



Ricardo de Sousa Farias

Classificação e Segmentação da Região Patológica da COVID-19 em Radiografias do Tórax Usando Redes Neurais Convolucionais

João Pessoa – PB

Março/2023



Ricardo de Sousa Farias

Classificação e Segmentação da Região Patológica da COVID-19 em Radiografias do Tórax Usando Neurais Redes Convolucionais

Dissertação de Mestrado submetida à Coordenação do Programa de Pós-Graduação em Engenharia Elétrica do Instituto Federal da Paraíba, como requisito necessário para a obtenção do grau de Mestre em Ciências no Domínio da Engenharia Elétrica.

Área de Concentração: Telecomunicações.

Linha de Pesquisa: Processamento de Sinais

Carlos Danilo Miranda Regis, Prof. Dr.

Orientador

José Manoel de Seixas, Prof. Dr.

Coorientador

João Pessoa - PB, Março de 2023

© Ricardo de Sousa Farias – ricardofariasxyz@gmail.com

Dados Internacionais de Catalogação na Publicação (CIP)
Biblioteca Nilo Peçanha do IFPB, *campus* João Pessoa

F224c Farias, Ricardo de Sousa.

Classificação e segmentação da região patológica da covid-19 em radiografias do tórax usando redes neurais convolucionais / Ricardo de Sousa Farias. – 2023.

122 f. : il.

Dissertação (Mestrado em Engenharia Elétrica) - Instituto Federal de Educação da Paraíba / Programa de Pós-Graduação em Engenharia Elétrica (PPGEE), 2023.

Orientação : Profº D.r Carlos Danilo Miranda Regis.

Coorientação : Profº D.r José Manoel de Seixas.

1. COVID – 19. 2. Radiografia do tórax. 3. Rede neural convolucional - MultiResUnet. 4. Classificação e segmentação de imagens 5. Aprendizado profundo de máquina. I. Título.

CDU 616-036.21:004.932(043)



MINISTÉRIO DA EDUCAÇÃO
SECRETARIA DE EDUCAÇÃO PROFISSIONAL E TECNOLÓGICA
INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA DA PARAÍBA

PROGRAMA DE PÓS-GRADUAÇÃO *STRICTO SENSU*
MESTRADO EM ENGENHARIA ELÉTRICA

RICARDO DE SOUSA FARIAS

**CLASSIFICAÇÃO E SEGMENTAÇÃO DA REGIÃO PATOLÓGICA DA COVID-19 EM RADIOGRAFIAS DO
TÓRAX USANDO REDES NEURAIAS CONVOLUCIONAIS**

Dissertação apresentada como requisito para obtenção do título de Mestre em Engenharia Elétrica, pelo Programa de Pós- Graduação em Engenharia Elétrica do Instituto Federal de Educação, Ciência e Tecnologia da Paraíba – IFPB - Campus João Pessoa.

Aprovado em 31 de março de 2023.

Membros da Banca Examinadora:

Dr. Carlos Danilo Miranda Regis

IFPB - PPGEE

Dr. José Manoel Seixas

UFRJ

Dr. Patric Lacouth Da Silva

IFPB - PPGEE

Dra. Suzete Elida Nobrega Correia

IFPB - PPGEE

João Pessoa/2023

Documento assinado eletronicamente por:

- **Carlos Danilo Miranda Regis**, PROFESSOR ENS BASICO TECN TECNOLOGICO, em 31/03/2023 16:16:12.
- **Patric Lacouth da Silva**, PROFESSOR ENS BASICO TECN TECNOLOGICO, em 31/03/2023 16:27:24.
- **José Manoel Seixas**, PROFESSOR DE ENSINO SUPERIOR NA ÁREA DE ORIENTAÇÃO EDUCACIONAL, em 01/04/2023 16:48:19.
- **Suzete Elida Nobrega Correia**, PROFESSOR ENS BASICO TECN TECNOLOGICO, em 03/04/2023 19:01:57.

Este documento foi emitido pelo SUAP em 22/03/2023. Para comprovar sua autenticidade, faça a leitura do QRCode ao lado ou acesse <https://suap.ifpb.edu.br/autenticar-documento/> e forneça os dados abaixo:

Código 407156
Verificador: fff6272e87
Código de Autenticação:



Av. Primeiro de Maio, 720, Jaguaribe, JOÃO PESSOA / PB, CEP 58015-435
<http://ifpb.edu.br> - (83) 3612-1200

Agradecimentos

- ☛ Gostaria de iniciar os meus agradecimentos agradecendo a Deus, por me permitir estar aqui hoje, pois sem Ele não teria chegado tão longe. Sei que foi pela Sua graça que adquiri o conhecimento necessário para realizar este trabalho e completar esta fase da minha vida. Reconheço Deus como o meu guia e protetor! Que toda honra e glória sejam dadas a Ele!
- ☛ Agradeço aos meus pais por todo o apoio ao longo de toda a minha trajetória. Também me ensinaram a ser disciplinado e a buscar com empenho todos os meus sonhos. Dedico-lhes todo o meu reconhecimento e agradecimento por dedicarem a vida de vocês para me oferecerem sempre o melhor.
- ☛ Agradeço à minha irmã, minha namorada e amigos. Obrigado pelos momentos de incentivo, conselhos e motivação. Vocês me ajudaram a chegar aqui e sempre acreditaram que eu poderia ir além.
- ☛ Agradeço ao meu orientador, Carlos Danilo, pelos ensinamentos, conselhos, apoio e confiança em meu trabalho. Agradeço pela paciência, compreensão e pelas chamadas de atenção para manter o foco. Obrigado por me incentivar a ir além.
- ☛ Agradeço ao meu coorientador, Seixas, por toda a contribuição na construção deste trabalho, com suas ideias, correções, sugestões, comentários e suporte.
- ☛ Agradeço à banca examinadora por aceitarem avaliar este trabalho e compartilhar seus conhecimentos, acrescentando valor a este trabalho.
- ☛ Por fim, agradeço a todo o corpo docente do IFPB, pela excelente formação acadêmica e profissional que me ofereceram.

“Em tudo dai graças, porque esta é a vontade de Deus em Cristo Jesus para convosco.”

