

Radiomics for the Classification of Radiological Patterns in Computed Tomography

Uso da Radiômica para Classificação de Padrões Radiológicos em Tomografia Computadorizada

Heloísa Dias Viotto^{1*}, Felipe Duarte Silva¹, Guilherme Eduardo Gonçalves da Silva¹, Heloisa de Oliveira Pinho¹, Ricardo Quer do Nascimento Filho¹, Lucas Ferrari de Oliveira¹

Resumo: Pulmonary diseases are a major issue today, being one of the leading causes of death worldwide. Because of this, it is essential to develop methods to improve their diagnosis. This study arises from this problem, with the aim of identifying radiomic features in Computed Tomography (CT) and to evaluate the performance of different machine learning algorithms, including KNN, SVM, RF, and MLP. The results obtained show satisfactory values, with specificity standing out. SVM had the best performance with, with a mean specificity of 96.25% and a mean sensitivity of 85.05% both with a small standard deviation.

Keywords: radiomic — radiological patterns — machine learning

Resumo: As doenças pulmonares são um grande problema na atualidade, sendo uma das principais causas de morte no mundo. Devido a isso, é essencial que se desenvolva formas para melhorar o seu diagnóstico. Este estudo tem a finalidade de identificar características radiômicas em imagens de Tomografia Computadorizada (TC) e avaliar o desempenho de diferentes algoritmos de aprendizado de máquina, incluindo KNN, SVM, RF e MLP. Os resultados alcançados possuem valores satisfatórios, com destaque para a especificidade. O SVM teve o melhor resultado com uma especificidade média de 96,25% e uma sensibilidade média de 85,05%, ambos com valores baixos de desvio padrão.

Palavras-Chave: radiômica — padrões radiológicos — aprendizado de máquina

¹ Departamento de Informática, Universidade Federal do Paraná (UFPR), Brasil

*Corresponding author: hdv23@inf.ufpr.br

DOI: <http://dx.doi.org/10.22456/2175-2745.150929> • Received: 14/10/2025 • Accepted: 03/11/2025

CC BY-NC-ND 4.0 - This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.

1. Introdução

Segundo estudos da Organização Mundial da Saúde (OMS), cerca de 1,8 milhão de pessoas morrem por ano devido ao câncer de pulmão, sendo que esta e as doenças pulmonares em geral estão entre as principais causas de morte no mundo [1, 2]. Em 2021, dados globais indicaram que cinco das dez principais causas de óbito estavam relacionadas a problemas pulmonares [3].

No Brasil, o câncer de traqueia, brônquios e pulmões foi responsável pela morte de aproximadamente 28 mil pessoas no ano de 2021, de acordo com o Instituto Nacional de Câncer (INCA) [4]. Isso mostra que estudos nas áreas de diagnóstico e tratamento de doenças pulmonares são essenciais para a atualidade, com impacto em milhares de vidas.

Os métodos de detecção de doenças pulmonares incluem Tomografia por Emissão de Pósitrons (PET), Imagem por Ressonância Magnética (RM), Tomografia Computadorizada

(TC) e radiografias de tórax. Dentre esses, a TC destaca-se devido à sua maior nitidez e por ser livre de distorções em relação às demais [5]. Com os avanços tecnológicos, tornou-se possível converter essas imagens para o formato digital, permitindo o desenvolvimento de ferramentas computacionais que aceleram e tornam mais precisa a identificação de padrões patológicos [2, 6, 7].

Nesse contexto, os sistemas de diagnóstico auxiliado por computador (*Computer-Aided Diagnosis* – CAD) têm se mostrado promissores ao oferecer uma segunda opinião ao radiologista, reduzindo incertezas e apoiando a tomada de decisão clínica. Essas ferramentas funcionam como suporte ao especialista, sendo o diagnóstico final sempre de responsabilidade médica [8].

Diante desses avanços, pesquisadores exploram maneiras de acelerar o processo de identificação de anomalias de forma rápida e precisa [7, 9]. Para continuar o desenvolvimento de sistemas CAD, técnicas como processamento de imagens,

radiômica, redes neurais (*Deep Learning*) e algoritmos de classificação são amplamente utilizados [10].

Dentre essas técnicas, a radiômica desempenha papel central, pois possibilita a extração de características quantitativas presentes em imagens médicas (TC, RM, PET) para avaliar o fenótipo do tumor. Tais características como textura, formato e cor, auxiliam tanto sobre o diagnóstico como o prognóstico em vários tipos de doenças [11, 12]. As etapas para a extração incluem: aquisição de dados, segmentação de imagens, extração de dados e análise/modelagem. Todas as etapas são implementadas por bibliotecas de programação já consagradas, como por exemplo, a *PyRadiomics* [13, 14].

Após a extração, os dados gerados são utilizados por classificadores com o objetivo de categorizar os padrões pulmonares mais comuns em pacientes com doenças pulmonares difusas, tais como: *não pulmão*, *normal*, *enfisema*, *vidro fosco*, *fibrose* e *micronódulos*.

