memória RAM: 64GB;
- Placa de video: RTX 4060;

3.5 Pré Processamento

Nesta seção, será apresentado o pré-processamento que teve de ser feito nos métodos.

3.5.1 Procedimentos Comuns

O banco de dados já disponibiliza uma divisão do banco para treinamento e validação, que foi utilizada para os métodos nos arquivos `TBX11K_train.txt` e `TBX11K_val.txt`. Essa divisão fornecida por Liu (Liu *et al.*, 2023) segue a seguinte configuração de imagens: Treinamento (TB-600; Doente-3000; Saudável-3000); Validação (TB-200; Doente-800; Saudável-800). Um dos motivos para a escolha dessa divisão foi, em parte, a simplicidade de utilizá-la na reprodução dos testes deste projeto, a facilidade para outras pessoas utilizarem a mesma divisão de imagens no futuro e, assim, poder ter continuidade na análise dos métodos. A divisão de imagens já segue uma aproximação ideal de 80% para treino e 20% para validação.

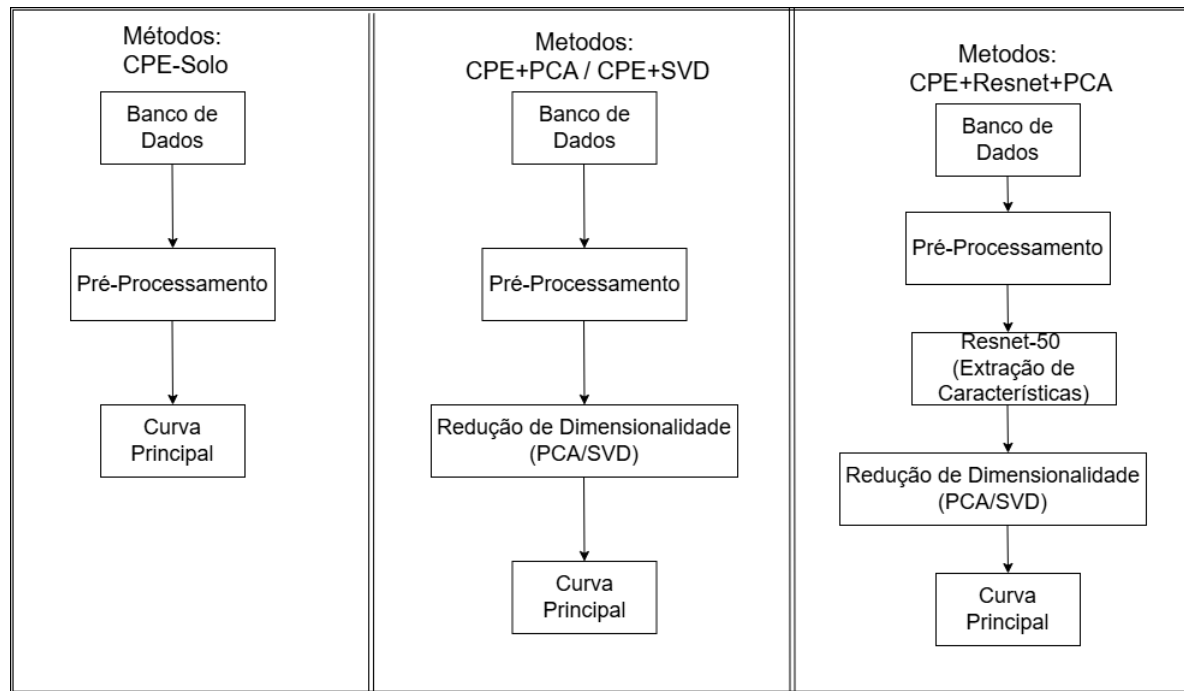
Observa-se, neste caso, ainda um significativo desbalanceamento na base de dados. A fim de minimizar o problema que esse desbalanceamento nas classes causa aos testes, será utilizado um *Stratified K-Fold*, com divisão estratificada de tamanho 5, utilizando a semente de randomização 42 para todos os métodos. O *Strat K-Fold* é um método de validação cruzada que divide os dados em K grupos, onde K deve ser maior que 1. O lado positivo dessa validação é que ela mantém a porcentagem de dados entre as classes, minimizando, assim, os problemas decorrentes do desbalanceamento do banco de dados (Szeghalmy; Fazekas, 2023), como a possibilidade de escolher um *fold* sem nenhuma imagem de uma classe.

3.5.2 Pré-processamento Específico dos Métodos Propostos

Essas imagens passaram por uma transformação para tons de cinza e um *resize* para 224×224 para os métodos CPE+PCA e CPE+SVD. Inicialmente, foi realizada a normalização dos valores das imagens, seguida pela aplicação da operação de *flatten*, que transforma a representação matricial da imagem em um vetor unidimensional. Devido as restrições de hardware que as curvas principais exigem no projeto (construção das curvas), como o alto uso da RAM, sendo necessário que todos os dados sejam carregados na RAM, o método CPE-Solo exigiu um *resize* de 32x32, demonstrando assim um limite físico desse método de analisar imagens.

O método CPE+Resnet+PCA utiliza uma Resnet-50 para fazer a extração de características; com base nessas características, utiliza-se um PCA antes de construir uma Curva Principal. A Figura 3.1 mostra um diagrama dos métodos propostos.

Figura 3.1 – Diagrama dos Métodos propostos



Fonte: Do Autor

3.5.3 Pré-processamento nas Arquiteturas Comparativas

O pré-processamento adotado segue o padrão utilizado nas principais arquiteturas modernas treinadas em ImageNet:

- normalização com médias e desvios padrão do ImageNet;
- conversão para tensores;
- redimensionamento para 224×224;
- convertidas para RGB;

Esse procedimento mantém-se consistente em todos os modelos avaliados para evitar viés entre as arquiteturas. Esse padrão está alinhado às recomendações presentes em trabalhos especializados no processamento de raios-X de tórax (Pramana *et al.*, 2024).

3.6 Método CPE+PCA: Curvas Principais Especialistas com PCA

Analizando os dados após o pré-processamento, podemos observar na Figura 2.2 que cada imagem passou a ser representada por um vetor contendo aproximadamente 5000 valores de pixel. Para simplificar os dados, foi utilizado um PCA. Para definir a quantidade de componentes finais, foi necessário observar, no gráfico, o nível de variância e localizar seu "cotovelo". Realizando uma série de testes, tanto de modo iterativo para encontrar a taxa de variância de 95%, quanto por meio de tentativa e erro para fazer uma calibração mais acurada, todos os valores coincidiram no valor de 210 componentes para o melhor resultado.

Em seguida, para cada *fold*, será montado um modelo de Curvas Principais para cada classe, utilizando a implementação descrita na Seção 2.4. Através de uma série de testes, chegou-se aos melhores resultados disponíveis na Tabela 3.1

Tabela 3.1 – Parâmetros utilizados na Abordagem I (CPE+PCA).

Algoritmos	Parâmetros	Valores
Análise de Componentes Principais	Número de Componentes	210
Curvas Principais	Número de Segmentos	55
	Alfa	1,8
	Lambda	1

Para certificar-se de que o treinamento obteve predições corretas e não tendenciosas, será selecionado o melhor modelo com as CPs que apresentaram a melhor acurácia no treinamento. A partir deste ponto, as CP escolhidas serão testadas usando o grupo de dados de validação.

3.7 Método CPE+SVD: Curvas Principais Especialistas com SVD

Como o SVD apresenta os mesmos componentes que estão na Figura 2.2, os componentes necessários para encontrar o melhor resultado são semelhantes ao método com PCA. Através de uma série de testes, chegou-se aos melhores resultados disponíveis na Tabela 3.2

Tabela 3.2 – Parâmetros utilizados na Abordagem II (CPE+SVD).

Algoritmos	Parâmetros	Valores
Decomposição por Valores Singulares	Número de Componentes	215
Curvas Principais	Número de Segmentos	48
	Alfa	1,8
	Lambda	2,8

3.8 Método CPE-Solo: Curvas Principais Especialistas

Esse método serve para demonstrar os resultados utilizando apenas uma curva principal para cada classe, sem nenhum tipo de redução de dimensões. Os parâmetros utilizados podem ser observados na Tabela 3.3.

Tabela 3.3 – Parâmetros utilizados na Abordagem III (CPE-Solo).

Algoritmos	Parâmetros	Valores
Curvas Principais	Número de Segmentos	20
	Alfa	1,8
	Lambda	1,3

3.9 Método CPE+Resnet+PCA: Curvas Principais Especialistas com extração de características usando Resnet-50 com PCA

Esse método utiliza-se de uma Resnet-50 para fazer a extração de características das imagens, e seguida esses dados passam por uma PCA para reduzir as dimensões e com esses dados reduzidos se monta as CP. Os parâmetros utilizados podem ser observados na Tabela 3.4.

Tabela 3.4 – Parâmetros utilizados na Abordagem IV (CPE+Resnet+PCA).