1 Tessalonicenses 5:18

Resumo

A COVID-19 se disseminou em larga escala, sobrecarregando os sistemas de saúde. Embora os pulmões sejam o órgão mais afetado pela doença, é importante destacar que o vírus pode atingir outros órgãos. Para avaliar a condição dos pulmões, é comum realizar radiografias de tórax. O objetivo desta pesquisa é aplicar técnicas de aprendizado de máquina para auxiliar na análise de radiografias de tórax de pacientes com COVID-19. Neste trabalho, foi utilizada a rede neural convolucional MultiResUNet para segmentação pulmonar em imagens de radiografias do tórax. Foram usadas 6.600 imagens, sendo 2.200 de cada classe (COVID-19, Normal e Não COVID-19). Posteriormente, foram utilizadas oito redes neurais convolucionais pré-treinadas para classificar as imagens segmentadas (COVID-19, Normal e Não COVID-19) com o modelo MultiResUNet. A comparação foi realizada entre os modelos na classificação binária (COVID-19 e Não COVID-19) com 4.900 imagens (20% para teste) e na classificação multiclases (COVID-19, Normal e Não COVID-19) com 7.350 imagens (20% para teste). A MultiResUNet também foi usada para segmentar a infecção COVID-19 (área patológica) em 2.950 imagens (20% para teste), enquanto a abordagem LIME foi usada para segmentar a infecção em imagens de COVID-19 usando o melhor modelo obtido na classificação. Todos os treinamentos realizados foram utilizando a validação cruzada. A segmentação da região de infecção causada pela COVID-19 pode auxiliar os profissionais de saúde na tomada de decisão e no tratamento dos pacientes. Os resultados obtidos na segmentação pulmonar foram: acurácia de 98,1%, valor da intersecção sobre união (IoU) de 92,4% e valor do coeficiente de Dice de 96,1%. Na classificação de imagens, a VGG16 obteve os melhores resultados na classificação binária. Na classificação multiclases, a VGG16 alcançou os melhores resultados em três das cinco medidas utilizadas. Na segmentação da infecção com a MultiResUNet, foi alcançada uma acurácia de 94,2%, valor de IoU de 67,3% e valor de Dice de 80,4%. Na segmentação da infecção, foram observadas dificuldades em segmentar corretamente a área patológica. Os resultados alcançados com a segmentação da infecção utilizando LIME foram inferiores aos obtidos com a MultiResUNet.

Palavras-Chave: COVID-19, radiografia do tórax, rede neural convolucional — MultiResUNet, classificação e segmentação de imagens, aprendizado profundo de máquina.

Abstract

COVID-19 has spread on a large scale, overwhelming health systems. Although the lungs are the organ most affected by the disease, it is important to note that the virus can reach other organs. To assess the condition of the lungs, it is common to take chest X-rays. The objective of this research is to apply machine learning techniques to assist in the analysis of chest X-rays of patients with COVID-19. In this work, the MultiResUNet convolutional neural network was used for lung segmentation in chest X-ray images. 6,600 images were used, 2,200 of each class (COVID-19, Normal and Non-COVID-19). Subsequently, eight pre-trained convolutional neural networks were used to classify the segmented images (COVID-19, Normal and Non-COVID-19) with the MultiResUNet model. The comparison was performed between the models in the binary classification (COVID-19 and Non-COVID-19) with 4,900 images (20% for testing) and in the multiclass classification (COVID-19, Normal and Non-COVID-19) with 7,350 images (20 % to test). MultiResUNet was also used to segment the COVID-19 infection (pathological area) into 2950 images (20% for testing), while the LIME approach was used to segment the infection into COVID-19 images using the best model obtained in the classification. All training carried out was using cross-validation. Segmentation of the region of infection caused by COVID-19 can help health professionals in decision-making and treatment of patients. The results obtained in lung segmentation were: accuracy of 98.1%, intersection over union (IoU) value of 92.4% and value of the Dice coefficient of 96.1%. In image classification, VGG16 achieved the best results in binary classification. In the multiclass classification, VGG16 achieved the best results in three of the five measures used. In the segmentation of infection with MultiResUNet, an accuracy of 94.2%, IoU value of 67.3% and Dice value of 80.4% were achieved. In the segmentation of the infection, difficulties were observed in correctly segmenting the pathological area. The results achieved with infection segmentation using LIME were inferior to those obtained with MultiResUNet.

Key-Words: COVID-19, chest radiography, convolutional neural network — MultiResUNet, image classification and segmentation, deep machine learning.

Lista de Siglas e Abreviaturas

OMS – Organização Mundial da Saúde

Sars-Cov-2 – Síndrome respiratória aguda grave 2; do inglês, *Severe acute respiratory syndrome coronavirus 2*

ML – Aprendizado de máquina; do inglês, *Machine Learning*

DL – Aprendizado de máquina profundo; do inglês, *Deep Learning*

RT-PCR – Reação de transcriptase reversa seguida de reação em cadeia da polimerase

CXR – Radiografia de tórax; do inglês, *Chest X-ray*

CNN – Rede neural convolucional; do inglês, *Convolutional Neural network* ou *ConvNet*

AUC – Área sob a curva; do inglês, *Area Under the Curve*

RNA – Rede Neural Artificial

SDRA – Síndrome do desconforto respiratório agudo

UTI – Unidade de Terapia Intensiva

KNN – K-ésimo Vizinho mais Próximo; do inglês, *k-nearest neighbors*

SVM – Máquina de vetor suporte; do inglês; do inglês, *Support Vector Machine*

DRR – Radiografias reconstruídas digitalmente; do inglês, *Digitally Reconstructed Radiographs*

MLP – Perceptron de Múltiplas Camadas, do inglês, *Multilayer perceptron*

ReLU – Unidades Lineares Retificadas; do inglês, *Rectified Linear Units*

ROI – Região de Interesse; do inglês, *Region of Interest*

FCN – Redes totalmente convolucionais; do inglês, *Fully Convolutional Networks*

VP – Verdadeiro Positivo

VN – Verdadeiro Negativo

FP – Falso Positivo

FN – Falso Negativo

DICE – Coeficiente de Dados; do inglês, *Dice Coefficient*

IoU – Intersecção sobre União; do inglês, *Intersection Over Union*

CLAHE – Equalização de histograma adaptativo limitada por contraste; do inglês, *Contrast Limited Adaptive Histogram Equalization*

LIME – Explicações agnósticas de modelos interpretáveis locais, *Local Interpretable Model-Agnostic Explanations*

Lista de Figuras

3.1	Estrutura de uma rede neural artificial simples.	15
3.2	Estrutura de uma rede neural profunda simplificada.	16
3.3	Estrutura de uma CNN.	16
3.4	Operação de convolução.	18
3.5	Funcionamento do <i>max pooling</i> com um filtro 2×2	19
3.6	Função de ativação ReLu.	20
3.7	Função de ativação <i>sigmoid</i>	21
3.8	Função de ativação <i>softmax</i>	21
3.9	Arquitetura MultiResUNet.	25
3.10	Características de um bloco MultiRes.	26
3.11	Caminho de Res com filtros 1×1 e 3×3	27
4.1	Amostras de radiografias da classe COVID-19 com as respectivas máscaras da base COVID-QU-Ex Dataset.	34
4.2	Amostras de radiografias da classe Não COVID-19 com as respectivas máscaras da base COVID-QU-Ex Dataset.	35
4.3	Amostras de radiografias da classe Normal com as respectivas máscaras da base COVID-QU-Ex Dataset.	35
4.4	Amostras de radiografias de COVID-19 com máscaras da base QaTa-COV19 Dataset.	35
4.5	Amostras de imagens com aplicação do filtro CLAHE e aumento de dados.	39
4.6	Passos para obter o modelo a partir da MultiResUNet em cada teste.	43
4.7	Passos para obter imagens segmentadas partir do modelo MultiResUNet obtido.	44
4.8	Passos para classificar as imagens.	46
4.9	Segmentação da infecção COVID-19 com o LIME.	48

