

Análise Comparativa entre Abordagens Convencionais e Métodos Orientados por Zonas de Aeração no Apoio ao Diagnóstico de COVID-19

Andressa G. Moreira

*Departamento de Engenharia de
Teleinformática (DETI)*

Universidade Federal do Ceará (UFC)

Fortaleza, Brazil

andressagomes@alu.ufc.br

Antonio E. P. Alves

*Departamento de Engenharia de
Teleinformática (DETI)*

Universidade Federal do Ceará (UFC)

Fortaleza, Brazil

ermeson.alves@lesc.ufc.br

Gabriel V. de Oliveira

*Departamento de Engenharia de
Teleinformática (DETI)*

Universidade Federal do Ceará (UFC)

Fortaleza, Brazil

gbarielvaz@gmail.com

Pedro Crosara Motta

*Departamento de Engenharia
Biomédica (COPPE)*

Universidade Federal do Rio de Janeiro (UFRJ)

Rio de Janeiro, Brazil

pedro.motta@lesc.ufc.br

Paulo C. Cortez

*Departamento de Engenharia de
Teleinformática (DETI)*

Universidade Federal do Ceará (UFC)

Fortaleza, Brazil

cortez@lesc.ufc.br

Resumo—A COVID-19 ainda representa um desafio global de saúde. Nesse contexto, a Tomografia Computadorizada (TC) de tórax desempenhou um papel fundamental durante a pandemia. No entanto, a interpretação dessas imagens é complexa. Diante disso, este trabalho tem como objetivo desenvolver um fluxo de trabalho para segmentação pulmonar e classificação de casos de COVID-19. Para isso, foram realizados experimentos com técnicas de *transfer learning* aplicadas tanto à segmentação quanto à classificação, explorando também a análise dos percentuais de aeração pulmonar. A combinação da arquitetura Res-Net50 com FPN obteve resultados promissores na segmentação, alcançando 99,50% de acurácia no teste, enquanto a Efficient-NetB7 apresentou o melhor desempenho na classificação, com acurácia de 96,39%.

Index Terms—COVID-19, classificação, segmentação, aeração, aprendizado de máquina.

I. INTRODUÇÃO

A COVID-19 é uma doença infecciosa que representa um desafio significativo para a saúde pública global [12]. Além dos impactos diretos na saúde da população, a pandemia ocasionou efeitos prejudiciais expressivos na economia e na qualidade de vida das pessoas [8]. De acordo com o relatório da Organização Mundial da Saúde (OMS), no Brasil, mais de 37,7 milhões de pessoas foram infectadas pela COVID-19, com uma taxa de mortalidade de 1,86% [18].

Nesse contexto, a Tomografia Computadorizada (TC) de tórax teve um papel fundamental ao longo da pandemia, sendo utilizada no diagnóstico primário, na resolução de

questões clínicas e na avaliação de possíveis complicações. No entanto, a interpretação dessas imagens apresenta desafios consideráveis, pois os achados radiológicos são frequentemente inespecíficos e podem ser observados em diversas outras doenças pulmonares [14, 6].

Além disso, a sensibilidade e a especificidade da TC variam significativamente entre os estudos, e sua interpretação está fortemente condicionada ao grau de experiência do radiologista. A ausência de validação ampla de diretrizes e a variabilidade entre os profissionais contribuem para uma considerável inconsistência nos laudos, dificultando a tomada de decisão clínica [14, 6].

Portanto, este trabalho tem como objetivo desenvolver um fluxo de trabalho para realizar a segmentação pulmonar e a classificação de casos de COVID-19 a partir de imagens de TC do tórax. Para isso, são empregadas redes neurais profundas tanto na segmentação da região pulmonar quanto na classificação. Além da abordagem direta com imagens originais, também são exploradas estratégias baseadas na análise das zonas de aeração pulmonar, com o intuito de enriquecer a avaliação clínica.

Para atingir os objetivos, foram conduzidos experimentos organizados em duas macro etapas: (i) a segmentação da região pulmonar em imagens de TC e (ii) a classificação entre casos de COVID-19 e pulmões saudáveis. Além disso, foram exploradas diferentes estratégias para a tarefa de classificação, incluindo a abordagem direta, utilizando as imagens originais, e uma abordagem baseada na análise das zonas de aeração pulmonar, derivadas da quantificação das áreas pouco aeradas.

Desse modo, os resultados obtidos foram analisados e comparados por meio de métricas quantitativas de desempe-

Os autores agradecem ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), sob o processo Número 15447/2023-0, a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES)- Código 001 e a Fundação Cearense de Apoio ao Desenvolvimento Científico e Tecnológico (Funcap) por oferecer suporte a esta pesquisa.

no, incluindo acurácia, precisão, *recall* e *F1-Score* para as tarefas de classificação, além da métrica de Interseção sobre União (IoU) e da Distância de *Hausdorff* para a avaliação da segmentação pulmonar.

Logo, as principais contribuições deste trabalho incluem:

- (i) o desenvolvimento de um fluxo de trabalho que integra a segmentação pulmonar com técnicas de classificação de casos de COVID-19 em imagens de TC;
- (ii) a avaliação comparativa de estratégias de classificação, incluindo abordagens baseadas diretamente nas imagens segmentadas e métodos orientados pela quantificação das zonas de aeração pulmonar.

O restante deste trabalho está estruturado da seguinte forma: a Seção II apresenta os trabalhos relacionados encontrados na literatura. A Seção III descreve as etapas metodológicas adotadas na condução deste estudo. Na Seção IV, são discutidos e analisados os resultados obtidos por meio dos experimentos realizados. Por fim, a Seção V apresenta as considerações finais e propõe possíveis direções para trabalhos futuros.