Entre as abordagens para a classificação de imagens, a mais utilizada é a rede neural convolucional (*Convolutional Neural Network* – CNN), pois apresenta um desempenho eficaz na identificação de padrões complexos das imagens médicas, inclusive no diagnóstico de doenças pulmonares [15]. Além disso, a CNN contribui para um aumento de sensibilidade na detecção do câncer de pulmão e auxílio na redução de falsos positivos [16, 17]. No entanto, há outros algoritmos comumente utilizados para métodos de classificação, como por exemplo o *Support Vector Machine* (SVM), *K-Nearest Neighbour* (KNN) e *Random Forest* (RF), além do próprio MLP (*Multi Layer Perceptron*).

Diferentemente da maioria dos trabalhos encontrados na literatura, este estudo propõe uma abordagem que combina a extração radiômica com validação cruzada por paciente, utilizando o método *StratifiedGroupKFold* da biblioteca *scikit-learn*, o qual é responsável por evitar que imagens de um mesmo paciente apareçam simultaneamente nos conjuntos de treino e teste. Além disso, é apresentada uma comparação sistemática entre quatro algoritmos supervisionados — *1-NN*, RF, SVM e MLP.

O experimento foi conduzido sobre uma base pública com cerca de 77 mil imagens de TC. As imagens foram segmentadas e processadas com a biblioteca *PyRadiomics*, resultando em descritores quantitativos utilizados para treinar e avaliar os classificadores. O desempenho foi avaliado por meio das métricas de sensibilidade, especificidade e F1-score do modelo, permitindo avaliar o desempenho de cada algoritmo na detecção de anomalias pulmonares.

A Seção 2 introduz conceitos utilizados na elaboração dos experimentos. Já a Seção 3 apresenta o que existe de trabalhos relacionados ao proposto. As Seções 4 e 5 mostram os recursos utilizados para a construção deste trabalho, bem como os resultados alcançados. Por fim, a Seção 6 conclui os achados deste trabalho.

2. Fundamentação Teórica

Esta seção visa introduzir brevemente os principais conceitos envolvidos neste estudo. A seção 2.1 aborda os fundamentos da radiômica, destacando suas etapas. Em seguida, a seção 2.2 apresenta os algoritmos de comparação utilizados. Já a Seção 2.3 descreve as classes do *dataset* utilizado.

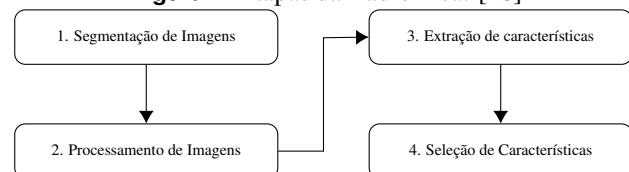
2.1 Radiômica

Com a abundante quantidade de dados produzidos na atualidade e o avanço da Inteligência Artificial (IA), procuram-se formas de interpretar o que é produzido. Nesse contexto, surge a radiômica, a qual busca extrair dados quantitativos de imagens médicas dificilmente percebidas pelos humanos utilizando noções matemáticas, como, por exemplo, em relação à distribuição espacial e à intensidade dos *pixels*. As características extraídas apresentam relação com a forma, cores, intensidade e textura da imagem analisada [18, 19].

As características radiômicas têm potencial para prever resposta de tratamento e estão indiretamente ligadas às ciências ômicas (genômica, transcriptômica e proteômica), além de poder ser utilizada com outros dados, como os clínicos e os de laboratórios, e usá-los em modelos de aprendizado de máquina. É importante ressaltar que a radiômica é um recurso que pode ser aplicada para auxiliar o diagnóstico, mas não como substituto de um profissional [19].

A radiômica pode ser definida por 4 passos como pode ser visto no diagrama apresentado na Figura 1.

Figura 1. Etapas da Radiômica. [18]



2.1.1 Segmentação de Imagens

Nesse processo, define-se a área de interesse (*Region of Interest (ROI)*) ou o volume de interesse (*Volume of Interest - VOI*) para imagens tridimensionais. Essa etapa é a definição de qual é a área em que as características serão extraídas, sendo essencial para a qualidade do resultado final [18].

2.1.2 Processamento de Imagens

A segunda etapa está relacionada a padronização da imagem em relação à intensidade dos níveis de cinza, espaçamento entre *pixels* e compartimentos de histograma, por exemplo. Também pode-se citar a interpolação, a re-segmentação, a discretização de intensidade e a extração do ROI. No caso desse estudo, isso foi realizado com a criação de um arquivo YAML utilizado no extrator de características da biblioteca *PyRadiomics*. [18]

2.1.3 Extração de Características

Nessa etapa várias técnicas são utilizadas para extrair características radiômicas. A segmentação feita anteriormente é uti-

lizada para calcular as características de forma (volume, área de superfície, esfericidade), além disso características de textura 2D e 3D também são extraídas. As características de textura descrevem padrões e heterogeneidade e as mais utilizadas são: contraste, entropia e homogeneidade da GLCM (*Grey Level Co-occurrence Matrix*) e características da GLRLM (*Grey Level Run Length Matrix*). Essas características serão melhor explicadas na Seção 4.3.2.

2.1.4 Seleção de Características

Nesta etapa as características mais relevantes são selecionadas de acordo com o problema em questão [18].

2.2 Algoritmos de Classificação

2.2.1 K-Nearest-Neighbours (KNN)

O algoritmo do K vizinhos mais próximos é um algoritmo de aprendizado de máquina supervisionado que utiliza o pressuposto de que itens iguais ocupam um espaço semelhante em um plano de N dimensões. Portanto, dado um item desconhecido (no caso deste estudo, uma imagem de TC), o KNN irá definir a sua classe de acordo com a classificação dos K itens mais próximos no plano de dimensão N cujos os dados de treinamento utilizados a priori estão representados [20, 21].