5.1	Médias das medidas de avaliação na segmentação pulmonar (Teste 1).	54
5.2	Médias das medidas de avaliação na segmentação pulmonar (Teste 2).	55
5.3	Médias das medidas de avaliação na segmentação pulmonar (Teste 3).	55
5.4	Médias das medidas de avaliação na segmentação pulmonar (Teste 4).	56
5.5	Resultado da segmentação pulmonar na faixa de 90% a 99,9% de IoU.	58
5.6	Resultado da segmentação pulmonar na faixa de 80% a 89,9% de IoU.	58
5.7	Resultado da segmentação pulmonar na faixa de 70% a 79,9% de IoU.	59
5.8	Resultado da segmentação pulmonar na faixa de 60% a 69,9% de IoU.	59
5.9	Exemplos de segmentação pulmonar utilizando as máscaras geradas pelo modelo MultiResUnet.	60
5.10	Segmentação da região de interesse das três classes utilizadas para classificação de imagens.	60
5.11	Matrizes de confusão classificação binária da VGG19 e VGG16.	62
5.12	Matrizes de confusão classificação binária da InceptionV3 e DenseNet201.	63
5.13	Matrizes de confusão classificação binária da Xception e ResNet50.	64
5.14	Matrizes de confusão classificação binária da InceptionResNetV2 e NASNetLarge.	64
5.15	Matrizes de confusão classificação multiclases da VGG19 e VGG16.	67
5.16	Matrizes de confusão classificação multiclases da InceptionV3 e DenseNet201.	68
5.17	Matrizes de confusão classificação multiclases da ResNet50 e Xception.	68
5.18	Matrizes de confusão classificação multiclases da InceptionResNetV2 e NASNetLarge.	69
5.19	Médias das medidas de avaliação na segmentação da infecção com a MultiResUNet (Teste 1).	70
5.20	Médias das medidas de avaliação na segmentação da infecção com a MultiResUNet (Teste 2).	71
5.21	Médias das medidas de avaliação na segmentação da infecção com a MultiResUNet (Teste 3).	72
5.22	Médias das medidas de avaliação na segmentação da infecção com a MultiResUNet (Teste 4).	72
5.23	Resultado da segmentação da infecção com a MultiResUNet na faixa de 90% a 99% de IoU.	75

5.24	Resultado da segmentação da infecção com a MultiResUNet na faixa de 80% a 89,9% de IoU.	75
5.25	Resultado da segmentação da infecção com a MultiResUNet na faixa de 70% a 79,9% de IoU.	76
5.26	Resultado da segmentação da infecção com a MultiResUNet na faixa de 60% a 69,9% de IoU.	76
5.27	Resultado da segmentação da infecção com a MultiResUNet na faixa de 50% a 59,9% de IoU.	77
5.28	Resultado da segmentação da infecção com a MultiResUNet na faixa de 40% a 49,9% de IoU.	77
5.29	Resultado da segmentação da infecção com a MultiResUNet na faixa de 30% a 39,9% de IoU.	78
5.30	Resultado da segmentação da infecção com a MultiResUNet na faixa de 20% a 29,9% de IoU.	78
5.31	Resultado da segmentação da infecção com a MultiResUNet na faixa de 10% a 19,9% de IoU.	79
5.32	Resultado da segmentação da infecção com a MultiResUNet na faixa de 0% a 9,9% de IoU.	79
5.33	Médias das medidas de avaliação na segmentação da infecção com o LIME. . .	80
5.34	Resultado da segmentação da infecção na faixa de 80% a 99.9% de IoU com LIME.	81
5.35	Resultado da segmentação da infecção na faixa de 70% a 79.9% de IoU com LIME.	82
5.36	Resultado da segmentação da infecção na faixa de 60% a 69,9% de IoU com LIME.	82
5.37	Resultado da segmentação da infecção na faixa de 50% a 59,9% de IoU com LIME.	83
5.38	Resultado da segmentação da infecção na faixa de 40% a 49,9% de IoU com LIME.	84
5.39	Resultado da segmentação da infecção na faixa de 30% a 39,9% de IoU com LIME.	84

5.40	Resultado da segmentação da infecção na faixa de 20% a 29,9% de IoU com LIME.	85
5.41	Resultado da segmentação da infecção na de 10% a 19,9% de IoU com LIME. .	85
5.42	Resultado da segmentação da infecção na faixa de 0% a 9,9% de IoU com LIME.	86
5.43	Comparativo da segmentação obtida pelo LIME em imagens com segmentação.	87
5.44	Alguns outros exemplos do comparativo da segmentação obtida pelo LIME em imagens com segmentação.	87

Lista de Tabelas

4.1	Parâmetros utilizados no aumento de dados.	38
4.2	Configurações do treinamento da MultiResUNet.	40
4.3	Divisão dos 4 testes realizados para realizar a segmentação.	41
4.4	Divisão dos testes realizados.	45
5.1	Resultados dos testes realizados em segmentação pulmonar nas medidas de IoU e Dice.	57
5.2	Resultados dos testes realizados em segmentação pulmonar nas demais medidas.	57
5.3	Resultados segmentação da infecção por faixas usando a medida IoU.	57
5.4	CNNs utilizadas para classificação binária.	61
5.5	Resultados das CNNs na classificação binária de imagens com o pulmão segmentado entre COVID-19 e Não COVID-19.	61
5.6	Tabela resumo das matrizes de confusão na classificação binária	65
5.7	CNNs utilizadas para classificação multiclases.	66
5.8	Resultados das CNNs com as imagens com o pulmão segmentado na classificação multiclases COVID-19, Normal e Não COVID-19.	66
5.9	Tabela resumo das matrizes de confusão na classificação multiclases.	69
5.10	Resultados dos testes realizados em segmentação da infecção nas medidas de IoU e Dice.	73
5.11	Resultados dos testes realizados em segmentação da infecção nas demais medidas.	73
5.12	Resultados segmentação da infecção por faixas usando a medida IoU.	74
5.13	Resultados segmentação da infecção por faixas usando a medida IoU.	80
5.14	Resultados segmentação da infecção por faixas usando a medida IoU.	86