II. TRABALHOS RELACIONADOS

Maiello et al. [10], propuseram um algoritmo para segmentação tridimensional de pulmões, baseado na arquitetura UNet [13]. O treinamento e validação do algoritmo foram baseados em varreduras de pacientes de UTI com diferentes doenças pulmonares, incluindo COVID-19. No entanto, para discernir entre duas possíveis arquiteturas de rede, foi criado um conjunto de dados com exames de porcos e posteriormente foi realizada a transferência de aprendizado. Como forma de avaliação do sistema de segmentação proposto, a correta determinação dos níveis de aeração pulmonar e do volume pulmonar foram critérios chaves. Os autores destacam que essas medidas derivadas da segmentação, como os níveis de aeração, serviram para averiguar o método proposto.

Em [15], os autores realizaram uma análise dos pulmões de pacientes com COVID-19. Para quantificar o volume pulmonar afetado pela doença descrito como “parênquima comprometido” utilizaram valores específicos de unidades *Hounsfield* (UH) para segmentar áreas de interesse, distinguindo tecido comprometido de tecido não afetado. Um detalhe importante é que os limites mínimo e máximo de UH empregados estão muito próximos aos intervalos de UH correspondentes a tecido nada aerado e pouco aerado, conforme descrito na literatura. O estudo mostrou que a porcentagem de parênquima pulmonar comprometido foi significativamente maior em pacientes com COVID-19 do que em indivíduos saudáveis, apresentando médias de 8,126% e 1,432%, respectivamente.

Da mesma forma, os autores de [7], utilizaram uma análise de tomografia computadorizada, análoga à realizada no [15], empregando faixas de unidades *Hounsfield* (UH) específicas para pulmão comprometido e normal. Entretanto, o insight obtido a partir desses dados foi explorado de forma distinta: em vez de apenas observar pacientes com COVID-19, os autores demonstraram, com base em métricas como a curva ROC e a curva precisão-recall, que existe uma relação entre

a necessidade de suporte ventilatório e a extensão da área pulmonar comprometida.

O artigo de [1] investigou a quantificação do volume pulmonar bem aerado (WAL) em pacientes com pneumonia por COVID-19, com base em tomografias computadorizadas (TC) realizadas na admissão hospitalar. Definido pela faixa de atenuação entre -950 e -700 HU, o WAL foi analisado como possível preditor de desfechos clínicos adversos, como admissão em UTI ou óbito. Os resultados mostraram que volumes reduzidos de pulmão bem aerado estavam significativamente associados a pior prognóstico. Valores de WAL inferiores a 70–75% ou volumes abaixo de 2,9 L foram identificados como fortes preditores de complicações. O estudo conclui que a quantificação automatizada do WAL na TC inicial é um marcador independente e útil para estratificação de risco em pacientes com COVID-19.

Embora diversos estudos analisem isoladamente a aeração pulmonar em exames de tomografia de pacientes com COVID-19, neste trabalho é analisado a integração do conhecimento clínico com técnicas avançadas de visão computacional. Por meio do uso de redes neurais profundas, foi realizada a segmentação automatizada dos pulmões e a extração de padrões de aeração, que são então utilizados como base para a classificação entre casos de COVID-19 e pulmões saudáveis. Essa abordagem permite quantificar alterações pulmonares associadas à doença e fornecer informações mais objetivas sobre o comprometimento pulmonar.

III. METODOLOGIA

Esta seção descreve as etapas metodológicas empregadas na condução dos experimentos. Serão abordados detalhes sobre o fluxo dos experimentos, os conjuntos de dados utilizados, pré-processamento e aumento de dados, as técnicas aplicadas no treinamento dos modelos de segmentação e classificação da COVID-19, o cálculo e criação das máscaras das zonas de aeração do pulmão, além das métricas adotadas para a avaliação do desempenho.

A. Abordagens dos Experimentos

Os experimentos foram divididos em duas etapas principais: segmentação pulmonar em imagens de TC e classificação entre casos de COVID-19 e pulmões saudáveis. Inicialmente, foram aplicadas técnicas de pré-processamento e aumento de dados para enriquecer o conjunto de treinamento. Em seguida, modelos de segmentação baseados em arquiteturas pré-treinadas foram treinados e aplicados ao conjunto COVIDxCT para extrair as regiões pulmonares. Para a classificação, foram testadas duas abordagens: classificação direta das imagens segmentadas e análise das zonas de aeração pulmonar. Por fim, os resultados foram avaliados e comparados utilizando métricas de desempenho. A Figura 1 apresenta as etapas realizadas no fluxo de trabalho proposto.