2.2.2 Random Forest (RF)

O Random Forest é um algoritmo que combina o resultado de N árvores de decisão para gerar um único. Ao contrário das árvores de decisão que utilizam todo o conjunto de dados de entrada, o random forest seleciona algumas características para ser utilizado [22].

2.2.3 Support Vector Machine (SVM)

O *Supported Vector Machine* é um algoritmo de aprendizado de máquina supervisionado comumente utilizado em problemas de classificação linear e não linear. O algoritmo busca encontrar o melhor hiperplano que permite separar duas classes por meio de margens que estão em ambos os seus lados. Essas margens são limitadas por pontos, denominados de vetores de suporte. Sendo assim, o objetivo é encontrar a maior margem entre os pontos mais próximo das classes, sendo o hiperplano o separador ideal entre as margens. Tratando-se de um cenário multiclasse, o algoritmo utiliza hiperplanos com N dimensões, o qual são definidas com base em funções matemáticas, denominadas *Kernel SVM* [23, 24].

2.2.4 Multi Layer Perceptron (MLP)

O *Multi Layer Perceptron* - MLP é uma rede neural composta por múltiplos neurônios, os quais formam três tipos de camadas: entrada, saída e ocultas. A primeira camada, a qual os dados são introduzidos, é a de entrada. Os dados passam em um único sentido para as camadas ocultas, que processam os dados e, por fim, uma resposta é devolvida pelos neurônios que formam a camada de saída [20, 21].

2.3 Doenças Pulmonares

Esta Subseção apresenta o conceito de quatro achados radiológicos que são classes do nosso estudo, são elas enfisema,

vidro fosco, fibrose e micronódulos. Há ainda outras duas classes, o não pulmão e o pulmão saudável, que são capazes de serem definidas como um tecido qualquer do corpo humano capturado pela tomografia computadorizada e um tecido pulmonar saudável, respectivamente.

2.3.1 Enfisema

O enfisema pulmonar é uma condição em que há um aumento fora do normal do espaços aéreos após os bronquíolos terminais. O tecido pulmonar responsável pelas trocas gasosas é prejudicado permanentemente e sem formação de cicatrizes [25].

2.3.2 Vidro Fosco

O vidro fosco é um padrão visto em tomografias computadorizadas em que é possível ver aumento na densidade dos pulmões, no entanto, ainda é possível visualizar os vasos sanguíneos e as vias aéreas [26].

2.3.3 Fibrose

A fibrose pulmonar ocorre quando há uma redução na funcionalidade do pulmão devido ao surgimento de fibrose (tecido cicatricial) no tecido que envolve os alvéolos pulmonares, órgão responsável pelo troca gasosa [27].

2.3.4 Micronódulos

Os micronódulos pulmonares são formações de diâmetro médio de até 6 milímetros que podem ser encontrados separadamente ou agrupados. Eles possuem características similares aos de nódulos, mas são mais difíceis de serem detectados devido a sua menor estrutura [28].

3. Trabalhos Correlatos

O diagnóstico das doenças pulmonares intersticiais (*Interstitial Lung Diseases* — ILDs) representa um desafio recorrente devido à sobreposição entre padrões radiológicos e à dependência de métodos especializados como a tomografia computadorizada de alta resolução (*High-Resolution Computed Tomography* — HRCT). Embora eficaz, a HRCT requer interpretação multidisciplinar e, por vezes, validação com métodos invasivos. Nesse cenário, abordagens baseadas em inteligência artificial vêm ganhando espaço por seu potencial em reduzir a variabilidade interobservador e ampliar o suporte diagnóstico.

Estudos recentes concentram-se majoritariamente em quatro frentes: (i) radiômica com aprendizado supervisionado tradicional, (ii) redes neurais convolucionais aplicadas a imagens HRCT, (iii) integração multimodal de dados clínicos e de imagem, e (iv) modelos aplicados ao câncer pulmonar. Apesar dos avanços, persiste uma lacuna no uso de imagens radiográficas convencionais — mais acessíveis na prática clínica — e na avaliação comparativa de diferentes algoritmos supervisionados sob esquemas de validação que considerem a dependência entre amostras.

Este trabalho contribui ao utilizar uma base pública composta por radiografias torácicas anotadas, adotando validação

cruzada estratificada por paciente e comparando o desempenho de quatro algoritmos aplicados a atributos radiômicos extraídos automaticamente.

3.1 Modelos baseados em radiômica e aprendizado supervisionado

He et al. [29] propuseram um modelo preditivo de progressão acelerada de ILD em pacientes com dermatomiosite anti-MDA5 positiva. A partir de atributos extraídos com a biblioteca *PyRadiomics*, treinaram um modelo de regressão logística sobre imagens HRCT, atingindo uma área sob a curva (AUC) superior a 0,90 em validações externas.

Haga et al. [30] investigaram a relação entre texturas radiômicas de HRCT e infiltração celular em biópsias pulmonares. O modelo de classificação binária alcançou *F1-score* de 0,73, sugerindo que as características radiômicas podem refletir alterações histopatológicas sem necessidade de procedimentos invasivos.