Algoritmos	Parâmetros	Valores
Análise de Componentes Principais	Número de Componentes	475
Curvas Principais	Número de Segmentos	5
	Alfa	1,8
	Lambda	2,8

3.10 Métodos comparativos

Todas as imagens nesses métodos foram redimensionadas para 224×224 pixels e normalizadas utilizando as estatísticas do ImageNet, uma prática comum em estudos contemporâneos de classificação de imagens médicas com CNNs e Transformers (Bista *et al.*, 2023; Morís, 2024). Além disso, sua última camada foi substituída por uma camada linear com o número de classes do TBX11K, conforme a prática comum em modelos pré-treinados aplicados a imagens de tórax.

3.10.1 VGG16

A VGG16, proposta por Simonyan & Zisserman (2014), amplamente utilizada em aplicações médicas (Simonyan; Zisserman, 2014; Liu *et al.*, 2023), baseia-se em blocos convolucionais empilhados com filtros 3×3 e camadas de pooling intermediárias. Utilizou-se a rede com pesos pré-treinados, mantendo o extrator de características com seus pesos fixos durante o treinamento e treinando apenas a camada final. Esse procedimento reduz o custo computacional e é particularmente útil em cenários com conjuntos reduzidos, como relatado em estudos recentes sobre tuberculose (Kotei; Thirunavukarasu, 2023).

3.10.2 Resnet-50

A implementação da arquitetura ResNet-50 utilizada neste trabalho foi obtida a partir da biblioteca Torchvision, integrante do ecossistema do PyTorch. Especificamente, foi empregado o módulo *torchvision.models*, conforme disponibilizado na documentação oficial da biblioteca.

A ResNet-50 emprega conexões residuais que permitem redes mais profundas sem degradação. Seguindo o mesmo protocolo da VGG16, apenas a camada classificadora final foi substituída e treinada, mantendo as demais camadas iguais. Esse modelo tem sido amplamente utilizado em aplicações de saúde, incluindo detecção de pneumonia e tuberculose (Kotei; Thirunavukarasu, 2023).

3.10.3 Densenet-121

A DenseNet-121 utiliza conexões densas entre camadas, facilitando o fluxo de gradientes e aumentando a reutilização de características. Sua aplicação em raios-X é frequentemente destacada pela boa relação entre desempenho e eficiência computacional, como sugerem trabalhos recentes (Morís, 2024).

3.10.4 ViT

O Vision Transformer representa uma mudança de paradigma ao substituir convoluções por mecanismos de atenção global. O modelo ViT-Base/16 foi utilizado, com embeddings de patches de 16×16 e 12 camadas Transformer. Estudos recentes reforçam o potencial de Transformers para imagens médicas, principalmente quando adaptados ou otimizados para cenários de poucos dados, como no RepViT-CXR (Ahmed, 2025). Assim como nas CNNs, o ViT teve todas as camadas congeladas, exceto a camada final substituída para classificação.

3.10.5 Hiperparâmetros dos modelos comparativos

Os hiperparâmetros foram padronizados na medida do possível, buscando estabelecer condições experimentais comparáveis entre os métodos avaliados. O otimizador Adam foi utilizado em todos os modelos. A Tabela 3.5 resume os valores empregados.

3.11 Seleção do Melhor Modelo

Após os 5 folds, identifica-se o modelo com a maior acurácia média de validação. Esse modelo foi, então, utilizado para a avaliação final no conjunto independente TBX11K_val, ga-

Tabela 3.5 – Hiperparâmetros utilizados nos experimentos.

Parâmetros	Valores
Taxa de aprendizado	5×10^{-5}
Batch size	32
Épocas por fold	5 (3-ViT)
Loss function	CrossEntropyLoss

mantendo a separação rigorosa entre os dados de treino e os dados de teste, preservando a validade estatística e evitando sobreajuste.

3.12 Avaliação Final

A avaliação foi realizada exclusivamente com o melhor modelo obtido de cada arquitetura. Foram geradas:

- matriz de confusão normalizada;
- acurácia final;
- precisão, revocação e F1-score por classe;
- tempo médio de inferência por imagem.

Esse procedimento segue recomendações recentes de avaliação rigorosa para aplicações médicas (Kotei; Thirunavukarasu, 2023; Liu *et al.*, 2023).

3.13 Monitoramento de Emissões e Custo Computacional

Para mensurar o impacto ambiental e energético de todo o treinamento, todos os experimentos foram monitorados com a biblioteca CodeCarbon (Courty *et al.*, 2024). Foram registradas emissões estimadas de CO₂-eq, consumo energético e tempo total do treinamento. Essa abordagem segue diretrizes recentes de sustentabilidade em IA (Ahmed, 2025; Kotei; Thirunavukarasu, 2023), permitindo avaliar não apenas a acurácia dos modelos, mas também seu custo computacional, em consonância com princípios da IA Sustentável.

O rastreador estimou:

- Energia consumida durante o treinamento
- Emissões de CO₂ equivalente

Essas informações foram automaticamente registradas em um arquivo CSV, permitindo análises posteriores sobre sustentabilidade computacional.

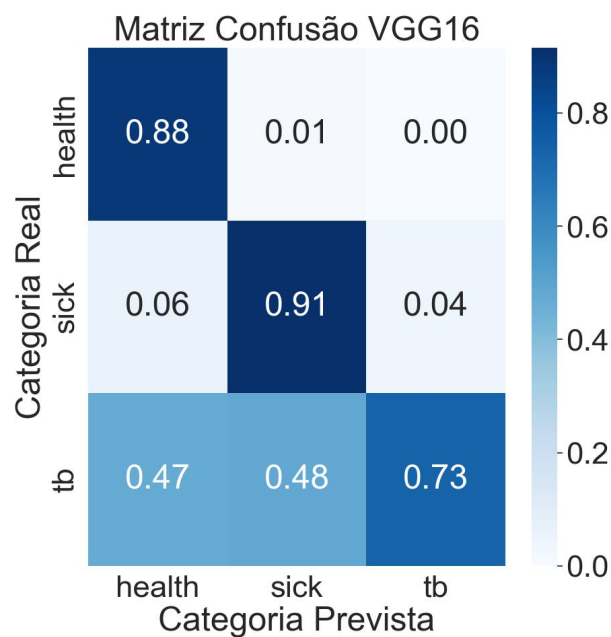
4 RESULTADOS E DISCUSSÃO

Neste capítulo, serão apresentados e analisados os resultados obtidos. Como foi testada uma série de métricas, este capítulo foi dividido em uma série de seções para cada métrica medida. Foi dada maior atenção aos resultados a respeito da classe TB, pois, devido a sua complexidade, é a maior causadora de outliers nos testes e é o foco deste trabalho.

4.1 Matriz de confusão

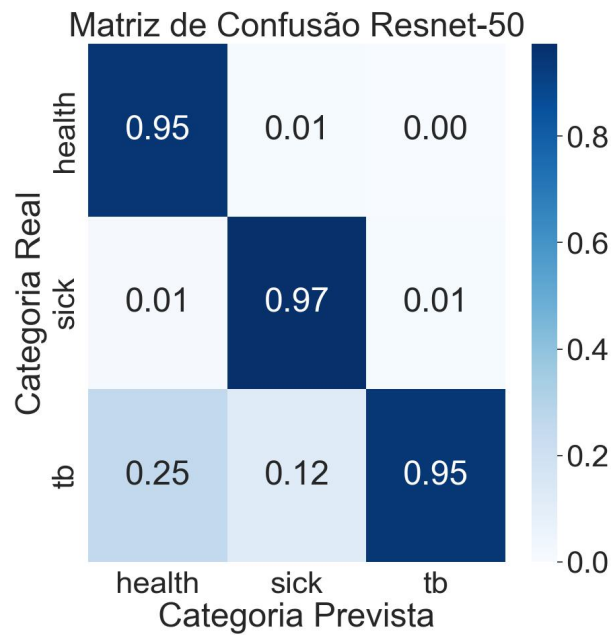
Seguem as matrizes de confusão dos modelos. Com essa representação gráfica, percebe-se a dificuldade de prever casos de TB com segurança, mostrando uma redução significativa na precisão para casos com TB.