B. Base de Dados

Para a tarefa de segmentação pulmonar, foram combinados três conjuntos de dados públicos disponíveis na literatura

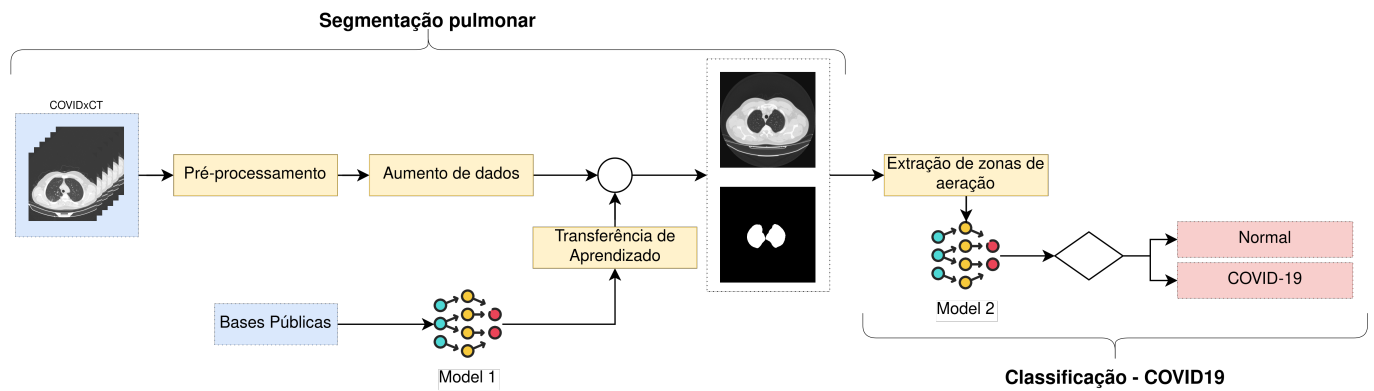


Figura 1: Fluxograma Metodológico.

Coronacases [5], *Kaggle* [2] e *Medical Segmentation* [9], totalizando 3.677 imagens de tomografias computadorizadas e suas respectivas máscaras binárias pulmonares. Para padronização, todas as imagens nos formatos DICOM ou NIFTI foram convertidas para PNG, com valores normalizados para o intervalo de unidades de Hounsfield entre 0 e 255. A Figura 2 apresenta exemplos das imagens e suas máscaras correspondentes de cada conjunto de dados.

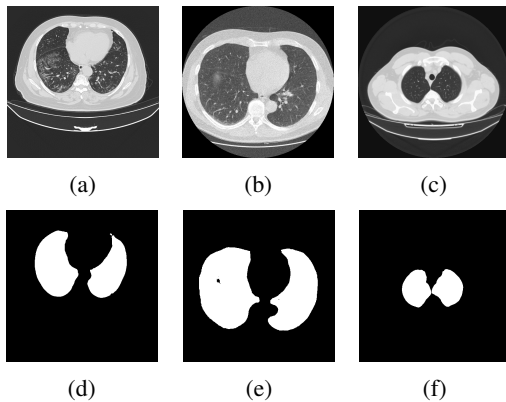


Figura 2: Amostras dos conjuntos de dados *Coronacases*, *Kaggle* e *Medical Segmentation*. A linha superior exibe as imagens originais e a inferior apresenta as respectivas máscaras pulmonares. Fonte: [5, 2, 9].

Na etapa de classificação, foi utilizado o conjunto de dados COVIDxCT [4], disponível publicamente na literatura, composto por imagens de TC no formato PNG. Para o treinamento, foram selecionadas aleatoriamente 2.000 imagens de casos positivos para COVID-19 e 2.000 imagens da classe normal. Ademais, 1.000 imagens de cada classe foram utilizadas para validação e teste. A Figura 3 apresenta exemplos de imagens das classes Normal e COVID-19.

Ademais, a Tabela I reúne as principais informações dos conjuntos de dados utilizados neste trabalho, indicando a respectiva tarefa à qual foi empregado e o total de imagens.

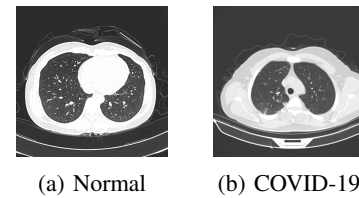


Figura 3: Amostras da base de dados COVIDxCT. Fonte: [4].

Tabela I: Informações dos Conjuntos de Dados.

Base de Dados	Segmentação	Classificação	Total de Imagens
<i>Coronacases</i> [5]	✓	✗	2.581
<i>Kaggle</i> [2]	✓	✗	267
<i>Medical Seg.</i> [9]	✓	✗	829
COVIDxCT [4]	✗	✓	6.000

C. Pré-processamento

O pré-processamento de imagens foi aplicado com o objetivo de transformar os dados brutos em imagens limpas, para melhorar a qualidade visual e reduzir artefatos indesejados. Neste projeto, as imagens provenientes dos bancos de dados foram convertidas dos formatos DICOM ou NIFTI para PNG, redimensionadas para 224×224 pixels e normalizadas canal a canal com base nas médias e desvios padrão, a fim de padronizar os valores de pixel e facilitar o aprendizado do modelo.

D. Aumento de Dados

Para ampliar o tamanho e a variabilidade do conjunto de dados de treinamento, mitigar o *overfitting* e aperfeiçoar a capacidade de generalização dos modelos, foram aplicadas técnicas de aumento de dados por meio de transformações nas imagens. Neste projeto, foram utilizadas transformações afins, por exemplo, inversão horizontal aleatória, rotação dentro do intervalo de -15 a 15 graus e desfoque gaussiano com kernel 3×3 , gerando variações artificiais que tornam o modelo mais robusto.

E. Segmentação

Na tarefa de segmentação da região pulmonar em imagens de tomografia computadorizada do tórax, foi adotada uma arquitetura do tipo *encoder-decoder*. Neste trabalho, como codificador, utilizou-se a rede ResNet50 pré-treinada, responsável por extrair e codificar a imagem de entrada em representações hierárquicas de características. Para o decodificador, foi empregada a arquitetura FPN, a fim de reconstruir as representações, projetando os atributos discriminativos aprendidos de volta ao espaço de pixels, possibilitando a segmentação das estruturas pulmonares.