Bernardi et al. [31] utilizaram atributos radiômicos combinados com o algoritmo RF para distinguir padrões de usual interstitial pneumonia (UIP) de outras ILDs, com AUC de 0,96. Gil et al. (2024), por sua vez, aplicaram *AdaBoost* para prever toxicidade pulmonar em pacientes com ILD submetidos à radioterapia estereotáxica (SABR), destacando a aplicabilidade clínica das abordagens supervisionadas clássicas.

3.2 Modelos baseados em redes neurais convolucionais (CNNs)

Yoon et al. [32] propuseram um modelo baseado em *CNN* para subtipagem automática de ILDs com imagens HRCT. O desempenho do modelo foi comparável ao de especialistas humanos, considerando métricas de acurácia e sensibilidade clínica. Em experimentos reportados, o modelo alcançou uma AUC de 0,945 na fase de treinamento e 0,845 em teste, demonstrando alta capacidade discriminativa.

Kim et al. [32] avaliaram a robustez de modelos *CNN* treinados para detecção do padrão de UIP, utilizando dados de diversas instituições. A validação multi-institucional evidenciou boa capacidade de generalização entre domínios clínicos distintos, reforçando o potencial de escalabilidade dos modelos baseados em *deep learning*.

3.3 Modelos Multimodais

Liu et al. [33] apresentaram um modelo *multimodal* que integra características radiômicas com variáveis clínicas estruturadas (como idade, sintomas e biomarcadores laboratoriais) para prever o risco de desenvolvimento de ILD em pacientes pós-COVID-19. A fusão dos dois domínios de dados elevou a acurácia e a especificidade do modelo. Estudos semelhantes com abordagens multimodais relataram AUCs em torno de 0,84 na fase de teste, sugerindo bom desempenho clínico potencial.

3.4 Modelos aplicados a outras doenças pulmonares

Embora com foco distinto, Rohmah et al. [34] compararam diferentes classificadores — SVM, RF, *Decision Tree* e *Naïve Bayes* — na detecção de câncer de pulmão com base em atributos extraídos de imagens radiográficas. O modelo SVM apresentou melhor desempenho, com acurácia de 96,5%, reforçando a aplicabilidade de métodos supervisionados em imagens torácicas bidimensionais.

3.5 Classificação de Tecidos Pulmonares

Bagesteiro et al. [35] propuseram um estudo que fez a classificação de tecidos pulmonares em imagens de tomografia computadorizada. Utilizando a técnica de CLBP, para a extração de atributos, os autores utilizaram *Support Vector Machine* (SVM). O modelo SVM com o CLBP apresentou o desempenho de sensibilidade de 0,7899, especificidade de 0,9300 na AUC de 0,9340.

A Tabela 1 resume os principais aspectos dos estudos mencionados nesta seção.

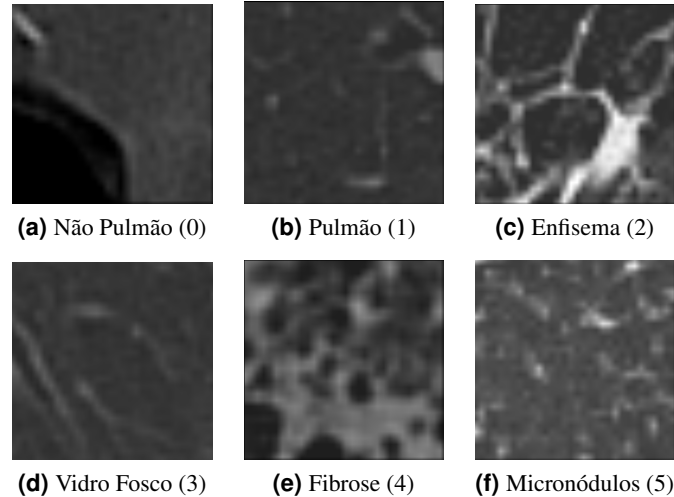
Tabela 1. Resumo de estudos em radiômica aplicada à ILD.

Autor(es)	Foco	Método
He et al. (2024)	ILD anti-MDA5+	<i>PyRadiomics</i> + regressão
Haga et al. (2024)	Infiltração celular	<i>PyRadiomics</i> + classificação binária
Yoon et al. (2023)	Subtipagem ILD	<i>CNN</i> + HRCT
Kim et al. (2023)	Detecção de UIP	<i>CNN</i> + validação multi-institucional
Bernardi et al. (2022)	UIP vs. outras ILDs	<i>Random Forest</i>
Gil et al. (2024)	Toxicidade em SABR-ILD	<i>AdaBoost</i>
Liu et al. (2023)	ILD pós-COVID-19	Fusão de dados clínicos + radiômica
Rohmah et al. (2024)	Câncer de pulmão	<i>SVM</i>
Bagesteiro et al. (2024)	Câncer de pulmão (CT)	<i>CLBP</i> + <i>SVM</i>

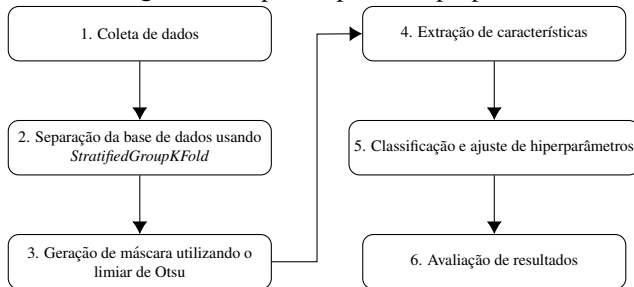
4. Materiais e Métodos

A Figura 3 apresenta o fluxo de execução dos experimentos, composto por seis etapas principais:

- coleta de dados;
- separação da base de dados por *StratifiedGroupKFold*;
- segmentação das imagens por meio de máscara gerada com o limiar de Otsu;

Figura 2. Imagens das classes utilizadas.