Figura 4.1 – Matriz de confusão referente a validação do melhor fold para o classificador VGG16



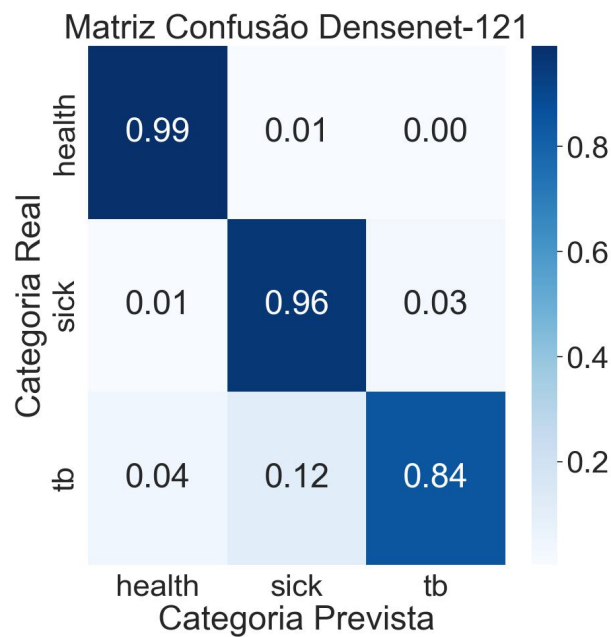
Fonte: Do Autor

Figura 4.2 – Matriz de confusão referente a validação do melhor fold para o classificador Resnet-50



Fonte: Do Autor

Figura 4.3 – Matriz de confusão referente a validação do melhor fold para o classificador Densenet-121



Fonte: Do Autor

Figura 4.4 – Matriz de confusão referente a validação do melhor fold para o classificador ViT16

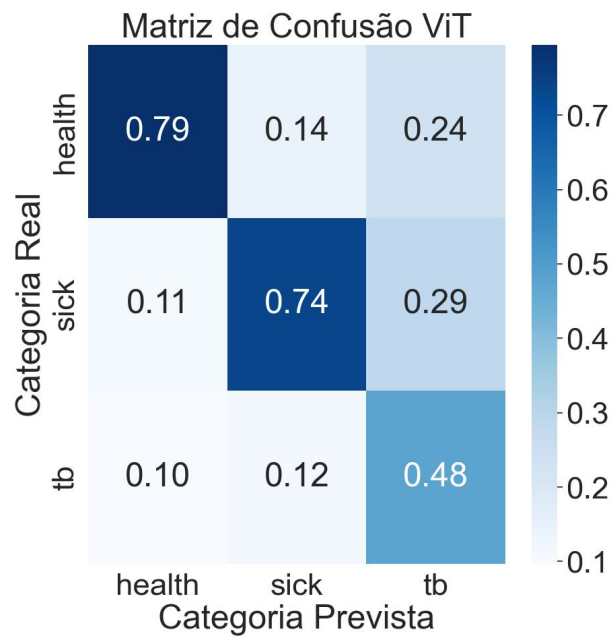


Figura 4.5 – Matriz de confusão referente a validação do melhor fold para o classificador CPE+PCA

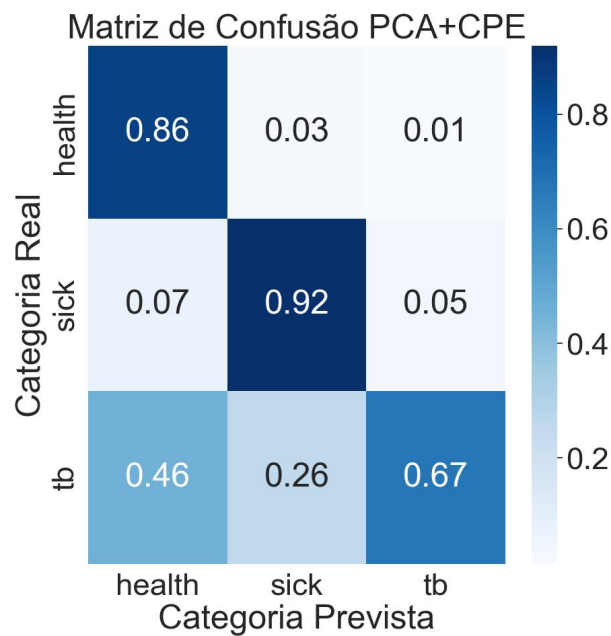


Figura 4.6 – Matriz de confusão referente a validação do melhor fold para o classificador CPE+SVD

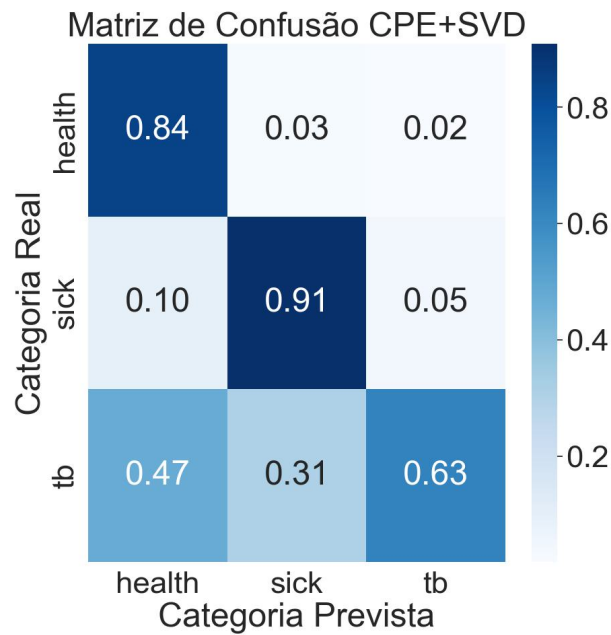


Figura 4.7 – Matriz de confusão referente a validação do melhor fold para o classificador CPE-Solo

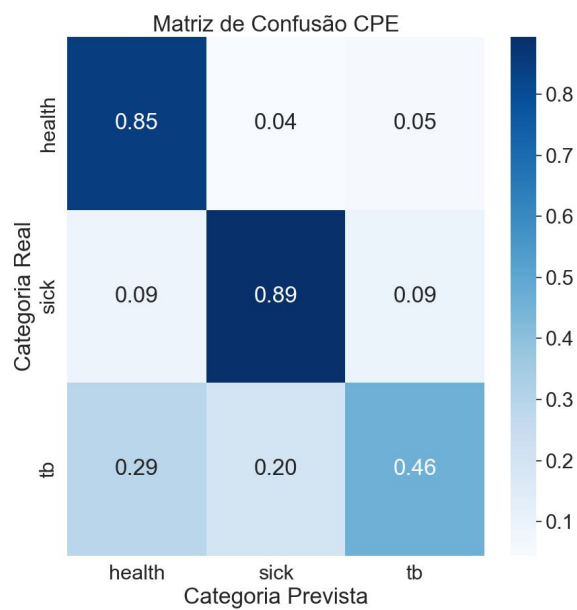
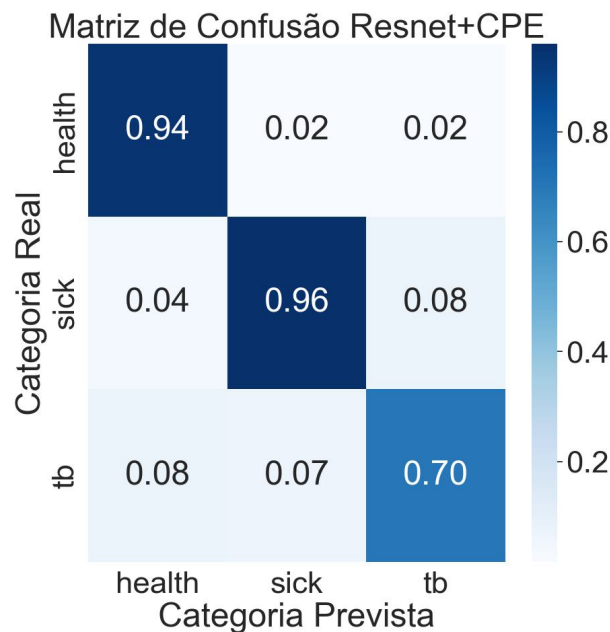


Figura 4.8 – Matriz de confusão referente a validação do melhor fold para o classificador CPE+Resnet+PCA



Fonte: Do Autor

4.2 Tempo de processamento e custo computacional

Na tabela 4.1, observa-se uma diferença expressiva no tempo total de treinamento entre os métodos analisados. Enquanto as arquiteturas profundas apresentam tempos significativamente elevados, o método proposto, baseado em Curvas Principais Especialistas, apresenta um custo computacional substancialmente menor. Esse comportamento era esperado, uma vez que o modelo não envolve propagação reversa, como nas redes neurais profundas, nem um número elevado de parâmetros, o que se reflete em uma complexidade assintótica inferior de treinamento quando comparado às arquiteturas tradicionais, como VGG16, ResNet-50, DenseNet-121 e ViT. Assim, além de reduzir o consumo energético e o tempo de execução, o método se mostra mais eficiente em termos de escalabilidade computacional.

4.2.1 Tempo de inferência

Como mencionado por Kocak *et al.* (2025), a pegada de carbono associada à inferência de aplicações baseadas em aprendizado profundo pode alcançar valores de até 1000 vezes

Tabela 4.1 – Tempo total de treinamento dos métodos comparados

Método	Tempo (horas)
VGG16	10,53 h
ResNet-50	4,51 h
DenseNet-121	4,67 h
Vision Transformer	19,53 h
CPE+PCA	0,98 h
CPE+SVD	0,93 h
CPE-Solo	2,92 h
CPE+Resnet+PCA	3,82 h

superiores aos do treinamento. Diante desse cenário, foram avaliados o tempo e as emissões associadas à etapa de inferência dos métodos propostos e comparativos. Os resultados demonstraram ganhos expressivos de eficiência, com reduções substanciais no tempo de processamento por imagem, alcançando ganhos de velocidade de até 8 vezes em relação aos modelos baseados em aprendizado profundo. Na proposta CPE+Resnet+PCA, devido a curva criada ser mais simples que as outras abordagens que também utilizam curvas, pode-se observar uma melhora significativa no tempo de inferência.