Sumário

1	Introdução	1
1.1	Motivação	3
1.1.1	Contribuições	3
1.2	Objetivo Geral	4
1.2.1	Objetivos Específicos	4
1.3	Estrutura do Texto	4
2	Estado da Arte	6
2.1	Segmentação em imagens de radiografias de tórax	6
2.2	Classificação das imagens de radiografias de tórax	9
3	Fundamentação Teórica	12
3.1	COVID-19	12
3.2	Aprendizado supervisionado	13
3.3	Redes Neurais Artificiais	14
3.3.1	Redes Neurais Convolucionais	15
3.4	Segmentação de imagens	22
3.4.1	Segmentação da infecção COVID-19	27
3.5	Classificação das imagens de radiografias	27
3.5.1	VGG16 e VGG19	28
3.5.2	InceptionV3	28
3.5.3	DenseNet201	29
3.5.4	ResNet50	29
3.5.5	Xception	30
3.5.6	InceptionResNetV2	30

3.5.7	NASNetLarge	30
3.6	Aprendizado por transferência	31
3.7	Explicações agnósticas de modelos interpretáveis locais (LIME)	31
4	Materiais e Metodologia	33
4.1	Base de Dados	33
4.1.1	COVID-QU-Ex Dataset	34
4.1.2	QaTa-COV19 Dataset	34
4.2	Pré-processamento de dados utilizados na segmentação de imagens	36
4.3	Divisão dos conjuntos de dados nas etapas de segmentação	37
4.3.1	Conjunto de validação	37
4.4	Aumento do conjunto de dados de treinamento para segmentação	37
4.5	Segmentações utilizando a MultiResUNet	40
4.5.1	Treinamento da MultiResUNet	41
4.6	Classificação das imagens de radiografias de tórax	44
4.6.1	Pré-processamento de dados utilizados na classificação de imagens	44
4.6.2	CNNs utilizadas	45
4.6.3	Construção do modelo CNN	45
4.6.4	Treinamento das CNNs para classificação de imagens	46
4.7	Segmentação da infecção causada pela COVID-19 usando o LIME	47
4.8	Medidas de desempenho para avaliar os modelos	48
5	Resultados	53
5.1	Segmentação pulmonar com a MultiResUNet	53
5.2	Classificação binária das imagens de radiografias de tórax (COVID-19 e Não COVID-19)	61
5.3	Classificação multiclases das imagens radiografias de tórax (COVID-19, Normal e Não COVID-19)	65
5.4	Segmentação da infecção causada pela COVID-19	70
5.4.1	Segmentação da Infecção COVID-19 com a MultiResUNET	70
5.4.2	Segmentação da Infecção COVID-19 com o LIME	78
5.4.3	Segmentação da infecção COVID-19 com MultiResUNet vs LIME	84

6 Conclusão	88
6.1 Sugestões de Trabalhos Futuros	90
6.2 Artigos Publicados	91
Referências Bibliográficas	104

Em 11 de março de 2020, a Organização Mundial de Saúde (OMS) declarou o surto de SARS-CoV-2 como uma pandemia, devido ao crescente número de casos fora da China (Elfiky, 2020). No Brasil, o primeiro paciente foi diagnosticado em fevereiro do mesmo ano (Freitas *et al.*, 2020).

A pandemia da COVID-19 afetou todo o mundo, várias pessoas ficaram em quarentena devido ao contágio, o que acabou afetando a vida de todos, incapacitando as indústrias, afetando a economia mundial. Diversos campos como social, econômico e político foram atingidos pela pandemia, o que causou diversos problemas em todo o mundo. No mundo, até março de 2023 o número de infectados ultrapassou 760 milhões de pessoas e mais de 6 milhões de mortes foram registradas (OMS, 2023).

A busca por agilidade na identificação da COVID-19, impulsionam o desenvolvimento de formas alternativas para o auxílio no diagnóstico da doença, como modelos de aprendizado profundo (*Deep learning* — DL). Aplicações de DL são empregadas em tarefas de análise de imagens médicas, como classificação e segmentação (Toennies, 2012). Os bons resultados alcançados por esses modelos têm ganhado cada vez mais espaço nas diversas áreas da saúde, como na identificação de Alzheimer e tumor cerebral por meio da ressonância magnética (Silva *et al.*, 2019; Ranjbarzadeh *et al.*, 2021); detecção de câncer de mama em imagens de mamografia (Al-Antari *et al.*, 2020); detecção da tuberculose, pneumonia e COVID-19 por imagens de radiografias de tórax (Lakhani & Sundaram, 2017; Gupta *et al.*, 2021); câncer de pulmão com imagens de tomografias computadorizadas (Riquelme & Akhloufi, 2020); dentre outras.

As radiografias de tórax são comumente requeridas para os pacientes com sintomas pulmonares agudos e críticos por possuírem menor custo, velocidade relativa, portabilidade e acessibilidade a baixa exposição a radiação (Baratella *et al.*, 2020). Embora sejam menos sensíveis do que a tomografia computadorizada (TC), a radiografia é utilizada como exame de

triagem em pacientes com suspeita ou confirmação de COVID-19. (Chamorro *et al.*, 2021). Comparadas às radiografias de tórax, as TCs apresentam um custo mais elevado, além de serem menos ágeis e terem menor disponibilidade. Por esses motivos, as radiografias de tórax são exames de imagens iniciais para verificar o quadro do paciente. A partir delas consegue-se ter uma visão do quanto o pulmão está deteriorado (Jiao *et al.*, 2021), sendo utilizadas no diagnóstico e acompanhamento de pacientes com COVID-19 (Rousan *et al.*, 2020).

A interpretação bem-sucedida dos resultados por imagens no contexto da detecção de COVID-19 é um desafio significativo devido às semelhanças visuais com outras doenças pulmonares, como a pneumonia (Aggarwal *et al.*, 2022) (Hertel & Benlamri, 2022). Essa dificuldade pode ser atribuída à similaridade dos sintomas e características radiológicas entre essas doenças, dificultando a identificação manual da COVID-19 em imagens de radiografias de tórax. Além disso, a interpretação manual é uma tarefa demorada e suscetível a erros humanos.

Diante desse cenário, os pesquisadores têm recorrido ao uso das redes neurais convolucionais, uma técnica amplamente utilizada para segmentar e detectar COVID-19 em imagens médicas. Vários estudos exploram o potencial das redes neurais convolucionais (*Convolutional Neural Networks* — CNN) na detecção e classificação de COVID-19 em imagens de radiografias de tórax. Essas redes mostraram resultados promissores e estão sendo cada vez mais exploradas visando automatizar soluções para a detecção de COVID-19 (Aggarwal *et al.*, 2022), (Tahir *et al.*, 2021a), (Arias-Garzón *et al.*, 2021), (Costa *et al.*, 2021), (Chowdhury *et al.*, 2020), (Leão *et al.*, 2021), entre outros.