Logo, durante a etapa de segmentação, a arquitetura *encoder-decoder* recebe como entrada uma imagem de tomografia computadorizada e a respectiva máscara pulmonar, utilizada como referência no treinamento. Como resultado, o modelo gera uma máscara predita, indicando a área segmentada, correspondente à região pulmonar.

Ademais, foi empregada a técnica de validação cruzada *k-fold*, com $k = 10$. Nessa abordagem, o conjunto de dados é particionado aleatoriamente em 10 *folds*. Inicialmente, os dados foram divididos em 80% para treinamento e 20% para teste, sendo que 10% da porção de treinamento foi reservada para validação. Ao final do processo, o desempenho do modelo foi obtido pela média dos resultados retornados em cada uma das iterações.

Desse modo, para o treinamento dos modelos, foram utilizadas 50 épocas, com critério de parada antecipada (*early stopping*) configurado com paciência de 5 épocas. Além disso, os foram adotados os seguintes hiperparâmetros:

- Tamanho do lote (*Batch Size*): 4;
- Função de perda: *Lovász*;
- Otimizador: Adam;
- Taxa de aprendizado = 0,0005.

F. Classificação

No processo de classificação binária, foram utilizadas arquiteturas de redes neurais convolucionais (CNN) pré-treinadas, com o objetivo de estimar a probabilidade de uma amostra pertencer a uma determinada classe. Para isso, aplicou-se a técnica de *Transfer Learning*, aproveitando o conhecimento previamente adquirido por modelos CNN treinados em grandes conjuntos de dados, com pesos pré-treinados do ImageNet. Neste projeto, as arquiteturas EfficientNetB7, ResNet50 e VGG16, todas com pesos pré-treinados no ImageNet, foram empregadas para classificar imagens entre casos de COVID-19 e normais.

Neste trabalho, o treinamento dos modelos de classificação foi realizado utilizando o conjunto de dados COVIDxCT, que já disponibiliza uma divisão pré-definida entre os subconjuntos de treinamento, validação e teste. Portanto, não foi necessário aplicar a validação cruzada *k-fold*.

Logo, para otimizar o desempenho dos modelos, foram ajustados um conjunto de hiperparâmetros. Cada modelo foi treinado por 10 épocas, definidas como o número de passagens completas pelo conjunto de dados durante o treinamento. A função de perda utilizada foi a Entropia Cruzada, amplamente

reconhecida para problemas de classificação binária e multi-classe. O otimizador escolhido foi o Adam, com uma taxa de aprendizado fixada em 0,0001. O treinamento foi realizado com um *batch size* de 4 amostras por iteração.

G. Percentuais de Aeração

Para o cálculo das zonas de aeração pulmonar, foi aplicado um modelo de segmentação da região pulmonar no banco de dados COVIDxCT, com o objetivo de preservar exclusivamente os voxels contidos dentro da máscara pulmonar.

Durante o processo de quantificação da aeração em imagens de TC, tanto em pulmões normais quanto em casos de COVID-19, foram extraídos atributos quantitativos, como a porcentagem da área relativa dos voxels, classificados com base em sua densidade, expressa em unidades *Hounsfield* (UH).

A quantificação teve início com a conversão das imagens, originalmente em formato PNG, para valores simulados em UH. Essa conversão foi realizada por meio de uma função de mapeamento linear, que transformou os níveis de cinza (0 a 255) em uma faixa de densidade de -1100 a 100 UH, aproximando-se das condições reais de densidade tecidual pulmonar. As faixas de densidade consideradas, conforme descrito na literatura [3, 17], são as seguintes: área de hiperaeração (-1.100 a -951 UH); normalmente aerada (-950 a -501 UH); pouco aerada (-500 a -101 UH); e não aerada (-100 a +100 UH).

Além disso, foram geradas máscaras coloridas com pseudo cores, que destaca visualmente as diferentes zonas de aeração, facilitando a interpretação das regiões pulmonares normais e alteradas. Os resultados finais incluem imagens segmentadas com codificação por cores de acordo com as faixas de densidade, acompanhadas da quantificação percentual e da área relativa de cada zona de aeração, permitindo uma avaliação visual e objetiva do estado pulmonar.

H. Classificação Baseada em Máscaras de Aeração

Uma segunda abordagem para a classificação binária entre casos de COVID-19 e pulmões normais foi desenvolvida com base em máscaras de aeração extraídas de imagens de TC. Dentre as zonas de aeração definidas por faixas de densidade em unidades *Hounsfield* (UH), foi selecionada especificamente a região pouco aerada, por estar mais associada a alterações pulmonares típicas da COVID-19.

A Figura 4 ilustra o processo de criação da máscara de aeração utilizada na classificação. Inicialmente, foi gerada uma imagem binária correspondente às regiões pouco aeradas do pulmão. Em seguida, a imagem de entrada para os modelos de classificação foi construída por meio da fusão RGB de três elementos, na qual: o canal R (vermelho) corresponde à imagem original segmentada em escala de cinza, o canal G (verde) representa a multiplicação da imagem segmentada pela máscara das áreas pouco aeradas e o canal B (azul) é composto apenas pela máscara das áreas pouco aeradas. Para realçar as regiões de interesse, foi aplicado um realce adaptativo, aumentando a visibilidade das alterações pulmonares relevantes. As imagens RGB compostas pelas fusões

dos elementos foram utilizadas como entrada para modelos pré-treinados, como ResNet-50, EfficientNetB7 e VGG16 no processo de classificação binária.