Tabela 4.2 – Tempo médio de análise por imagem (média $\pm$ desvio padrão), com 4 casas decimais.

Método	Tempo por imagem (s)
VGG16	0.1052 $\pm$ 0.0111
ResNet-50	0.0569 $\pm$ 0.0033
DenseNet-121	0.0722 $\pm$ 0.0055
ViT	0.0949 $\pm$ 0.0024
CPE+PCA	0.0204 $\pm$ 0.0075
CPE+SVD	0.0204 $\pm$ 0.0075
CPE-Solo	0.0130 $\pm$ 0.0022
CPE+Resnet+PCA	0.0019 $\pm$ 0.0010

4.2.2 Complexidade computacional

As redes profundas possuem treinamento proporcional ao número de parâmetros e camadas (tipicamente entre $O(n \cdot p)$ e $O(n \cdot p^2)$ dependendo da arquitetura), enquanto a abordagem CPE requer apenas a extração de componentes e projeções em subespaço.

Embora o algoritmo K-Segments apresente custo assintótico de $O(kn^2)$ (Verbeek; Vlasis; Kröse, 2002), no contexto desta implementação, a etapa de PCA torna-se a mais custosa devido à alta dimensionalidade inicial dos dados ($d \approx 5000$). A redução para apenas 210 componentes principais diminui significativamente o custo subsequente do classificador, fazendo com que o tempo total seja dominado pelo processo de decomposição espectral da PCA.

A decomposição espectral utilizada pela PCA possui custo computacional aproximado de $O(nd^2 + d^3)$, considerando n amostras e d variáveis, segundo (Bishop, 2006; Shlens, 2014).

4.3 Consumo energético estimado e Pegada de carbono (COO₂eq)

Analisando a Tabela 4.3, o método proposto apresenta uma grande diferença no gasto de energia, enquanto os outros métodos tem um consumo de energia similar. E isso se reflete na Tabela 4.4, onde o método proposto tem uma pegada de carbono entre ≈ 20 e ≈ 4 vezes menor que as arquiteturas de aprendizado profundo.

Tabela 4.3 – Tabela do gasto de energia (kwh)

Métodos	Energia consumida (kWh)
VGG16	0.699 kWh
Resnet50	0.300 kWh
DenseNet-121	0.310 kwh
ViT	1.296 kWh
CPE+PCA	0.065 kWh
CPE+SVD	0.061 kWh
CPE-Solo	0.194 kWh
CPE+Resnet+PCA	0.309 kWh

Tabela 4.4 – Emissões de COO₂e por modelo avaliado

Métodos	Pegada de Carbono/Kg CO₂eq
VGG16	0.111 Kg COO ₂ eq
Resnet50	0.048 Kg COO ₂ eq
DenseNet-121	0.049 Kg COO ₂ eq
ViT	0.206 Kg COO ₂ eq
CPE+PCA	0.010 Kg COO₂eq
CPE+SVD	0.010 Kg COO₂eq
CPE-Solo	0.031 Kg COO₂eq
CPE+Resnet+PCA	0.00056 Kg COO₂eq

4.4 Análise do Teste Estatístico

Devido aos métodos VGG16, ResNet50 e DenseNet121 apresentarem valores superiores de acurácia e precisão em relação aos métodos propostos, é necessário avaliar se as diferenças observadas entre os métodos de classificação são estatisticamente significativas. Os métodos passaram por um teste de Friedman, que indicou uma heterogeneidade no desempenho dos métodos para acurácia e precisão TB ($p < 0,05$). Por causa disso, os métodos passaram por um teste pós-hoc de Nemenyi para ter uma análise mais aprofundada da relação entre os métodos, conforme recomendado para comparações múltiplas em cenários com vários algoritmos avaliados sobre os mesmos conjuntos de dados.

O pós-teste de Nemenyi revelou alguns padrões principais ao focar nos métodos baseados em Curvas Principais Especialistas (CPE).

1 - Equivalência entre os métodos CPE:

Os três métodos de interesse, CPE-Solo, CPE+PCA e CPE+SVD, não apresentaram diferenças estatisticamente significativas entre si, tanto em precisão quanto em acurácia. Os valores de p para essas comparações foram consistentemente elevados (próximos de 1), indicando que a aplicação de técnicas de redução de dimensionalidade (PCA ou SVD) não alterou

de forma significativa o desempenho do classificador CPE. Sugere-se que o método CPE é intrinsecamente estável e que sua eficácia independe da escolha específica da técnica de redução de dimensão.

2 - Comparação dos métodos CPE com redes profundas:

Os métodos baseados em CPE demonstraram desempenho competitivo quando comparados às arquiteturas de redes neurais profundas avaliadas.

Para os métodos ViT e VGG16, não houve diferenças estatisticamente significativas em relação a nenhum dos métodos CPE, tanto em precisão quanto em acurácia. Esse resultado demonstra que os métodos propostos, mesmo em sua forma mais simples, são capazes de alcançar desempenho estatisticamente equivalente ao de arquiteturas de deep learning amplamente utilizadas e de maior complexidade.

Para a ResNet50 e a DenseNet121, observou-se uma diferença significativa apenas quando comparadas ao CPE-Solo, indicando um maior desempenho dessas redes em relação à versão não projetada do método CPE. No entanto, essa diferença não se manteve ao comparar ResNet50 e DenseNet121 com CPE+PCA ou CPE+SVD, o que sugere que o uso de uma etapa de redução de dimensionalidade torna o método CPE estatisticamente equivalente às redes profundas mais robustas analisadas.

Tabela 4.5 – Resultados da Acurácia (média $\pm$ desvio padrão) dos métodos

Método	Acurácia (média $\pm$ desvio padrão)
VGG16	0.8845 $\pm$ 0.0070
ResNet50	0.9506 $\pm$ 0.0089
DenseNet121	0.9615 $\pm$ 0.0119
ViT	0.7656 $\pm$ 0.0073
CPE+PCA	0.8717 $\pm$ 0.0172
CPE+SVD	0.8694 $\pm$ 0.0123
CPE-Solo	0.8165 $\pm$ 0.0092
CPE+Resnet+PCA	0.9165 $\pm$ 0.0039

Tabela 4.6 – Resultados da Precisão TB(média $\pm$ desvio padrão) dos métodos

Método	Precisão TB(média $\pm$ desvio padrão)
VGG16	0.6820 $\pm$ 0.1929
ResNet50	0.8695 $\pm$ 0.0843
DenseNet121	0.8162 $\pm$ 0.1067
ViT	0.0962 $\pm$ 0.1318
CPE+PCA	0.6208 $\pm$ 0.0592
CPE+SVD	0.6177 $\pm$ 0.0738
CPE-Solo	0.3955 $\pm$ 0.0340
CPE+Resnet+PCA	0.6521 $\pm$ 0.0431

Tabela 4.7 – Resultados da Análise Estatística: Teste de Friedman e Pós-hoc Nemenyi para Acurácia e Precisão TB.

Métrica	p-value (Friedman)	Significância ($\alpha = 0.05$)
Acurácia	0.00009	Significativo
Precisão TB	0.00016	Significativo

Tabela 4.8 – Matriz de p-valores do pós-teste Nemenyi para Acurácia.

	VGG16	ResNet50	DenseNet121	ViT	CPE+PCA	CPE+SVD	CPE-Solo	CPE+Resnet+PCA
VGG16	1.000000	0.905096	0.478211	0.163298	0.997234	0.999949	0.577610	0.999664
ResNet50	0.905096	1.000000	0.990645	0.004720	0.577610	0.766379	0.053044	0.999664
DenseNet121	0.478211	0.990645	1.000000	0.000227	0.163298	0.297637	0.004720	0.974578
ViT	0.163298	0.004720	0.000227	1.000000	0.478211	0.297637	0.990645	0.195221
CPE+PCA	0.997234	0.577610	0.163298	0.478211	1.000000	0.999949	0.905096	0.886305
CPE+SVD	0.999949	0.766379	0.297637	0.297637	0.999949	1.000000	0.766379	0.941347
CPE-Solo	0.577610	0.053044	0.004720	0.990645	0.905096	0.766379	1.000000	0.481766
CPE+Resnet+PCA	0.999664	0.999664	0.974578	0.195221	0.886305	0.941347	0.481766	1.000000

4.5 Discussão dos resultados

Nesta seção, são discutidas as diferenças de desempenho entre as arquiteturas avaliadas, com ênfase nos fatores que explicam o comportamento distinto de cada modelo. A análise considera tanto os indicadores quantitativos, como a acurácia, a precisão, a variabilidade entre

Tabela 4.9 – Matriz de p-valores do pós-teste Nemenyi para Precisão TB.