- extração de características radiômicas com a biblioteca *PyRadiomics*;
- classificação utilizando 11-Nearest Neighbors (11-NN), RF, SVM e MLP, com ajuste de hiperparâmetros; e
- avaliação dos resultados.

Figura 3. Etapas do processo proposto.

4.1 Conjunto de dados

Foram utilizadas 77.835 imagens de regiões de interesse (ROIs) obtidas a partir de exames de tomografia computadorizada da base pública ILD Database (HUG-ILD/MedGIFT), que reúne cortes de alta resolução anotados por especialistas em diferentes padrões pulmonares [35]. As imagens estão distribuídas em sete classes clínicas: *não pulmão* (0), *normal* (1), *enfisema* (2), *vidro fosco* (3), *fibrose* (4), *micronódulos* (5) e *consolidação* (6). Esta última foi desconsiderada devido ao número reduzido de amostras ($n = 60$), o que inviabilizaria um treinamento consistente dos modelos e poderia comprometer a avaliação dos resultados. A Figura 2 apresenta exemplos das classes utilizadas.

4.2 Separação dos dados

A divisão entre conjuntos de treino e de teste foi realizada utilizando a técnica *StratifiedGroupKFold* ($k = 5$), disponível

na biblioteca *scikit-learn* [36]. Esse método combina dois critérios: (i) mantém a proporção original das classes em cada *fold* (estratificação), e (ii) garante que todas as amostras provenientes de um mesmo paciente (grupo) fiquem restritas a apenas um dos subconjuntos, evitando sobreposição entre treino e teste. Essa abordagem minimiza o risco de *data leakage*, isto é, reduz a chance de o modelo aprender informações não generalizáveis por exposição prévia a dados correlacionados.

Durante os testes iniciais, percebemos que um dos folds não possuía uma das classes, impossibilitando encontrar o AUC. Devido a isso, foi utilizado um código que todos os *folds* tivessem ao menos uma amostra de cada uma das classes, variando o argumento *random_state* da função.

4.3 Radiômica

4.3.1 Segmentação

Cada ROI foi convertida para escala de cinza e, em seguida, submetida ao método de limiarização automática de Otsu [37]. Esse procedimento define um valor de corte T^* capaz de separar o parênquima pulmonar do fundo com base na distribuição de intensidades. O limiar ótimo é obtido pela maximização da variância entre classes (σ_B^2), conforme a Equação 1, em que $\omega_i(T)$ representa a proporção de *pixels* em cada classe $i \in \{0, 1\}$, e $\mu_i(T)$ corresponde à média de intensidades de cada grupo.

$$T^* = \arg \max_T \sigma_B^2(T) \quad (1)$$

Optamos por esse processo de segmentação justamente para isolar o tecido pulmonar do fundo e evitar que regiões irrelevantes interferissem nos descritores. Diferentemente de uma máscara “totalmente branca”, que incluiria todos os *pixels* da ROI sem distinção, a aplicação do Otsu permite refinar o conteúdo analisado, preservando apenas a região de interesse.

4.3.2 Extração de características

As máscaras resultantes foram utilizadas no *feature extractor* da biblioteca *PyRadiomics* [38]. Todas as categorias de descritores documentadas foram calculadas, com exceção das de forma, que não fornecem informação útil em ROIs retangulares (por não representarem estruturas anatômicas reais).

Foram extraídas 94 características de textura, abrangendo diferentes aspectos estatísticos da imagem: *First-Order Statistics* (19), *Gray Level Co-occurrence Matrix* (GLCM) (24), *Gray Level Run Length Matrix* (GLRLM) (16), *Gray Level Size Zone Matrix* (GLSZM) (16), *Neighbouring Gray Tone Difference Matrix* (NGTDM) (5) e *Gray Level Dependence Matrix* (GLDM) (14). Esses atributos constituem a representação quantitativa das texturas pulmonares utilizada nas etapas seguintes de classificação. Porém, utilizamos somente 93 das 94 características, pois a característica *Sum Average* é o dobro da *Joint Average* (ambas da GLCM), sendo necessário utilizar apenas uma delas.

Visão geral das características radiômicas. Os descritores de primeira ordem e de textura (GLCM, GLRLM, GLSZM, NGTDM e GLDM) resumem a distribuição de intensidades e a organização espacial dos tons de cinza em múltiplas escalas, capturando contraste, homogeneidade, granularidade e variação local do parênquima. Essa representação quantitativa é suficiente para diferenciar padrões como o vidro fosco, o enfisema e a fibrose.

4.4 Classificação e ajuste de hiperparâmetros

Foram avaliados quatro algoritmos supervisionados — KNN, RF, SVM e MLP — seguindo o mesmo esquema de cinco *folds* gerados pelo *StratifiedGroupKFold*.