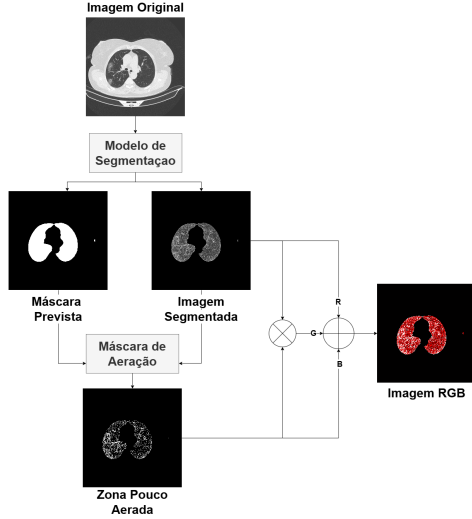


Figura 4: Criação de imagem RGB para classificação binária.

I. Métricas de Avaliação

Para uma avaliação precisa do desempenho dos algoritmos de DL, é essencial empregar métricas de avaliação apropriadas. Dessa forma, um conjunto de métricas são utilizadas para determinar se o modelo está de fato atendendo aos requisitos estabelecidos.

1) *Matriz de Confusão*: Trata-se de uma tabela cruzada, onde suas linhas registram o número de ocorrências para a classe real e as colunas apresentam os rótulos preditos por um classificador. Um caso particular de uso dessa matriz é a matriz de confusão para classificação binária. No exemplo abaixo, as classes são: Normal e COVID-19. Os valores das entradas são: VN (verdadeiros negativos), FP (falsos positivos), FN (falsos negativos) e VP (verdadeiros positivos). Com base nessas quantidades, é possível determinar algumas métricas de qualidade dos modelos.

2) *Métricas com base na matriz de confusão*: Em contextos clínicos, é fundamental priorizar métricas que penalizem mais fortemente os Falsos Negativos, uma vez que o impacto de um diagnóstico omitido (FN) é geralmente mais grave do que o de um diagnóstico incorreto (FP). Embora a acurácia seja uma métrica geral amplamente utilizada, ela pode ser inadequada em cenários com classes desbalanceadas, sendo recomendável considerar também outras métricas, como o *recall*.

Desse modo, a acurácia (1) indica a proporção de classificações corretas em relação ao total de amostras. A precisão (2) mede a proporção de amostras corretamente classificadas como positivas entre todas as classificadas como positivas. O *recall* (3) quantifica a proporção de amostras positivas corretamente identificadas. Já o *F1-Score* (4) representa a média harmônica entre a precisão e o *recall*, oferecendo um equilíbrio entre ambos os aspectos.

$$Acuracia(Acc) = \frac{VP + VN}{VP + VN + FP + FN} \quad (1)$$

$$Precisao(Prec) = \frac{VP}{VP + FP} \quad (2)$$

$$Recall = \frac{VP}{VP + FN} \quad (3)$$

$$F_1score = \frac{2 \times Prec \times Recall}{Prec + Recall} \quad (4)$$

No contexto da avaliação de resultados de segmentação semântica de imagens médicas, é necessário considerar não apenas a precisão da classificação mas também a correteza da localização [11]. Diante disso, muitas vezes as métricas acima podem não ser adequadas. Se a predição do modelo for uma máscara de segmentação completamente escura, a acurácia ainda pode ter um valor alto (dependendo da região de interesse). Assim, muitos trabalhos atuais têm empregado métricas mais apropriadas para essa tarefa, como a Interseção sobre União (IoU) (5) e a Distância de *Hausdorff* (6) [16].

$$IoU = \frac{VP}{VP + FP + FN} \quad (5)$$

$$Hausdorff = \max(h(A, B), h(B, A)) \quad (6)$$

$$h(A, B) = \max_{a \in A} \min_{b \in B} \|a - b\| \quad (7)$$

Nas Equações 6 e 7, A e B são conjuntos de pontos.

IV. RESULTADOS

Nesta seção, são analisados e discutidos em detalhe os resultados obtidos a partir da metodologia aplicada à segmentação pulmonar e classificação de COVID-19 em imagens de tomografias computadorizadas do tórax.

A. Segmentação

Algumas das métricas apresentadas, foram usadas para avaliação dos resultados da segmentação da região pulmonar em imagens de TC.

As métricas quantitativas de avaliação são fundamentais para mensurar o desempenho do modelo de segmentação. No contexto da segmentação pulmonar, a acurácia representa a proporção de pixels corretamente classificados em relação ao total. O *F1-Score*, corresponde à média harmônica entre o *recall*, que reflete a capacidade do modelo em identificar todos os pixels pertencentes à região pulmonar e a precisão, que avalia a proporção de pixels segmentados que realmente pertencem ao pulmão. A métrica IoU indica o grau de sobreposição entre a região pulmonar real e a segmentada, sendo um indicativo da qualidade da segmentação. Por fim, a Distância de *Hausdorff* quantifica o erro na delimitação das bordas, refletindo a precisão na detecção dos contornos pulmonares.

A Tabela II apresenta os resultados de desempenho do modelo baseado na arquitetura ResNet50 combinada com FPN na segmentação da região pulmonar em imagens de TC, considerando os 10 folds utilizados no processo de validação cruzada. A acurácia média de 99,5% reflete a capacidade do modelo em classificar corretamente os pixels da imagem. O *F1-Score*, com média de 97,8%, demonstra um equilíbrio entre precisão e *recall*, evidenciando a capacidade de identificação dos pixels pulmonares e a redução de falsos positivos. A métrica IoU, com média de 95,6%, demonstra a sobreposição entre as regiões segmentadas e as áreas pulmonares reais, sugerindo uma segmentação precisa. Por fim, a Distância de *Hausdorff*, com valor médio de aproximadamente 3,95, indica o desempenho na delimitação das bordas pulmonares, aspecto fundamental para uma segmentação confiável.