	VGG16	ResNet50	DenseNet121	ViT	CPE+PCA	CPE+SVD	CPE-Solo	CPE+Resnet+PCA
VGG16	1.000000	0.905096	0.990645	0.053044	0.948498	0.975835	0.297637	0.974578
ResNet50	0.905096	1.000000	0.999459	0.000820	0.297637	0.383365	0.013423	0.997811
DenseNet121	0.990645	0.999459	1.000000	0.004720	0.577610	0.675740	0.053044	0.999978
ViT	0.053044	0.000820	0.004720	1.000000	0.478211	0.383365	0.990645	0.132863
CPE+PCA	0.948498	0.297637	0.577610	0.478211	1.000000	0.999999	0.905096	0.886305
CPE+SVD	0.975835	0.383365	0.675740	0.383365	0.999999	1.000000	0.844041	0.941347
CPE-Solo	0.297637	0.013423	0.053044	0.990645	0.905096	0.844041	1.000000	0.372133
CPE+Resnet+PCA	0.974578	0.997811	0.999978	0.132863	0.886305	0.941347	0.372133	1.000000

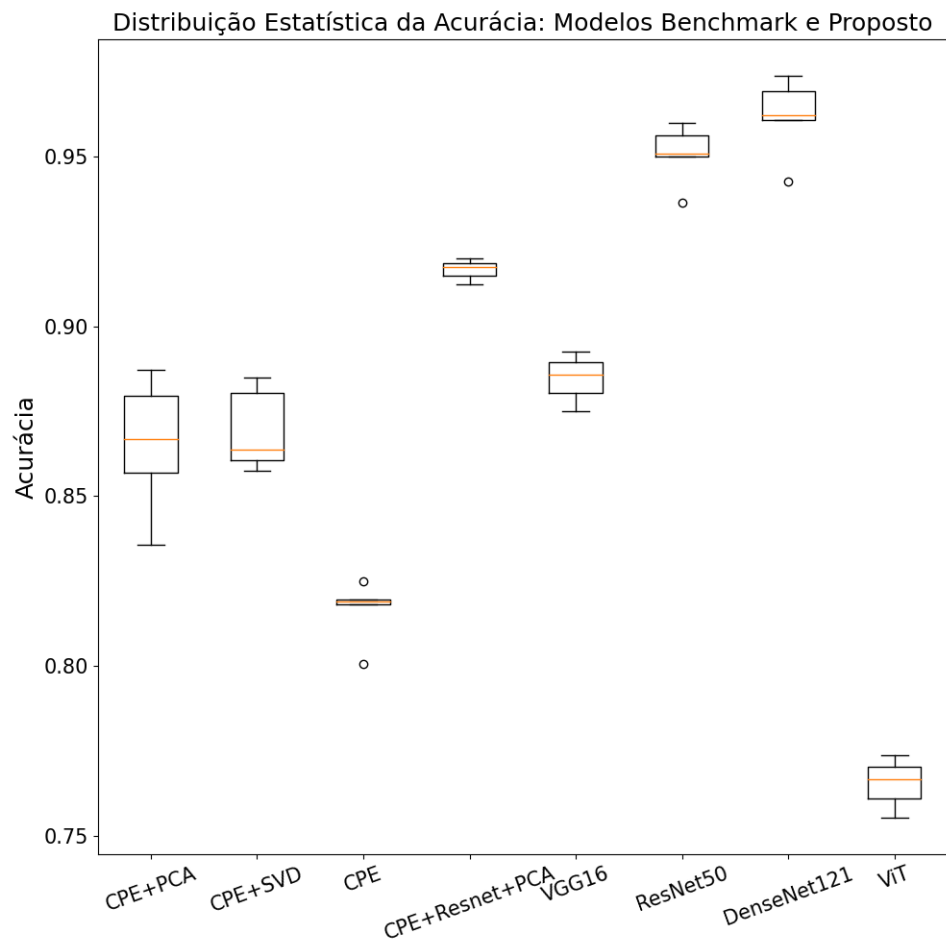


Figura 4.9 – Distribuição das acurácias dos métodos de classificação.

folds e a significância estatística, quanto os aspectos estruturais das redes que podem justificar as tendências observadas. Considerando que a classe de tuberculose (TB) constitui o principal foco deste estudo e representa o maior desafio do conjunto de dados, uma atenção especial é dada ao comportamento dessa classe ao longo dos experimentos.

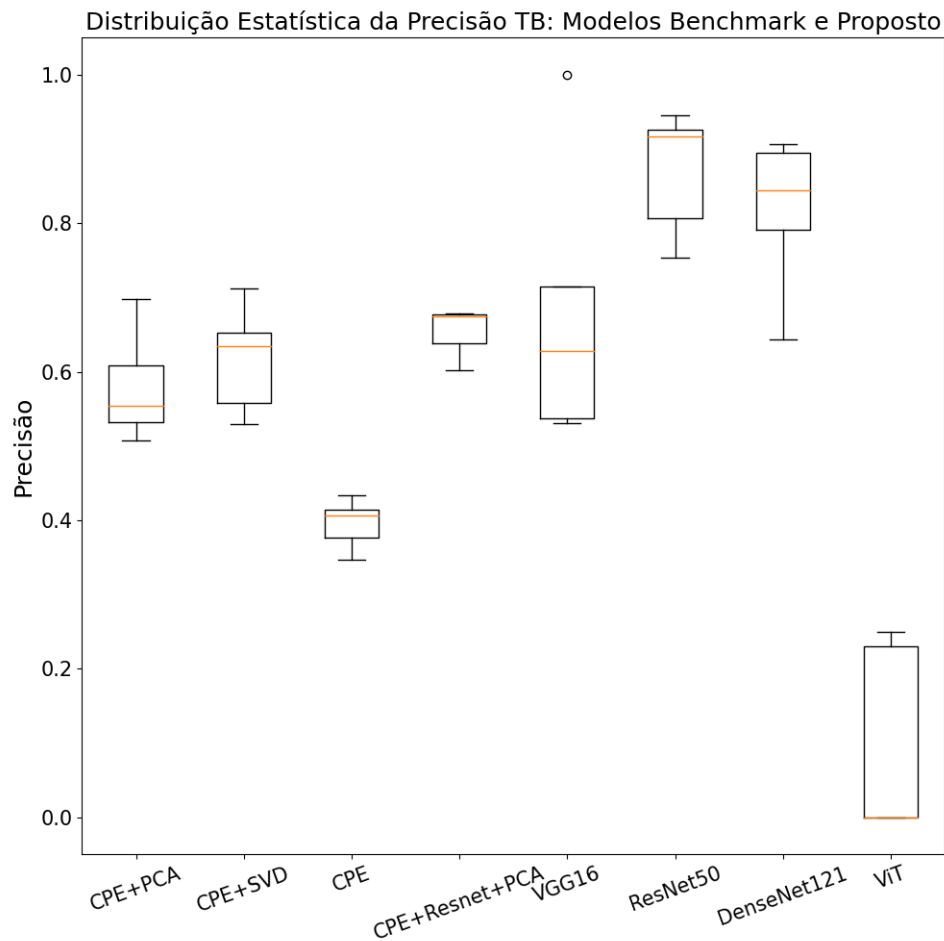


Figura 4.10 – Distribuição das precisões para TB dos métodos de classificação.

Para garantir imparcialidade nos testes, todas as arquiteturas foram treinadas exclusivamente com o conjunto TBX11K, sem qualquer pré-treinamento externo. Os resultados da Tabela 4.10 mostram que a classe TB permanece como o principal desafio do conjunto de dados, apresentando consistentemente os menores valores de Recall e F1-score em todos os métodos avaliados. Essa dificuldade decorre tanto da alta variabilidade entre radiografias de TB quanto da menor representatividade dessa classe no conjunto. Métodos baseados em redes profundas conseguem capturar padrões sutis ao custo de maior complexidade computacional, enquanto o CPE modela apenas tendências estruturais predominantes, o que limita sua capacidade de capturar variações raras ou atípicas.