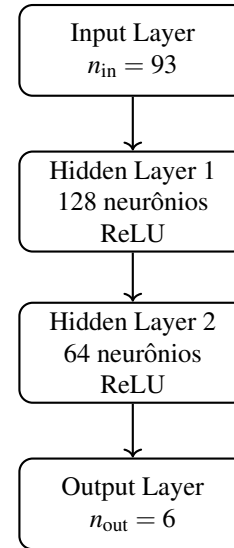
Para cada combinação de hiperparâmetros o modelo foi treinado e avaliado nos cinco *folds*:

- KNN: número de vizinhos $K = 1, 3, 5, 7, 9, 11, 13, 15, 21$.
- RF: $n_estimators = 100, 200, 300, 400$; e $max_depth = 5, 10, 20, 30, 40, 50$.
- SVM (kernel RBF): custo $C = 0,01, 0,1, 1, 10, 100, 500, 1000$ com $\gamma = scale$.

Para cada combinação treinou-se o modelo na partição de treino e avaliou-se na partição de teste do respectivo *fold*. Os F_1 médios das combinações foram comparados, e selecionou-se aquela que obteve o maior F_1 médio. Isso resultou em: $k = 11$ para o KNN; 300 árvores com profundidade máxima de 30 para a RF; para a SVM, kernel RBF com $C = 10$ e $\gamma = scale$; e, para o MLP foi utilizado 30 épocas de treinamento com *learning rate* de $1 \times e^{-3}$ juntamente com otimizador Adam e *Cross Entropy Loss*. Na Figura 4, é possível encontrar a arquitetura do MLP utilizado, o qual soma 20.678 parâmetros.

Com esses hiperparâmetros fixados, cada classificador foi novamente treinado e avaliado *fold a fold*; os valores obtidos são apresentados na Tabela 2.

Figura 4. Arquitetura do MLP



4.5 Avaliação dos resultados

As métricas sensibilidade, especificidade e AUC foram utilizadas para avaliar o desempenho dos classificadores [39]. A sensibilidade (também chamada de *recall*) quantifica a proporção de casos positivos corretamente identificados pelo modelo, sendo essencial em cenários onde falsos negativos são críticos. A especificidade mede a capacidade do classificador em reconhecer corretamente os casos negativos, reduzindo o número de falsos positivos [40]. Já a Área Sob a Curva (*Area Under Curve ROC* - AUC-ROC) representa a área total sob a curva ROC [41]. Para cada métrica, foi reportada a média aritmética e o desvio padrão dos cinco *folds*, proporcionando uma estimativa mais estável da capacidade de generalização.

5. Resultados e Discussão

A Tabela 2 apresenta os resultados alcançados. Os modelos se mostraram mais específicos que sensíveis, ou seja, o modelo gera poucos resultados falsos positivos. Por outro lado, os modelos não são tão sensíveis, deixando de indicar algumas condições reais (falsos negativos).

Tabela 2. Média e Desvio Padrão dos Resultados dos classificadores.

	Sensibilidade %	Especificidade %	AUC %
KNN	80,56±0,10	95,14±0,02	85,18±0,03
RF	84,54±0,07	96,14±0,02	89,48±0,04
SVM	85,05±0,04	96,26±0,01	92,89±0,03
MLP	77,35±0,06	95,47±0,01	92,24±0,04

Analisando os resultados individuais de cada modelo, destaca-se a SVM, com os melhores valores médios de sensibilidade, especificidade e AUC e com pouca variação entre os *folds*. Já o MLP teve o menor resultado na sensibilidade e um desvio padrão mais alto, juntamente com o KNN. No entanto, apresentou uma AUC bom, ficando atrás somente do SVM,

mostrando que em algum fold ele teve um melhor resultado, porém seu desempenho não foi consistente.

O RF teve resultados próximos ao SVM, indicando que também pode ser uma boa escolha, no entanto, seus resultados são menos constantes que o do melhor modelo. Por fim, o KNN obteve resultados intermediários nas métricas e um alto desvio padrão na sensibilidade.

Em comparação ao trabalho desenvolvido anteriormente [35], o resultado apresentado pelo SVM foi superior na sensibilidade com 85,05% contra 78,99% e na especificidade 96,25% contra 93%. A métrica de AUC foi a única que a nova pesquisa não conseguiu alcançar resultados melhores com 92,89% contra 93,40%. Porém, as duas métricas em que tivemos melhores resultados foram em torno de 3 até 6 pontos percentuais superiores.

6. Conclusão

Nesta pesquisa, avaliamos o desempenho de algoritmos de classificação (KNN, RF, SVM e MLP) em um conjunto de dados formado por recortes de imagens de tomografia computadorizada de pulmão com 6 classes distintas. Para isso, utilizamos os dados provenientes de características de textura da técnica conhecida como radiômica.

Após os testes, é possível afirmar que o melhor modelo foi o SVM, o qual exibiu os melhores resultados nas três métricas avaliadas, além de menor variabilidade, seguido pelo RF, MLP e o KNN, o qual apresentou os menores valores e a maior variabilidade entre as execuções.

Contribuição dos Autores

Heloísa Dias Viotto foi responsável pela conceituação, metodologia, software, administração do projeto e redação (rascunho original); Felipe Duarte Silva contribuiu com a curadoria de dados, metodologia e redação (rascunho original); Guilherme Eduardo Gonçalves da Silva atuou na redação (rascunho original, revisão e edição); Heloisa de Oliveira Pinho contribuiu com a redação (rascunho original); Ricardo Quer do Nascimento Filho foi responsável pela investigação e redação (rascunho original); Lucas Ferrari de Oliveira contribuiu com a conceituação, supervisão e redação (revisão e edição).