Tabela II: Resumo das Métricas de Segmentação para os 10 Folds da Validação Cruzada.

<i>Fold</i>	<i>Acc (%)</i>	<i>F1-Score (%)</i>	<i>IoU (%)</i>	<i>Hausdorff</i>
1	99,50	97,70	95,50	3,84
2	99,50	97,70	95,50	4,03
3	99,50	97,80	95,70	3,86
4	99,50	97,80	95,60	4,06
5	99,60	98,00	96,10	3,62
6	99,60	98,00	96,00	3,89
7	99,50	97,80	95,70	3,86
8	99,40	97,60	95,20	4,03
9	99,40	97,40	94,90	4,08
10	99,50	97,70	95,40	4,17
Média	99,50 ± 0,05	97,78 ± 0,17	95,56 ± 0,33	3,95 ± 0,15

Além da avaliação quantitativa, os resultados visuais complementam a análise, evidenciando a eficácia da arquitetura utilizada na segmentação da região pulmonar. A Figura 5 ilustra a segmentação de imagens de TC baseada na combinação dos modelos ResNet50 e FPN. A primeira coluna apresenta as máscaras originais da região pulmonar. A segunda coluna mostra as máscaras previstas pelo modelo, enquanto a terceira coluna exibe as imagens de TC com a sobreposição dos contornos das máscaras. Nessas imagens, o contorno em verde corresponde à segmentação de referência (máscara original), e o contorno em vermelho representa a segmentação gerada pela rede. Vale ressaltar a semelhança entre as segmentações previstas e as máscaras originais, combinada aos resultados quantitativos, evidencia a precisão da arquitetura na tarefa de segmentação pulmonar.

Por fim, o modelo de segmentação foi aplicado a um conjunto de dados externo, não utilizado durante o treinamento e que não possui máscaras pulmonares disponíveis, com o objetivo de gerar segmentações automáticas da região pulmonar, essenciais para a análise das zonas de aeração. A Figura 6 ilustra os resultados obtidos com o modelo baseado na arquitetura ResNet50 e FPN aplicado ao conjunto de dados COVIDxCT. A primeira coluna exibe a imagem original, a segunda mostra a máscara pulmonar prevista pelo modelo e a terceira apresenta a imagem segmentada com a região pulmonar de interesse destacada.



Figura 5: Resultados visuais da segmentação da região pulmonar em imagens de TC, utilizando a arquitetura ResNet50 e FPN.

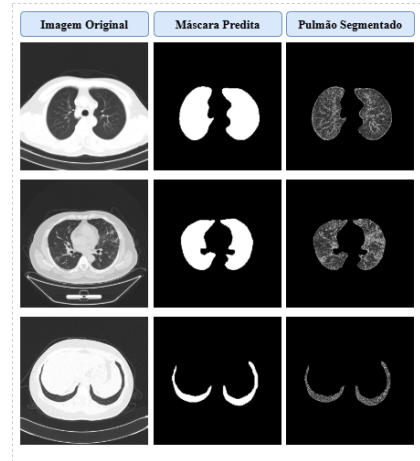


Figura 6: Segmentação da região pulmonar aplicado ao conjunto de dados COVIDxCT, utilizando a arquitetura ResNet50 e FPN.

B. Classificação

No contexto da classificação binária, a análise das métricas de desempenho é de suma importância para assegurar a confiabilidade do diagnóstico. Dessa forma, o desempenho dos modelos foi avaliado quantitativamente por meio das métricas de acurácia, precisão, *recall* e *F1-Score*. A acurácia indica a capacidade do modelo em classificar corretamente os casos de COVID-19 e os normais. A precisão avalia a proporção de verdadeiros positivos entre as amostras classificadas como COVID-19, sendo essencial para evitar alarmes falsos. O *recall* mede a capacidade do modelo em identificar corretamente os casos positivos reais, o que é crucial para não negligenciar diagnósticos de COVID-19. Por fim, o *F1-Score* representa o equilíbrio entre precisão e *recall*, oferecendo uma visão integrada do desempenho do modelo.

Neste trabalho, o treinamento dos modelos de classificação foi realizado utilizando o conjunto de dados COVIDxCT, que

já disponibiliza uma divisão pré-definida entre os subconjuntos de treinamento, validação e teste. Dessa forma, não foi necessário aplicar a validação cruzada *k-fold* para fins de comparação com benchmarks. Com base nessa estrutura, foi realizado o treinamento dos modelos pré-treinados EfficientNetB7, ResNet50 e VGG16 para a tarefa de classificação binária entre casos de COVID-19 e normais em imagens de TC.