As diferenças entre arquiteturas evidenciam um claro trade-off entre desempenho máximo e custo computacional. Modelos profundos, como DenseNet121 e ResNet50, apresentam métricas superiores, porém à custa de maior demanda energética e computacional. Em contraste, as abordagens baseadas em CPE oferecem tempos de inferência até dez vezes meno-

Tabela 4.10 – Comparação das abordagens propostas com outros métodos

Métodos	Acurácia	Precisão TB	Recall TB	F1 TB	Precisão Doente	Recall Doente	F1 Doente	Precisão Saudável	Recall Saudável	F1 Saudável
VGG16	0,89	0,73	0,43	0,55	0,91	0,90	0,91	0,88	0,99	0,93
Resnet-50	0,96	0,95	0,72	0,82	0,97	0,98	0,98	0,95	0,99	0,97
DenseNet-121	0,96	0,86	0,84	0,85	0,97	0,96	0,96	0,98	0,99	0,99
ViT	0,76	0,48	0,05	0,09	0,74	0,88	0,80	0,79	0,82	0,81
CPE+PCA	0,87	0,69	0,49	0,58	0,91	0,90	0,90	0,87	0,94	0,90
CPE+SVD	0,85	0,63	0,45	0,52	0,91	0,86	0,88	0,84	0,95	0,89
CPE-Solo	0,83	0,46	0,50	0,49	0,90	0,83	0,86	0,85	0,90	0,88
CPE+Resnet+PCA	0,0,92	0,75	0,82	0,78	0,94	0,91	0,93	0,94	0,96	0,95

res e consumo energético até vinte vezes inferior, mantendo um desempenho estatisticamente equivalente ao de arquiteturas como ViT e VGG16. Esse equilíbrio entre custo e desempenho sugere que métodos baseados em CPE podem ser mais adequados para cenários reais de triagem médica em ambientes com restrições de hardware, especialmente em regiões de baixa infraestrutura.

Nas abordagens CPEs, cada classe é modelada por curvas segmentadas, cujo número de segmentos reflete sua complexidade. Dada a natureza heterogênea da TB, tornou-se necessário utilizar um número maior de segmentos para representar adequadamente a classe. Após a construção das curvas, a classificação é realizada por meio de uma simples distância euclidiana entre a curva da imagem e os modelos de cada classe, o que explica a elevada eficiência computacional observada na Tabela 4.2. Devido a necessidade de todo o vetor estar na memória RAM, e sem um redimensionamento, as imagens nos testes da CPE-Solo foram forçadas a ter um redimensionamento agressivo para 32x32 pixels, ao mesmo tempo em que se limita o número de segmentos a 20, o que comprometeu seu desempenho. Essas limitações e seus resultados provam a necessidade de aplicar um pré-processamento adequado antes da extração das curvas. As Figuras 4.9 e 4.10 mostram que o método apresenta baixa variabilidade entre folds, apesar de seus valores absolutos permanecerem inferiores aos demais.

Comparando as abordagens híbridas, CPE+PCA e CPE+SVD exibem desempenhos muito próximos, conforme sugerido pelas Tabelas 4.8 e 4.9. Contudo, na análise das distribuições no boxplot, revela-se que o PCA apresenta maior estabilidade, reduzindo a variância dos resultados em relação ao SVD. Esse comportamento sugere que o PCA seleciona direções principais mais alinhadas à variabilidade discriminativa entre as classes, enquanto o SVD preserva parte do ruído estrutural da matriz de pixels, resultando em maior dispersão. Os testes estatísticos indicam que a redução de dimensionalidade exerce papel central na estabilização

do CPE: enquanto o CPE-Solo apresenta desempenho significativamente inferior aos modelos profundos, as versões CPE+PCA e CPE+SVD tornam-se estatisticamente equivalentes a arquiteturas mais robustas após a projeção para subespaços reduzidos. Isso reforça que a redução dimensional não apenas elimina redundâncias, mas direciona o modelo para regiões do espaço de características que favorecem a separabilidade entre classes, aumentando sua consistência entre folds.

A Tabela 4.11 apresenta uma comparação entre as abordagens propostas e métodos consolidados na literatura. Observa-se que modelos baseados em arquiteturas profundas para detecção e classificação, como Faster R-CNN, FCOS e Deformable DETR, alcançam acurácias entre 0,88 e 0,91, impulsionados principalmente pelo uso de backbones robustos, como *ResNet-50* com *Feature Pyramid Networks* (FPN). O melhor desempenho reportado é obtido pela abordagem baseada em *Normalization-Free Networks* combinada com *RandAugment*, que atinge 0,96, um reflexo direto do uso de arquiteturas altamente otimizadas e técnicas avançadas de aumento de dados. Em contraste, as abordagens propostas neste trabalho, baseadas em Curvas Principais com diferentes estratégias de redução de dimensionalidade (PCA e SVD), apresentam acurácias entre 0,85 e 0,87, valores competitivos, considerando que utilizam modelos substancialmente mais leves e não dependem de arquiteturas profundas de detecção. Embora não alcancem o desempenho dos métodos de ponta, as técnicas propostas situam-se no mesmo patamar de soluções intermediárias, como SSD e RetinaNet, com vantagens claras em eficiência e simplicidade.

Adicionalmente aos experimentos originalmente planejados neste trabalho, foi conduzida, durante o processo de revisão desta dissertação, uma análise exploratória utilizando uma abordagem híbrida CPE+ResNet+PCA; os resultados obtidos com essa abordagem são mais semelhantes aos métodos de aprendizado profundo, mas sem a pegada de carbono tão significativa.

Os achados também estão alinhados com princípios de IA Sustentável, que incentivam o uso de soluções energeticamente eficientes e com menor impacto ambiental. O CPE apresenta consumo computacional reduzido, menor pegada de carbono e alta velocidade de inferência, características desejáveis em aplicações clínicas reais e em ambientes hospitalares com recursos limitados. Esse conjunto de propriedades reforça o potencial do método como uma alternativa ambientalmente responsável.

Apesar dos resultados promissores, algumas limitações devem ser reconhecidas. Além disso, a persistente dificuldade na classificação da TB indica que características discriminativas

Tabela 4.11 – Comparação das abordagens propostas com outros métodos da literatura

Métodos	Backbones	Acurácia
SSD (Liu <i>et al.</i> , 2023)	VGGNet-16	0,84
RetinaNet (Liu <i>et al.</i> , 2023)	ResNet-50 w/ FPN	0,87
Faster R-CNN (Liu <i>et al.</i> , 2023)	ResNet-50 w/ FPN	0,89
FCOS (Liu <i>et al.</i> , 2023)	ResNet-50 w/ FPN	0,88
Deformable DETR (Liu <i>et al.</i> , 2023)	ResNet-50 w/ FPN	0,91
Normalization-free network and RandAugment (Acharya <i>et al.</i> , 2022)	NFNet	0,96
Abordagem CPE+PCA	Curvas Principais	0,87
Abordagem CPE+SVD	Curvas Principais	0,85
Abordagem CPE-Solo	Curvas Principais	0,83
Abordagem CPE+Resnet+PCA	Curvas Principais	0,92

essenciais podem estar distribuídas espacialmente de forma complexa, excedendo a capacidade de representação das curvas principais. Como direções futuras, recomenda-se ampliar os experimentos para bases de maior escala, integrar o CPE com técnicas auto-supervisionadas, explorar novos métodos de redução de dimensionalidade e investigar variantes dos K-Segmentos capazes de capturar descontinuidades anatômicas relevantes.

5 CONCLUSÃO

O presente trabalho teve como objetivo desenvolver e avaliar uma nova abordagem para o diagnóstico de TB na triagem de pacientes, fundamentada no uso de Curvas Principais (CPE). Os resultados experimentais obtidos demonstraram que a combinação entre CPE e técnicas de redução de dimensionalidade gera um classificador competitivo em relação aos métodos tradicionais da literatura, apresentando desempenho consistente e acurácia satisfatória. Dessa forma, esta dissertação atua como uma prova de conceito de que métodos parcimoniosos continuam relevantes, sobretudo diante do crescente interesse por soluções computacionais mais eficientes e com menor pegada de carbono.

Do ponto de vista da inovação, o método proposto apresentou contribuições em três pontos principais:

- (i) demonstrou a viabilidade do uso de CPE para a detecção de TB, indicando potencial de aplicação em outros problemas de reconhecimento de padrões;
- (ii) possui baixo custo computacional e reduzido tempo de processamento, o que favorece sua adoção em dispositivos com recursos limitados, como microcontroladores e smartphones, ampliando a possibilidade de universalização do diagnóstico em regiões com infraestrutura restrita;
- (iii) apresenta uma menor pegada de carbono em comparação a arquiteturas de deep learning, contribuindo para o desenvolvimento de técnicas alinhadas aos princípios da IA Sustentável.

Apesar de modelos profundos apresentarem acurácias superiores em alguns cenários, observou-se que seu processo de treinamento é substancialmente mais custoso em termos de tempo e energia. O caso do Vision Transformer (ViT) ilustra esse ponto, onde o tempo de treinamento foi tão elevado que foi necessário usar menos épocas, resultando em maior impacto ambiental e maior tempo de análise por imagem em comparação ao método baseado em CPE. Esses resultados fortalecem o debate sobre o equilíbrio entre desempenho, eficiência energética e responsabilidade ambiental no desenvolvimento de modelos de IA.

No que diz respeito às exigências computacionais, destaca-se que as CNNs, amplamente utilizadas em aplicações médicas, ainda enfrentam limitações quanto à sua implementação em ambientes IoT, embarcados e móveis. O método proposto, por demandar recursos significativamente menores, mostra-se mais adequado para tais cenários.