References

- [1] KEERTHI, P.; Y, S. S. S.; S, I. Lung cancer classification using deep neural networks. In: *2025 AI-Driven Smart Healthcare for Society 5.0*. [S.l.: s.n.], 2025. p. 1–6.
- [2] HAQUE, K. N. et al. Efficacy of an attention based lung cancer detection through hybrid cnn approach. In: *2024 27th International Conference on Computer and Information Technology (ICCIT)*. [S.l.: s.n.], 2024. p. 717–722.
- [3] World Health Organization. *The top 10 causes of death*. Genebra, Suíça, 2024. Fact sheet. Acesso em: 19 jun. 2025.
- Disponível em: <https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death>.
- [4] Instituto Nacional de Câncer (INCA). *Estatísticas de câncer: números*. Rio de Janeiro, Brasil, 2022. Portal Gov.br. Acesso em: 19 jun. 2025. Disponível em: <https://www.gov.br/inca/pt-br/assuntos/cancer/numeros>.
- [5] NARVEKAR, S. et al. A survey on detection of lung cancer using different image processing techniques. In: *2022 3rd International Conference on Intelligent Engineering and Management (ICIEM)*. [S.l.: s.n.], 2022. p. 13–18.
- [6] SU, H. et al. Tgmt: A terminology-guided multi-task approach for lung cancer ct report generation. In: *2024 46th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. [S.l.: s.n.], 2024. p. 1–5.
- [7] G, S. V.; PONRAJ, N.; L, D. P. Study on public chest x-ray data sets for lung disease classification. In: *2021 3rd International Conference on Signal Processing and Communication (ICPSC)*. [S.l.: s.n.], 2021. p. 54–58.
- [8] AZEVEDO-MARQUES, P. M. d. Diagnóstico auxiliado por computador na radiologia. *Radiologia Brasileira*, v. 34, p. 285–293, 2001.
- [9] PEREIRA, F. R.; OLIVEIRA, L. F. de. Desenvolvimento de um sistema cad para detecção de nódulo pulmonar em exames de tomografia computadorizada de tórax. In: *SBC. Anais do Congresso Brasileiro de Informática em Saúde (CBIS)*. Brasil, 2021.
- [10] DULF, E.-H. et al. Computer-aided diagnosis system for colorectal cancer. In: *2021 25th International Conference on System Theory, Control and Computing (ICSTCC)*. [S.l.: s.n.], 2021. p. 178–181.
- [11] AERTS, H. J. W. L. et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nature Communications*, v. 5, n. 1, p. 4006, 2014. ISSN 2041-1723. Disponível em: <https://doi.org/10.1038/ncomms5006>.
- [12] LIU, Z. et al. The applications of radiomics in precision diagnosis and treatment of oncology: Opportunities and challenges. *Theranostics*, v. 9, n. 5, p. 1303–1322, 2019. Disponível em: <https://doi.org/10.7150/thno.30309>.
- [13] SCHALL, B.; ANTY, R.; FILLATRE, L. Non-linear logistic regression applied to radiomics. In: *2024 46th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. [S.l.: s.n.], 2024. p. 1–4.
- [14] I, G. et al. Enhanced robustness in lung cancer classification from ct images using noise-resistant dynamic integral neural networks. In: *2025 International Conference on Emerging Technologies in Engineering Applications (ICETEA)*. [S.l.: s.n.], 2025. p. 1–6.