Neste trabalho, é proposto o uso da rede neural convolucional MultiResUNet para a segmentação pulmonar em radiografias do tórax em três classes (COVID-19, Normal e Não COVID-19). As imagens segmentadas obtidas são utilizadas para a classificação, utilizando oito redes neurais convolucionais pré-treinadas. Também são realizadas comparações entre os modelos de classificação binária e multiclases. Além disso, a MultiResUNet e o LIME (*Local Interpretable Model-Agnostic Explanations*) são usados para segmentar a região patológica causada pela infecção por COVID-19. A abordagem LIME utiliza o melhor modelo obtido na classificação. A segmentação da região de infecção causada pela COVID-19 pode auxiliar os profissionais de saúde na tomada de decisão e no tratamento dos pacientes.

1.1 – Motivação

O objetivo desta pesquisa é aplicar técnicas de aprendizado de máquina para auxiliar na análise de radiografias de tórax de pacientes com COVID-19, por se tratar de uma doença grave e altamente contagiosa que já afetou milhões de pessoas em todo o mundo. É crucial para os profissionais de saúde identificar as áreas afetadas pelo vírus em cada paciente, a fim de implementar o tratamento mais adequado e garantir o melhor prognóstico, e adaptar a estratégia de tratamento adequado.

Identificar as áreas afetadas pelo vírus, pode ser útil para o diagnóstico e o tratamento da COVID-19, mas também para o entendimento de outras doenças respiratórias e para o desenvolvimento de novas técnicas de imagem. Os resultados dessa pesquisa podem ajudar a desenvolver novos métodos de segmentação de infecções, bem como contribuir para o entendimento das características da doença.

1.1.1 – Contribuições

Neste trabalho, as partes principais são: segmentação do pulmão, classificação das imagens segmentadas usando CNNs e segmentação da infecção usando duas abordagens diferentes.

Com base nessas informações, as principais contribuições são:

- ▣ Implementação de uma solução para segmentar o pulmão em imagens de radiografias de tórax, usando a rede neural MultiResUNet.
- ▣ Implementação para classificar imagens de radiografias de tórax como COVID-19, Normal e Não COVID-19, usando diferentes modelos de CNN e comparando os resultados para ver qual tem a melhor sensibilidade.
- ▣ Implementação de uma solução para segmentar a região da infecção por COVID-19 em imagens de radiografias de tórax, usando duas abordagens diferentes: a MultiResUNet e o LIME.
- ▣ Contribuição para o conhecimento científico e técnico na área de processamento de imagem e reconhecimento de padrões, que podem ser aplicados em outras áreas além da detecção e segmentação da infecção de COVID-19. Isso pode gerar novas oportunidades de pesquisa e desenvolvimento em áreas relacionadas.

1.2 – Objetivo Geral

O presente trabalho visa explorar a eficiente de classificação e segmentação da região patológica da COVID-19 em imagens de radiografias do tórax, utilizando redes neurais convolucionais. O estudo visa avaliar o desempenho de diferentes modelos pré-treinados na classificação binária e multiclases das imagens, além de avaliar a precisão da segmentação da infecção utilizando a rede neural MultiResUNet e uma abordagem alternativa denominada LIME.

1.2.1 – Objetivos Específicos

- ▣ Estudar e implementar segmentações pulmonares de imagens médicas de radiografias utilizando a técnica de segmentação de imagens MultiResUNet.
- ▣ Estudar e aplicar modelos de redes neurais convolucionais para classificar as imagens de radiografias de tórax: COVID-19, Normal e Não COVID-19.
- ▣ Estudar e implementar segmentação da região patológica causada pela COVID-19 em imagens de radiografias para auxiliar na tomada de decisão e no tratamento de pacientes com esta infecção, para isso é utilizado a MultiResUNet e o LIME.
- ▣ Quantificar e avaliar o desempenho dos modelos gerados em relação à segmentação pulmonar, classificação das imagens e na identificação da região patológica causada pela COVID-19.

1.3 – Estrutura do Texto

Este trabalho de Dissertação de Mestrado está organizado em cinco capítulos que compreendem todos os aspectos teóricos acerca dos tópicos associados à área da pesquisa, bem como a metodologia empregada e os resultados obtidos neste trabalho.

Além deste capítulo introdutório, o Capítulo 2, trata do estado da arte na área de classificação e segmentação de imagens de COVID-19. Nesse capítulo, são introduzidos os trabalhos que buscam segmentação e classificação da COVID-19 em imagens de radiografias, como também trabalhos que buscam identificar a região patológica causada pela infecção COVID-19 em imagens de radiografias.

Na seção seguinte, no Capítulo 3, são apresentados os conceitos fundamentais necessários para o desenvolvimento deste trabalho, incluindo aprendizado de máquina, redes neurais artificiais, segmentação de imagens e classificação de imagens.

O Capítulo 4 descreve, em detalhes, os passos realizados em cada etapa do projeto, incluindo a coleta e pré-processamento de dados, a implementação dos algoritmos de segmentação e classificação, e a obtenção dos resultados.

Na seção seguinte, no Capítulo 5, são apresentados e explicados os resultados obtidos na segmentação pulmonar, classificação das imagens e na segmentação da infecção COVID-19.

Finalmente, no último capítulo, são apresentadas as conclusões finais do estudo, incluindo as principais contribuições do trabalho e sugestões para trabalhos futuros.

Há diversos estudos na literatura que abordam a detecção de doenças por meio de radiografia de tórax. Com o avanço bem-sucedido da tecnologia de aprendizado profundo no campo do processamento de imagens, a detecção de patologias por meio de inteligência artificial e aprendizado profundo se popularizou. De modo geral, os estudos presentes na literatura podem ser categorizados em diferentes grupos com relação aos métodos utilizados para extrair características e treinar modelos.

Existem vários tipos de testes para COVID-19 que estão sendo aplicado para detectar a COVID-19 em pacientes. Um desses testes, o teste de reação de transcriptase reversa seguida de reação em cadeia da polimerase (RT-PCR) é conceituado o padrão-ouro atual para diagnosticar a COVID-19 (Ai *et al.*, 2020). Entretanto, os testes de RT-PCR não são absolutos, problemas com a sensibilidade geral dos testes já foram relatados (Fang *et al.*, 2020).