A Tabela III apresenta o desempenho dos modelos EfficientNetB7, ResNet50 e VGG16 na tarefa de classificação binária entre casos de COVID-19 e normais. Em suma, o modelo EfficientNetB7 apresentou os melhores resultados em todas as métricas avaliadas. Com acurácia de 96,39%, o modelo apresentou o melhor desempenho para classificar corretamente tanto os casos positivos quanto os negativos. Além disso, atingiu precisão e *recall* de 96,39% e 96,40%, respectivamente, indicando não apenas a confiabilidade nas previsões positivas como também a capacidade de identificar corretamente os casos reais de COVID-19. O *F1-Score*, também foi superior no EfficientNetB7 (96,40%). Em comparação, a ResNet50 apresentou desempenho levemente inferior, com acurácia de 96,10% e *F1-Score* de 96,01%. Em contrapartida, o modelo VGG16 obteve os menores valores em todas as métricas, com destaque para o *F1-Score* de 94,62% e um valor de perda (*loss*) mais elevado (0,28), evidenciando menor eficácia na tarefa. Esses resultados indicam que, entre os modelos avaliados, o EfficientNetB7 é o mais adequado para a classificação automatizada de COVID-19.

Tabela III: Resumo das Métricas de Desempenho na Classificação de Casos de COVID-19 e Normais.

Métrica	EfficientNetB7	ResNet50	VGG16
Acurácia (%)	96,39	96,10	94,59
Precisão (%)	96,39	96,15	94,65
Recall (%)	96,40	96,10	94,59
<i>F1-Score</i> (%)	96,40	96,01	94,62
Loss	0,12	0,17	0,28

Adicionalmente, foram calculados os percentuais de aeração pulmonar e geradas as respectivas máscaras segmentadas, com o intuito de realizar a classificação binária entre casos de COVID-19 e normais. A Figura 7 ilustra um exemplo representativo do processo de segmentação, no qual apenas o tecido pulmonar é isolado e diferentes pseudo cores são aplicadas para evidenciar as distintas faixas de densidade. As faixas utilizadas neste trabalho correspondem a: área hiperaerada de -1.100 a -951 UH (amarelo), normalmente aerada de -950 a -501 UH (verde), pouco aerada de -500 a -101 UH (azul) e não aerada de -100 a +100 UH (vermelho). Dessa forma, observa-se uma predominância de áreas pouco aeradas na imagem de pacientes com COVID-19, enquanto regiões normalmente aeradas são mais predominantes no caso de pulmões saudáveis.

A Tabela IV apresenta um resumo das métricas de desempenho obtidas pelos modelos EfficientNetB7, ResNet50 e VGG16 na tarefa de classificação binária entre pulmões nor-

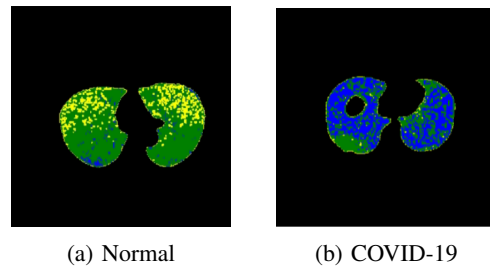


Figura 7: Máscaras de Aeração Pulmonar em Casos de COVID-19 e Pulmões Saudáveis.

mais e casos de COVID-19. Esses resultados foram alcançados a partir de uma abordagem baseada na fusão de canais RGB, na qual, o canal R corresponde à imagem original segmentada, o canal G representa a multiplicação da imagem original em escala de cinza pela máscara das zonas pouco aeradas e o canal B é composto apenas pela máscara das áreas pouco aeradas. Desse modo, o modelo ResNet50 apresentou o melhor desempenho geral, com acurácia de 90,30%, *F1-Score* de 90,26% e o menor valor de perda (0,31), indicando sua maior capacidade de generalização. Em contrapartida, os modelos EfficientNetB7 e VGG16 apresentaram valores levemente inferiores em acurácia e *F1-Score*.

Tabela IV: Resumo das Métricas de Desempenho na Classificação de Casos de COVID-19 e Pulmões Normais Baseado em Máscaras de Aeração.

Métrica	EfficientNetB7	ResNet50	VGG16
Acurácia (%)	89,10	90,30	89,80
Precisão (%)	90,44	90,95	89,85
Recall (%)	89,10	90,30	89,80
<i>F1-Score</i> (%)	89,01	90,26	89,80
Loss	0,43	0,31	0,33

Todavia, ao comparar a abordagem de máscaras de aeração com a classificação convencional, que utiliza apenas as imagens originais em escala de cinza, observa-se que a estratégia baseada na fusão de canais com ênfase nas zonas pouco aeradas resultou em desempenho inferior. Uma possível explicação para esse resultado é que a exclusividade das máscaras pouco aeradas pode ter limitado a quantidade de informação disponível para o modelo, ao negligenciar áreas relevantes do tecido pulmonar com outras densidades, como as normalmente aeradas. Dessa forma, a redução na diversidade de padrões texturais pode ter comprometido a capacidade dos modelos de capturar diferenças mais sutis entre os casos de COVID-19 e os pulmões saudáveis.

Ainda assim, essa abordagem possibilita quantificar alterações pulmonares associadas à doença, fornecendo informações objetivas e mensuráveis sobre os percentuais de aeração pulmonar. Esses dados podem servir como um importante complemento ao processo diagnóstico tradicional, auxiliando os médicos na tomada de decisão clínica e no monitoramento da progressão da doença.

V. CONCLUSÃO

Neste trabalho, foi proposto um fluxo para o diagnóstico de COVID-19 a partir de imagens de TC, envolvendo duas etapas principais: a segmentação da região pulmonar e a classificação binária entre casos positivos e pulmões saudáveis. Para isso, foram utilizadas técnicas de *transfer learning*, com arquiteturas como ResNet50 com FPN para segmentação e EfficientNetB7, VGG16 e ResNet50 para classificação. Além da abordagem convencional com imagens originais, explorou-se uma estratégia complementar baseada na análise das zonas de aeração pulmonar.