- [15] METAGAR, S. M.; SAYYED, F. R. Lung cancer detection and classification using machine learning algorithm(cnn). In: *2024 Second International Conference on Advances in Information Technology (ICAIT)*. [S.l.: s.n.], 2024. v. 1, p. 1–5.
- [16] de Margerie-Mellon, C.; CHASSAGNON, G. Artificial intelligence: A critical review of applications for lung nodule and lung cancer. *Diagnostic and Interventional Imaging*, v. 104, n. 1, p. 11–17, 2023. ISSN 2211-5684. Disponível em: <https://www.sciencedirect.com/science/article/pii/S221156842200225X>.
- [17] MIDDHA, S.; KUMAR, A. Advancements in early lung cancer detection using convolutional neural networks. In: *2023 International Conference on Artificial Intelligence for Innovations in Healthcare Industries (ICAIIHI)*. [S.l.: s.n.], 2023. v. 1, p. 1–6.
- [18] TIMMEREN, J. E. van et al. Radiomics in medical imaging - “how-to” guide and critical reflection. *Insights into Imaging*, 2020.
- [19] MAYERHOEFER, M. E. et al. Introduction to radiomics. *J Nucl Med.*, 2020.
- [20] POTGHAN, S.; RAJAMENAKSHI, R.; BHISE, A. Multi-layer perceptron based lung tumor classification. In: *2018 Second International Conference on Electronics, Communication and Aerospace Technology (ICECA)*. [S.l.: s.n.], 2018. p. 499–502.
- [21] YU, H. et al. Advancing precision in lung cancer subtyping: Integration of machine learning feature selection with mlp. In: *2024 7th International Conference on Advanced Algorithms and Control Engineering (ICAACE)*. [S.l.: s.n.], 2024. p. 514–518.
- [22] ROHMAH, A. et al. Comparison of machine learning model for lung cancer detection using classification techniques. In: *2024 International Conference of Adisutjipto on Aerospace Electrical Engineering and Informatics (ICAAEEI)*. [S.l.: s.n.], 2024. p. 1–6.
- [23] AL-OFARY, S.; ILHAN, H. O. Classification of pca based reduced deep features by svm for diagnosing lung and colon cancer. In: *2023 5th International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA)*. [S.l.: s.n.], 2023. p. 1–5.
- [24] GUAN, P. et al. Lung cancer gene expression database analysis incorporating prior knowledge with support vector machine-based classification method. v. 28, n. 1, p. 103, 2009. ISSN 1756-9966. Disponível em: <https://doi.org/10.1186/1756-9966-28-103>.
- [25] IRION, K. L. et al. Radiograma de tórax e tomografia computadorizada na avaliação do enfisema pulmonar. *Jornal Brasileiro de Pneumologia*, Sociedade Brasileira de Pneumologia e Tisiologia, v. 33, n. 6, p. 720–732, Nov 2007. ISSN 1806-3713. Disponível em: <https://doi.org/10.1590/S1806-37132007000600017>.
- [26] SANTOS, M. L. d. O. et al. Opacidades em vidro fosco nas doenças pulmonares difusas: correlação da tomografia computadorizada de alta resolução com a anatomopatologia. *Radiologia Brasileira*, Publicação do Colégio Brasileiro de Radiologia e Diagnóstico por Imagem, v. 36, n. 6, p. 329–338, Nov 2003. ISSN 0100-3984. Disponível em: <https://doi.org/10.1590/S0100-39842003000600003>.
- [27] BACELAR, V. K. S. S. et al. Fibrose pulmonar pós-covid-19: uma revisão integrativa. *OBSERVATÓRIO DE LA ECONOMÍA LATINOAMERICANA*, v. 21, n. 9, p. 13280–13294, set. 2023. Disponível em: <https://ojs.observatoriolatinoamericano.com/ojs/index.php/ole/article/view/1325>.
- [28] BANKIER, A. A. et al. Fleischner society: Glossary of terms for thoracic imaging. *Radiology*, v. 310, n. 2, p. e232558, 2024. PMID: 38411514. Disponível em: <https://doi.org/10.1148/radiol.232558>.
- [29] HE, Y. et al. Rpild-score: A quantitative ct radiomics-based model for predicting rapidly progressive interstitial lung disease in dermatomyositis with anti-mda5 antibody. *Respiratory Research*, v. 25, n. 1, p. 112, 2024.
- [30] HAGA, H. et al. Ct-based radiomics approach to evaluate lymphocytic infiltration in the lung of patients with idiopathic interstitial pneumonias. *Japanese Journal of Radiology*, v. 42, n. 4, p. 356–364, 2024.
- [31] BERNARDI, F. et al. Radiomic features for the differentiation of usual interstitial pneumonia and non-uip fibrosing interstitial lung diseases. *Journal of Imaging*, v. 8, n. 11, p. 317, 2022.
- [32] YOON, H. et al. Deep learning-based subtype classification of interstitial lung disease using high-resolution ct images. *IEEE Transactions on Medical Imaging*, 2023. A ser completado com dados reais caso existam.
- [33] LIU, F. et al. A multimodal approach combining radiomics and clinical features for predicting post-covid interstitial lung disease. *Medical Image Analysis*, 2023. A ser completado com dados reais caso existam.
- [34] ROHMAH, S. N. et al. Lung cancer classification using radiomics and machine learning techniques. In: *2024 International Conference on Advanced and Applied Engineering (ICAAE)*. [S.l.: s.n.], 2024. p. 1–6.
- [35] BAGESTEIRO, L. D.; OLIVEIRA, L. F.; WEINGAERTNER, D. Blockwise classification of lung patterns in unsegmented ct images. In: *28th IEEE International Symposium on Computer-Based Medical Systems (CBMS)*. [S.l.: s.n.], 2015. p. 177–182.
- [36] PEDREGOSA, F. et al. Scikit-learn: Machine learning in python. *Journal of Machine Learning Research*, v. 12, p. 2825–2830, 2011.
- [37] OTSU, N. A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, v. 9, n. 1, p. 62–66, 1979.

- [38] GRIETHUYSEN, J. J. M. van et al. Computational radiomics system to decode the radiographic phenotype. *Cancer Research*, v. 77, n. 21, p. e104–e107, 2017.
- [39] POWERS, D. M. W. Evaluation: From precision, recall and f-measure to roc, informedness, markedness and correlation. *Journal of Machine Learning Technologies*, v. 2, n. 1, p. 37–63, 2011.
- [40] KABIR, M.; TAHITI, T. J.; PROMÉ, T. A. A comparative study of certain convolutional neural network architectures for x-ray image analysis in bone fracture detection and identification. In: *2024 International Conference on Artificial Intelligence, Computer, Data Sciences and Applications (ACDSA)*. [S.l.: s.n.], 2024. p. 1–8.
- [41] KHAN, S. A.; RANA, Z. A. Evaluating performance of software defect prediction models using area under precision-recall curve (auc-pr). In: *2019 2nd International Conference on Advancements in Computational Sciences (ICACS)*. [S.l.: s.n.], 2019. p. 1–6.