No estudo realizado por Luo *et al.* (2020) com 4.653 participantes, observaram que os testes de RT-PCR têm uma sensibilidade de cerca de 71%. Na pesquisa de Kucirka *et al.* (2020) constatou-se que leva cerca de uma semana contando do início dos sintomas para que o teste de RT-PCR resulte a menor taxa de falsos negativos. Embora a taxa de falso-negativos permaneça alta em 21%. Dessa forma, quando os testes de RT-PCR são menos precisos, deixam margem para outros métodos possam ser aplicados no decorrer do percurso. A radiografia de tórax pode ser um destes métodos, podendo ser utilizada até em conjunto com outros testes como uma forma apoio ao diagnóstico do paciente.

2.1 – Segmentação em imagens de radiografias de tórax

No estudo de Tahir *et al.* (2021a), é abordado segmentação pulmonar e localização de COVID-19 com quantificação de infecção a partir de imagens de radiografias de tórax. Foi realizado um conjunto de experimentos utilizando as redes de segmentação estado-da-

arte, U-Net, U-Net ++ e Feature Pyramid Networks (FPN). A rede desenvolvida FPN com decodificador DenseNet121, após um processo iterativo, alcançou um desempenho superior para a segmentação da região pulmonar, com acurácia 99,12%, IoU de 96,11% e Coeficiente Dice de 97,99%. Além disso, as infecções de COVID-19, que foram localizadas usando a U-Net++ com codificador DenseNet121 que apresentou o melhor desempenho com valores de interseção sobre união *Intersection over Union — IoU* e Coeficiente Dice de 83,05% e 88,21%.

No trabalho de Alom *et al.* (2020) foi apresentado um método usando rede neural convolucional recorrente residual com transferência de aprendizado para a realização da detecção da COVID-19 e o modelo desenvolvido pelos autores, o NABLA-N, foi aplicado para a segmentação das regiões infectadas pela COVID-19 em imagens de radiografia e tomografia computadorizada, obtendo um percentual de 84,67% na precisão de testes para as imagens de radiografias e 98,78% para precisão da tomografia computadorizada.

Zhang *et al.* (2020a) propôs uma abordagem baseada em radiografias reconstruídas digitalmente (*Digitally Reconstructed Radiographs — DRR*) consiste em uma rede de segmentação e um módulo de adaptação de domínio, denominada DRR4Covid. A técnica de segmentação de rede neural desenvolvida utilizando DRRs com rótulos de tomografias computadorizadas foi adaptada para imagens de radiografias de tórax graças ao design do módulo de domínio. Sem usar nenhuma anotação de radiografia, a rede atingiu na classificação precisão de 94,9%, AUC (*Area Under the Curve*) de 98,7% e pontuação F1 de 94,7%. Na segmentação a precisão foi de 95,6%, AUC de 98,0% e pontuação F1 95,5% em um conjunto de teste com 558 casos normais e 558 casos positivos.

No trabalho de Arias-Garzón *et al.* (2021) utilizou-se a U-net para processar as imagens para segmentar o pulmão e a VGG19 e VGG16 com aprendizagem por transferência para classificar as imagens como positivas ou negativas para COVID-19. A interpretação dos resultados se dá por medidas de desempenho, bem como a utilização de mapas de calor com objetivo de encontrar o local de maior intensidade que representa a localização da infecção. Os melhores modelos alcançaram uma precisão de detecção de COVID-19 em torno de 97%.

A pesquisa de Kusakunniran *et al.* (2021) utilizou a arquitetura ResNet-101 como a rede principal com mais de 44 milhões de parâmetros. A rede foi treinada por completo sem utilizar transferência de aprendizado. As imagens utilizadas como entradas possuem uma resolução de 1500 × 1500 píxeis. É aplicado o mapa de calor sob a região de interesse do pulmão segmentado é construído para visualizar e enfatizar os sinais de COVID-19 em cada imagem de radiografia

de entrada. Para realizar a segmentação pulmonar foi utilizado a U-Net. Para classificar as imagens utilizaram-se duas abordagens com duas classes e três classes. COVID-19 e não COVID-19, e COVID-19, normal e outras anomalias. A solução proposta atinge sensibilidade, especificidade e precisão muito promissoras de 97%, 98% e 98%, respectivamente.

O trabalho proposto por Oliveira *et al.* (2021) realiza a classificação de duas categorias de imagens de radiografia do tórax, COVID-19 e não COVID-19. Os conjuntos de imagens utilizados possuem alta qualidade com grandes resoluções, para evitar vieses indesejados. Totalizando 7.875 imagens da classe COVID-19 e 16.841 imagens da classe não COVID-19. Na classe COVID-19, foram aplicadas 6.300 imagens para o treinamento e 1.575 para o teste, enquanto para a classe não COVID-19, 13.472 para o treinamento e 3.369 para o teste. Antes de classificar as imagens, foi aplicado a segmentação nas imagens utilizando o modelo U-Net para em seguida classificar. O modelo U-Net atingiu um IoU de 93,33%. A classificação ocorreu por meio de diversos modelos de redes neurais convolucionais, onde foi realizado um comparativo entre os modelos. Dos 25 modelos utilizados, 5 modelos obtiveram uma acurácia entre 91% e 92%, foram eles: VGG-11, ResNet-152, Densenet-161, Densenet-201 e GoogleNet. Por fim foi utilizado um mapa de calor como método de explicabilidade. Cada um dos cinco classificados que se destacaram, obtém-se um mapa de calor para cada. Dessa forma, é calculada a média entre os mapas de calor e conseqüentemente obtém-se o mapa de calor final.

Costa *et al.* (2021) realizaram a classificação de imagens de COVID-19 e normais. Sendo 1201 imagens da classe COVID-19 e 1232 da classe normal. Antes da classificação foi realizada uma segmentação nas imagens de radiografias. Para isso, foi utilizado o modelo MultiResUNet, sendo um modelo derivado da U-Net. O modelo MultiResUNet alcançou um Índice de Jaccard de 93,34% e Coeficiente de dados de 96,53%. As imagens segmentadas alimentaram os modelos COVNet, CNN e SVM (classe única). O modelo que mais se destacou foi a CNN quando comparado com a COVNet e SVM. A CNN alcançou a acurácia de 97,82 ($\pm 1, 11$), sensibilidade de 97, 03% ($\pm 1, 24$), especificidade de 99,66 ($\pm 0, 41$), pontuação F1 de 98, 40% ($\pm 0, 48$), Rrcall de 99, 63% ($\pm 0, 48$) e AUC de 98, 35% ($\pm 0, 51$).

No estudo de Sharma *et al.* (2022), foram utilizados os modelos de segmentação UNet e UNet+ em conjunto com oito modelos de classificação para detectar a COVID-19. O desempenho do sistema foi avaliada usando diversas métricas, incluindo Dice, Jaccard, AUC e ROC, e validada usando Grad-CAM em um framework de IA explicável para segmentar a COVID-19. O melhor modelo de segmentação obteve uma acurácia de 96,35%, perda de 0,15%,

Dice de 94,88%, Jaccard de 90,38% e AUC de 0,99. Já o melhor modelo de classificação baseado em segmentação, UNet+Xception, apresentou uma acurácia de 97,45%, precisão de 97,46%, recall de 97,45%, pontuação F1 de 97,43% e AUC de 0,998.