As dificuldades relacionadas ao acesso a bancos de dados reforçam a necessidade de iniciativas institucionais que promovam maior colaboração entre hospitais, pesquisadores e órgãos públicos. Parcerias, como um possível acordo com o SUS para acesso estruturado a imagens radiológicas, podem mitigar as limitações enfrentadas por pesquisas que dependem de dados médicos.

No âmbito do debate da IA de Confiança, a pesquisa também evidenciou a persistência de equívocos conceituais que influenciam debates públicos e processos regulatórios. Tais constatações ressaltam a importância de um diálogo interdisciplinar envolvendo especialistas, instituições, sociedade civil e formuladores de políticas, de modo a estabelecer normas fundamentadas e coerentes, bem como critérios sólidos para o desenvolvimento responsável e seguro de sistemas de IA.

Outra frente de pesquisa relevante consiste na avaliação clínica do método, especialmente por profissionais da área médica. A prática hospitalar demanda alto nível de precisão, transparência e rastreabilidade, e a validação por médicos permitirá verificar a aderência do modelo aos requisitos exigidos em ambientes reais.

Durante a investigação sobre IA Sustentável, observou-se que a busca por melhores resultados pode, em alguns casos, levar a decisões equivocadas ou interpretações excessivamente simplificadas, como a redução injustificada no número de exames realizados. Assim, reforça-se a importância de estabelecer limites éticos e critérios claros que conciliem sustentabilidade, qualidade diagnóstica e responsabilidade social.

Por fim, identificou-se que uma análise mais aprofundada da infraestrutura computacional pode contribuir significativamente para estimar de forma precisa a pegada de carbono dos métodos avaliados. Monitoramentos contínuos de CPU, GPU, memória e consumo energético durante o treinamento podem esclarecer variações observadas nos testes preliminares e revelar gargalos relevantes. Investigações adicionais sobre o impacto do uso de CUDA e de GPU também constituem caminhos promissores.

Em suma, como perspectivas futuras, espera-se aprimorar o método proposto, principalmente o método híbrido com Resnet, pois ele demonstrou uma pegada de carbono menor, e resultados melhores que as outras propostas, explorar estratégias otimizadas de pré-processamento e avaliar sua implementação prática em hardware. Tais avanços poderão consolidar o modelo como uma alternativa de baixo custo, transparente, reproduzível e energeticamente eficiente para o apoio ao diagnóstico de tuberculose em larga escala.

REFERÊNCIAS

- ACHARYA, V. *et al.* AI-Assisted tuberculosis detection and classification from chest X-Rays using a deep learning Normalization-Free network model. **Comput Intell Neurosci**, United States, v. 2022, p. 2399428, out. 2022.
- AHMED, F. Repvit-cxr: A channel replication strategy for vision transformers in chest x-ray tuberculosis and pneumonia classification. **ArXiv**, abs/2509.08234, 2025. Disponível em: <https://api.semanticscholar.org/CorpusID:281243782>.
- ALZUBAIDI, L. *et al.* Review of deep learning: concepts, cnn architectures, challenges, applications, future directions. **Journal of Big Data**, v. 8, 03 2021.
- BAIER, A. Trust and antitrust. **Ethics**, University of Chicago Press, v. 96, n. 2, p. 231–260, 1986. ISSN 00141704, 1539297X. Disponível em: <http://www.jstor.org/stable/2381376>.
- BISHOP, C. M. **Pattern Recognition and Machine Learning**. New York: Springer, 2006. ISBN 978-0387310732.
- BISTA, R. *et al.* Advancing tuberculosis detection in chest x-rays: A yolov7-based approach. **Information**, v. 14, n. 12, 2023. ISSN 2078-2489. Disponível em: <https://www.mdpi.com/2078-2489/14/12/655>.
- BORGES, F. E. de M. **One Class Classifier based on Principal Curves in Python**. 2024. Acesso em: 24 de Setembro de 2024. Disponível em: <https://pypi.org/project/ocpc-py/>.
- Brasil. Ministério da Saúde. Secretaria de Vigilância em Saúde. Departamento de Vigilância das Doenças Transmissíveis. **Manual de Recomendações para o Controle da Tuberculose no Brasil 2ª edição**. 2. ed. Brasília: Ministério da Saúde, 2019. Publicação técnica nacional. ISBN 978-85-334-2696-2.
- CANTRELL, S. **Top 3 Most Popular Neural Networks**. 2018. Acesso em: 20 de Setembro de 2022. Disponível em: <https://www.excella.com/insights/top-3-most-popular-neural-networks>.
- CARVALHO, J.; LOURENÇO, A. *et al.* Climate change, carbon dioxide emissions, and medical imaging contribution. **European Radiology**, Springer, 2023.
- Catholic News Service. **Pope Leo: A.I. inspired name choice**. 2025. https://www.youtube.com/watch?v=dz2n_ci3oyc. Vídeo no YouTube. Acesso em: 12 de maio 2025.
- CDC. **Centers for Disease Control and Prevention**. 2024. Acesso em: 20 de Setembro de 2024. Disponível em: <https://www.cdc.gov/tb/hcp/clinical-guidance/index.html>.
- CDC. **Centers for Disease Control and Prevention**. 2024. Acesso em: 20 de Setembro de 2024. Disponível em: <https://www.cdc.gov/tb/testing/index.html>.
- CHAPAGAIN, A.; YAQOOB, M. M. Tuberculosis in the 21st century: A narrative review. **Néphrologie & Thérapeutique**, v. 15, p. S33–S35, 2019. ISSN 1769-7255. Disponível em: <https://www.sciencedirect.com/science/article/pii/S176972551930015X>.
- CORNISH, L. **Interactive: Who's funding the COVID-19 response and what are priorities**. 2021. Acesso em: 20 de Setembro de 2022. Disponível em: <https://tinyurl.com/34ex556k>.

CORTEZ, A. O. *et al.* Tuberculosis in brazil: one country, multiple realities. **Jornal brasileiro de pneumologia : publicacao oficial da Sociedade Brasileira de Pneumologia e Tisiologia**, v. 47, n. 2, 2021. ISSN 1806-3713. Disponível em: <https://europepmc.org/articles/PMC8332839>.

COURTY, B. *et al.* **CodeCarbon: Track Carbon Emissions from Machine Learning Workloads (v2.4.1)**. Zenodo, 2024. Acesso em: 20 de Setembro de 2025. Disponível em: <https://doi.org/10.5281/zenodo.11171501>.

DEY, S. *et al.* An optimized fuzzy ensemble of convolutional neural networks for detecting tuberculosis from chest x-ray images. **Applied Soft Computing**, v. 114, p. 108094, 2022. ISSN 1568-4946. Disponível em: <https://www.sciencedirect.com/science/article/pii/S1568494621009881>.

DOSOVITSKIY, A. *et al.* An image is worth 16x16 words: Transformers for image recognition at scale. **arXiv preprint arXiv:2010.11929**, 2020.

EL-GHANY, S. A. *et al.* A robust tuberculosis diagnosis using chest x-rays based on a hybrid vision transformer and principal component analysis. **Diagnostics**, v. 14, n. 23, 2024. ISSN 2075-4418. Disponível em: <https://www.mdpi.com/2075-4418/14/23/2736>.

FALQUETO, L.; PAES, R.; PASSARO, A. Knn e rede neural convolucional para o reconhecimento de plataformas de petróleo em imagens sar do sentinel-1. **Aplicações Operacionais em Áreas de Defesa**, v. 24, p. 29–33, 09 2023.

FERNANDEZ, H. L. **Classificação de Navios Baseada em Curvas Principais**. 12–47 p. Dissertação (Mestrado) — Engenharia de Sistemas e Automação - Universidade Federal do Rio de Janeiro, 2005.

FERREIRA, D. D. *et al.* A new power quality deviation index based on principal curves. **Electric Power Systems Research**, v. 125, p. 8–14, 2015.

FURTADO, A. *et al.* A light deep learning algorithm for ct diagnosis of covid-19 pneumonia. **Diagnostics (Basel, Switzerland)**, v. 12, 06 2022.

GERIC, C. *et al.* The rise of artificial intelligence reading of chest x-rays for enhanced tb diagnosis and elimination. **The International Journal of Tuberculosis and Lung Disease**, v. 27, n. 5, p. 367–372, 2023. ISSN 1027-3719. Disponível em: <https://www.ingentaconnect.com/content/iuatld/ijtld/2023/00000027/00000005/art00005>.

GÜNAY, D. **Principal Component Analysis (PCA)**. 2023. <https://medium.com/@denizgunay/principal-component-analysis-pca-d8edf2bb6620>. Medium article, 19 min read. Disponível em: <https://medium.com/@denizgunay/principal-component-analysis-pca-d8edf2bb6620>.

HE, K. *et al.* Deep residual learning for image recognition. In: **Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)**. [S.l.: s.n.], 2016. p. 770–778.

HEALTHMANAGEMENT.ORG. **Sustainable Medical Imaging: Carbon Footprint of CT Scans and Value-Based Healthcare**. 2023. <https://healthmanagement.org>. Acesso em: 2025-02-21.