Entre as contribuições do estudo, destacam-se o desenvolvimento de um fluxo de segmentação eficiente, que alcançou 99,50% de acurácia e 95,6% de IoU, e a análise do uso máscaras de aeração, como recurso adicional para apoiar o diagnóstico. Embora essa abordagem tenha apresentado desempenho inferior ao da classificação direta, oferece informações interpretáveis e potencial para uso complementar.

Como trabalhos futuros, sugere-se a integração de dados clínicos e laboratoriais aos dados de imagem, além do desenvolvimento de uma rede neural híbrida que combine imagens de TC e percentuais de aeração pulmonar. Essa integração pode fortalecer a capacidade preditiva do sistema e oferecer suporte diagnóstico mais robusto e contextualizado.

REFERÊNCIAS

- [1] Davide Colombi et al. “Well-aerated lung on admitting chest CT to predict adverse outcome in COVID-19 pneumonia”. Em: *Radiology* 296.2 (2020), E86–E96.
- [2] *COVID-19 CT Segmentation Dataset*. <http://medicalsegmentation.com/covid19/>. Accessed: 2025-04-10. 2020.
- [3] P. A. Gevenois et al. “Pulmonary Emphysema Quantitative CT during Expiration”. Em: *Radiology* 199.3 (1996), pp. 825–829. DOI: 10.1148/radiology.199.3.8638012.
- [4] H. Gunraj et al. “COVID-Net CT-2: Enhanced Deep Neural Networks for Detection of COVID-19 from Chest CT Images through Bigger, More Diverse Learning”. Em: *Frontiers in Medicine* 8 (2022). DOI: 10.3389/fmed.2021.729287.
- [5] M. Junho et al. *COVID-19 CT Lung and Infection Segmentation Dataset*. Accessed: 2025-04-10. 2020. DOI: 10.5281/zenodo.3757476. URL: <https://zenodo.org/record/3757476>.
- [6] Saloni Laddha et al. “COVID-19 Diagnosis and Classification Using Radiological Imaging and Deep Learning Techniques: A Comparative Study”. Em: *Diagnostics* 12.8 (2022). ISSN: 2075-4418. DOI: 10.3390/diagnostics12081880. URL: <https://www.mdpi.com/2075-4418/12/8/1880>.
- [7] Ezio Lanza et al. “Quantitative chest CT analysis in COVID-19 to predict the need for oxygenation support and intubation”. Em: *European radiology* 30 (2020), pp. 6770–6778.
- [8] Mohamed Loey et al. “Fighting against COVID-19: A novel deep learning model based on YOLO-v2 with ResNet-50 for medical face mask detection”. Em: *Sustainable Cities and Society* 65 (2021), p. 102600. ISSN: 2210-6707. DOI: <https://doi.org/10.1016/j.scs.2020.102600>. URL: <https://www.sciencedirect.com/science/article/pii/S2210670720308179>.
- [9] K. Mader. *Finding and Measuring Lungs in CT Data*. <https://www.kaggle.com/kmader/finding-lungs-in-ct-data>. Accessed: 2025-04-10. 2019.
- [10] Lorenzo Maiello et al. “Automatic lung segmentation and quantification of aeration in computed tomography of the chest using 3D transfer learning”. Em: *Frontiers in physiology* 12 (2022), p. 725865.
- [11] Dominik Müller, Iñaki Soto-Rey e Frank Kramer. “Towards a guideline for evaluation metrics in medical image segmentation”. Em: *BMC Research Notes* 15.1 (2022), p. 210.
- [12] Amir Masoud Rahmani e Seyedeh Yasaman Hosseini Mirmahaleh. “Coronavirus disease (COVID-19) prevention and treatment methods and effective parameters: A systematic literature review”. Em: *Sustainable Cities and Society* 64 (2021), p. 102568. ISSN: 2210-6707. DOI: <https://doi.org/10.1016/j.scs.2020.102568>. URL: <https://www.sciencedirect.com/science/article/pii/S2210670720307861>.
- [13] Olaf Ronneberger, Philipp Fischer e Thomas Brox. *U-Net: Convolutional Networks for Biomedical Image Segmentation*. 2015. arXiv: 1505.04597 [cs.CV]. URL: <https://arxiv.org/abs/1505.04597>.
- [14] Archana Som et al. “Implementation of the Radiological Society of North America Expert Consensus Guidelines on Reporting Chest CT Findings Related to COVID-19: A Multireader Performance Study”. Em: *Radiology: Cardiothoracic Imaging* 2.5 (2020), e200276. DOI: 10.1148/ryct.2020200276.
- [15] Elif Soya et al. “Pixel-based analysis of pulmonary changes on CT lung images due to COVID-19 pneumonia”. Em: *Journal of Clinical Imaging Science* 12 (2022), p. 6.
- [16] Abdel Aziz Taha e Allan Hanbury. “Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool”. Em: *BMC medical imaging* 15 (2015), pp. 1–28.
- [17] S. R. Vieira et al. “A Lung Computed Tomographic Assessment of Positive Endexpiratory Pressure-induced Lung Overdistension”. Em: *American Journal of Respiratory and Critical Care Medicine* 158.5 (1998), pp. 1571–1577. DOI: 10.1164/ajrccm.158.5.9802101.
- [18] World Health Organization. *WHO COVID-19 Dashboard: Brazil*. Acesso em: 19 maio 2025. 2025. URL: <https://data.who.int/dashboards/covid19/cases?m49=076&n=o>.