2.2 – Classificação das imagens de radiografias de tórax

A maioria dos trabalhos anteriores relacionados concentra-se no diagnóstico. Um dos primeiros trabalhos com foco na detecção da COVID-19, utilizando imagens de radiografias de tórax, foi desenvolvido por Hemdan *et al.* (2020). O trabalho propôs um modelo chamado COVIDX-Net, onde abordou diferentes modelos neurais convolucionais, entre elas, as que se destacaram foram a VGG19 e DenseNet201 que obtiveram resultados acima de 90% de acurácia.

Khan *et al.* (2020) propôs o CoroNet, um modelo de CNN para detectar COVID-19 a partir de imagens de radiografias de tórax. O modelo proposto é baseado na arquitetura Xception pré-treinada no conjunto de dados ImageNet. O modelo proposto alcançou uma acurácia de 89,6%, precisão de 93% e recall de 98,2% para casos de 4 classes: COVID-19, pneumonia bacteriana, pneumonia viral e normal. Para classificação de 3 classes: COVID-19, pneumonia, pneumonia e normal. O modelo obteve acurácia de classificação de 95%, recall de 96.9%, especificidade de 97.5%, precisão de 95%. Na classificação binária entre COVID-19 e normal. O modelo obteve acurácia de classificação de 98.3%, recall de 99.3%, especificidade de 98.6%, precisão de 95%.

Outra pesquisa desenvolvida com imagens de raios-X do tórax foi a de Nour *et al.* (2020), nela foi projetada um modelo de uma rede neural convolucional (*Convolutional Neural Network* — CNN) com 5 camadas para extrair características discriminativas, essas características foram aplicadas como entradas para algoritmos de aprendizado de máquina, K-ésimo Vizinho mais Próximo (*k-nearest neighbors* — KNN), máquina de vetores de suporte (*Support Vector Machine* — SVM) e árvore de decisão. Dentre eles, o SVM obteve os melhores resultados, acurácia de 98,97%, sensibilidade de 89,39% e especificidade de 99,75%.

Abbas *et al.* (2021) detectaram a Covid-19 utilizando o método DeTraC e várias redes CNN para detectar anomalias em imagens de radiografia do tórax. O método de ajuste fino foi utilizado para re-treinar modelos CNNs. No estudo, foram utilizadas imagens de COVID-19,

SARS e normal. Como resultado, DeTraC com VGG19 alcançou a maior acurácia de 97,35%, sensibilidade de 98,23% e especificidade de 96,34%.

O trabalho de Gupta *et al.* (2021) propôs um modelo de Rede Neural Convolutiva (CNN) empilhada e integrada onde usa outras arquiteturas de CNN pré-treinadas, foram aplicadas na detecção de 3 classes, normal, pneumonia e COVID-19 atingindo uma acurácia de 99,08%. Na classificação binária, COVID-19 e não COVID-19 obtiveram uma acurácia de 99%.

A pesquisa de Chowdhury *et al.* (2020) realizou duas classificações: uma binária, COVID-19 × Normal e outra multiclases, Pneumonia Viral, Normal, e COVID-19 com e sem *data augmentation*¹ utilizando diversos modelos pré-treinados. No problema com duas classes sem *data augmentation*, dois modelos empataram, o ResNet18 e CheXNet atingiram uma acurácia de 99,41%, sensibilidade de 99,41% e especificidade de 91,41%. Com o *data augmentation*, o modelo DenseNet201 obteve os melhores resultados, onde a acurácia e sensibilidade foram 99,70%, especificidade de 99,55%. No problema de três classes sem *data augmentation* a CheXNet obteve uma acurácia de 97,24%, sensibilidade de 96,61% e especificidade de 98,31%, enquanto com *data augmentation* a DenseNet201 obteve acurácia de 97,94%, sensibilidade de 97,94% e especificidade de 98,80%.

No estudo de Khan & Aslam (2020) foram utilizados quatro modelos de CNNs, a DenseNet121, ResNet50, VGG16 e VGG19, para o diagnóstico de imagens de raios-X como COVID-19 ou Normal. No estudo constatou-se que a VGG16 e VGG19 superaram os outros dois de CNNs e alcançaram uma precisão de 99,3%.

No trabalho de Bhadra & Kar (2020) uma arquitetura CNN de múltiplas camadas profundas e menos complexas foi proposta. Foram utilizadas 1330 imagens de casos COVID-19, 1330 imagens de casos de pneumonia e 1330 imagens de casos normais. O conjunto de dados foi dividido em 80% para treinamento, 10% para validação e 10% para teste cego, foi aplicado a validação cruzada de 10 vezes. A precisão do teste cego foi de 99,1% pela técnica proposta, enquanto as redes pré-treinadas VGG16, VGG19, Mobilenet, InceptionV3, ResNet50 utilizadas para comparação atingiram 93,9%, 92,2%, 98,8%, 99,5% e 98,3%, respectivamente.

O trabalho de Leão *et al.* (2021) utiliza imagens de radiografias de tórax e investiga duas abordagens: classificação binária (COVID-19 e Pneumonia) e multiclases (COVID-19, Pneumonia e Normal). Foram utilizados 3 modelos de redes neurais convolucionais, um dos

¹Técnicas usadas para aumentar artificialmente a quantidade de dados adicionando cópias ligeiramente modificadas de dados já existentes.

modelos é treinado do zero sem transferência de aprendizado. O modelo sem transferência de aprendizado mostrou-se melhor no cenário com duas classes. A acurácia obtida de 85,37% contra 82,11% da VGG16 e 80,49% da ResNet50V2, enquanto a abordagem com transferência de aprendizado foi superior no cenário com três classes, obteve acurácia na VGG de 87,91%, rede treinada do zero de 86,26% e ResNet50V2 de 85,71%.

Neste trabalho, a MultiResUNet é utilizada para realizar segmentação pulmonar em imagens de radiografias do tórax. Os testes incluem a aplicação de aumento de dados e do filtro CLAHE (*Adaptive histogram equaliza*). As imagens segmentadas são então classificadas em três categorias: COVID-19, Normal e Não COVID-19, utilizando oito modelos (VGG16, VGG19, InceptionV3, DenseNet201, InceptionResNetV2, NASNetLarge, Xception e ResNet50) pré-treinados da biblioteca Keras. Além disso, a infecção por COVID-19 é segmentada, utilizando tanto a MultiResUNet quanto uma abordagem alternativa chamada LIME.