High-Level Expert Group on Artificial Intelligence, European Commission. **Ethics Guidelines for Trustworthy Artificial Intelligence**. [S.l.], 2019. Published 08 April 2019. Disponível em: <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>.

HUANG, G. *et al.* Densely connected convolutional networks. In: **Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)**. [S.l.: s.n.], 2017. p. 4700–4708.

HUY, V. T. Q.; LIN, C.-M. An improved densenet deep neural network model for tuberculosis detection using chest x-ray images. **IEEE Access**, v. 11, p. 42839–42849, 2023.

HYVÄRINEN, A.; KARHUNEN, J.; OJA, E. **Independent Component Analysis**. Wiley, 2001. (Adaptive and Cognitive Dynamic Systems: Signal Processing, Learning, Communications and Control). ISBN 9780471405405. Disponível em: <https://books.google.com.br/books?id=8EbJkwEACAAJ>.

KETZER, P. **O conceito de confiança em epistemologia do testemunho : distinguindo confiar de fiar-se**. Tese (Doutorado) — PUCRS, 2015. Faculdade de Filosofia e Ciências Humanas. Disponível em: <http://tede2.pucrs.br/tede2/handle/tede/6382>.

KOCAK, M. *et al.* Radiology ai and sustainability paradox: environmental, economic, and social dimensions. **Insights into Imaging**, Springer, v. 16, n. 1, p. 19, 2025.

KOTEL, E.; THIRUNAVUKARASU, R. A comprehensive review on advancement in deep learning techniques for automatic detection of tuberculosis from chest x-ray images. **Archives of Computational Methods in Engineering**, v. 31, p. 455–474, 2023. Disponível em: <https://api.semanticscholar.org/CorpusID:261321911>.

LIU, W. *et al.* Ssd: Single shot multibox detector. In: **Lecture Notes in Computer Science**. Springer International Publishing, 2016. p. 21–37. ISBN 9783319464480. ISSN 1611-3349. Disponível em: http://dx.doi.org/10.1007/978-3-319-46448-0_2.

Liu, Y. *et al.* Revisiting Computer-Aided Tuberculosis Diagnosis. **arXiv e-prints**, jul 2023.

MACIEL, E. L. *et al.* Tuberculosis: a deadly and neglected disease in the covid-19 era. **Jornal brasileiro de pneumologia : publicacao oficial da Sociedade Brasileira de Pneumologia e Tisiologia**, v. 48, n. 3, p. e20220056, June 2022. ISSN 1806-3713. Disponível em: <https://doi.org/10.36416/1806-3756/e20220056>.

MARIANO, D. **SVD – Decomposição por Valores Singulares com Python**. 2023. Acessado em Dia Mês Ano. Disponível em: <https://diegomariano.com/svd-decomposicao-por-valores-singulares-com-python/>.

MORAES, E. C. C. *et al.* Data clustering based on principal curves. **Advances in Data Analysis and Classification**, v. 14, n. 1, p. 77–96, Mar 2020. ISSN 1862-5355. Disponível em: <https://doi.org/10.1007/s11634-019-00363-w>.

MORÍS, D. e. a. Adapted generative latent diffusion models for accurate pathological analysis in chest x-ray images. **Medical & Biological Engineering & Computing**, 2024.

MUSINGUZI, D.; KATUMBA, A.; MURINDANYI, S. **PaliGemma-CXR: A Multi-task Multimodal Model for TB Chest X-ray Interpretation**. 2025. Acesso em: 20 de Setembro de 2025. Disponível em: <https://arxiv.org/abs/2503.00171>.

- OLIVEIRA, P. C. de. **Baldeação ideológica inadvertida e Diálogo**. [S.l.]: Catolicismo, 1965.
- OMS. **Global Tuberculosis report 2024**. 2024. Acesso em: 22 de Setembro de 2024.
- PACHECO, R. 2023. Acesso em: 20 de Setembro de 2024. Disponível em: <https://www25.senado.leg.br/web/atividade/materias/-/materia/157233>.
- PAI TEREZA KASAEVA, S. S. M. Covid-19's devastating effect on tuberculosis care — a path to recovery. **The New England Journal of Medicine**, 01 2022.
- PATTERSON, D. *et al.* Carbon emissions and large neural network training. **arXiv preprint arXiv:2104.10350**, 2021. Disponível em: <https://arxiv.org/abs/2104.10350>.
- PEARSON, K. On lines and planes of closest fit to systems of points in space. **Philosophical Magazine**, v. 2, n. 11, p. 559–572, 1901.
- PENG, T. *et al.* Detection of lung contour with closed principal curve and machine learning. **Journal of Digital Imaging**, v. 31, 02 2018.
- PETERSON TERENCE TOLAND, G. H. E. R. **Robots vs. COVID-19: how the pandemic is accelerating automation**. 2020. <https://www.kearney.com/web/global-business-policy-council/article/-/insights/robots-vs-covid-19-how-the-pandemic-is-accelerating-automation>.
- PHAM, V.-T.-N.; NGUYEN, Q.-C.; NGUYEN, Q.-V. Chest x-rays abnormalities localization and classification using an ensemble framework of deep convolutional neural networks. **Vietnam Journal of Computer Science**, v. 10, n. 01, p. 55–73, 2023. Disponível em: <https://doi.org/10.1142/S2196888822500348>.
- PRAMANA, A. A. G. Y. *et al.* **Few-Shot Learning Approach on Tuberculosis Classification Based on Chest X-Ray Images**. 2024. Acesso em: 20 de Outubro de 2024. Disponível em: <https://arxiv.org/abs/2409.11644>.
- RONNEBERGER, O.; FISCHER, P.; BROX, T. U-net: Convolutional networks for biomedical image segmentation. In: **Medical Image Computing and Computer-Assisted Intervention (MICCAI)**. [S.l.: s.n.], 2015. p. 234–241.
- RYAN, M. In ai we trust: Ethics, artificial intelligence, and reliability. **Science and Engineering Ethics**, v. 26, n. 5, p. 2749–2767, Oct 2020. ISSN 1471-5546. Disponível em: <https://doi.org/10.1007/s11948-020-00228-y>.
- SAÚDE, M. D. **Tuberculose**. 2024. Acesso em: 22 de Setembro de 2024. Disponível em: <https://www.gov.br/saude/pt-br/assuntos/saude-de-a-a-z/t/tuberculose>.
- SCHNEIDER, I. *et al.* **Life-Cycle Emissions of AI Hardware: A Cradle-To-Grave Approach and Generational Trends**. 2025. Acesso em: 20 de Setembro de 2025. Disponível em: <https://arxiv.org/abs/2502.01671>.
- SCHWARTZ, R. *et al.* Green ai. **Communications of the ACM**, v. 63, n. 12, p. 54–63, 2020.
- SHLENS, J. **A Tutorial on Principal Component Analysis**. [S.l.], 2014. ArXiv:1404.1100. Disponível em: <https://arxiv.org/abs/1404.1100>.

SIMONYAN, K.; ZISSERMAN, A. Very deep convolutional networks for large-scale image recognition. **arXiv preprint arXiv:1409.1556**, 2014.

STEGER, A. **Healthcare Automation Matters More Than Ever During a Pandemic**. 2020. <https://healthtechmagazine.net/article/2020/07/healthcare-automation-matters-more-ever-during-pandemic-perfcon>.

SZEGHALMY, S.; FAZEKAS, A. A comparative study of the use of stratified cross-validation and distribution-balanced stratified cross-validation in imbalanced learning. **Sensors**, v. 23, n. 4, 2023. ISSN 1424-8220. Disponível em: <https://www.mdpi.com/1424-8220/23/4/2333>.

TRAJMAN, A. *et al.* Tuberculosis. **The Lancet**, Elsevier, v. 405, n. 10481, p. 850–866, Mar 2025. ISSN 0140-6736. Disponível em: [https://doi.org/10.1016/S0140-6736\(24\)02479-6](https://doi.org/10.1016/S0140-6736(24)02479-6).

TREVOR, H.; WERNER, S. Principal curves. **Journal of the American Statistical Association**, Taylor & Francis, v. 84, n. 406, p. 502–516, 1989. Disponível em: <https://www.tandfonline.com/doi/abs/10.1080/01621459.1989.10478797>.

VERBEEK, J.; VLASSIS, N.; KRÖSE, B. A k-segments algorithm for finding principal curves. **Pattern Recognition Letters**, v. 23, n. 8, p. 1009–1017, 2002. ISSN 0167-8655. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0167865502000326>.

WATSON, I. **O Watson Health é a saúde mais inteligente**. 2022. Acesso em: 20 de Setembro de 2022. Disponível em: <https://www.ibm.com/br-pt/watson-health>.

ZAMAN, K. Tuberculosis: a global health problem. **Journal of health, population, and nutrition**, v. 28, n. 2, p. 111–113, April 2010. ISSN 1606-0997. Disponível em: <https://europepmc.org/articles/PMC2980871>.