Fundamentação Teórica

Neste capítulo será abordado uma visão geral sobre a COVID-19, aprendizado de máquina, redes neurais convolucionais, segmentação de imagens, classificação de imagens e os modelos utilizados na etapa de classificação.

3.1 – COVID-19

A COVID-19 é uma doença causada pela síndrome respiratória aguda grave coronavírus-2 (SARS-COV-2). O primeiro caso de COVID-19 foi relatado pela primeira vez no final de 2019 em Wuhan, na China, logo em seguida se espalhou por todo mundo em meses se tornando uma pandemia de nível global com crescimento exponencial. A OMS declarou como uma pandemia de emergência de saúde global por se tratar de um vírus altamente contagioso.

O vírus infecta principalmente a parte do sistema respiratório humano, levando a broncopneumonia grave com sintomas de febre, dispneia, tosse seca, fadiga, perda do olfato e paladar, dor de garganta e insuficiência respiratória, entre outros sintomas (Cheng *et al.*, 2020). Sendo transmitidos por pacientes assintomáticos e sintomáticos, disseminando por contato físico e gotículas respiratórias.

A COVID-19 em pessoas com sistema imunológico comprometido ou idosos têm maior probabilidade de desenvolver doenças graves, como insuficiência cardíaca, renal e choque séptico. Casos graves resultam em síndrome do desconforto respiratório agudo (SDRA) ou insuficiência respiratória completa, que requer suporte de ventilação mecânica e uma unidade de terapia intensiva (UTI) (Tahir *et al.*, 2021a). O que preocupa, pois, o número de UTIs disponíveis é limitado e muitas vezes é inferior ao número de pacientes em estado grave.

Para verificar o quão comprometido o pulmão se encontra, imagens de radiografias de tórax podem ser utilizadas, com objetivo de verificar a inflação do pulmão e tentar identificar as características para atribuir a qual doenças pertence. As características encontradas nos exames

radiológicos do tórax comuns da infecção incluem opacidade em vidro fosco e espessamento do septo interlobular nos pulmões, com correlações significativas entre o grau de inflamação pulmonar e os principais sintomas clínicos de COVID-19 (Maharjan *et al.*, 2021).

A tomografia computadorizada de tórax é melhor quando comparada com radiografia para ser a melhor modalidade para detectar e analisar alterações pulmonares. Entretanto, quando há um número grande de pacientes hospitalizados, a radiografia de tórax é o método radiológico mais comum para monitorar o curso rápido de COVID-19 (Fogante *et al.*, 2021).

Outra vantagem da imagem de radiografia é acessibilidade e a menor incidência de radiação ionizante, entretanto a identificação de COVID-19 em imagens de radiografias de pacientes com esta doença não são fáceis de interpretar, sendo a maior barreira no diagnóstico clínico. Portanto, é uma tarefa desafiadora para os radiologistas se especializarem no diagnóstico da COVID-19, especialmente em cidades com um número menor de radiologistas (Dilshad *et al.*, 2021).

As imagens de radiografias podem ser usadas para o diagnóstico clínico da COVID-19, contudo, as imagens de várias outras categorias de pneumonia viral são similares, difícil de distinguir a COVID-19 entre outras pneumonias virais. Além disso, também apresentam características similares com outras doenças inflamatórias pulmonares. O que pode levar a um diagnóstico errado, e causar algum dano ao paciente. Dessa forma, se faz necessário métodos computacionais inteligentes para auxiliar no diagnóstico clínico.

A aplicação de aprendizado de máquina pode conseguir realizar a triagem automatizada com antecedência. Úteis para auxiliar especialistas diagnosticarem a doença, uma vez que modelos computacionais podem aprender padrões que médicos não conseguem observar. Os padrões são úteis para ocorrer o aprendizado e posteriormente a classificação da doença.

3.2 – Aprendizado supervisionado

O aprendizado supervisionado é uma das técnicas mais utilizadas em Machine Learning para resolver problemas de classificação, regressão e outros. Nesse tipo de aprendizado, um algoritmo é treinado em um conjunto de dados rotulados, ou seja, um conjunto de dados que já possuem as respostas esperadas. O objetivo é que o algoritmo aprenda a mapear as entradas para as saídas esperadas, de modo que ele possa fazer previsões precisas em novos dados (Liu & Wu, 2012).

Um exemplo prático de aplicação do aprendizado supervisionado é a detecção de pneumonia em imagens de radiografia do tórax. Nesse caso, o conjunto de dados consiste em imagens de radiografia do tórax de pulmões, rotulados com a presença ou ausência de pneumonia. O algoritmo é treinado nesse conjunto de dados para identificar padrões que indicam a presença de pneumonia na imagem. Após treinado, o algoritmo pode ser usado para identificar a presença de pneumonia em novas imagens.

As redes neurais artificiais são uma das abordagens mais populares para a realização de aprendizado supervisionado. Essas redes são compostas por camadas de neurônios, que processam as entradas e geram as saídas esperadas (Haykin, 2001).

3.3 – Redes Neurais Artificiais

A rede neural artificial (RNA) surgiu na década de 40, quando McCulloch e Pitts desenvolveram o primeiro modelo neural, a partir desse momento resultou no interesse de diversos pesquisadores, desde então vários algoritmos com base em RNAs foram desenvolvidos e aplicadas em diversas áreas (Brahme, 2014).

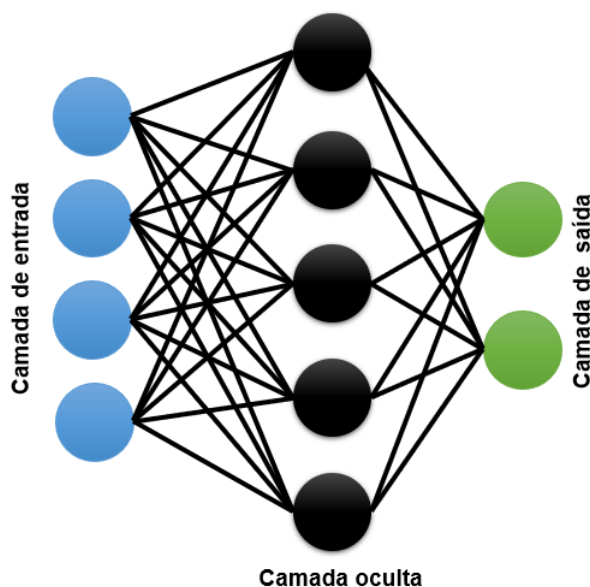
As RNAs têm como intuito resolver problemas complexos que demandam um alto processamento. Segundo Haykin (2001), uma RNA é um modelo computacional composto por vários neurônios artificiais interconectados. É uma técnica poderosa de aprendizado de máquina, frequentemente utilizada em tarefas de classificação, reconhecimento de padrões e previsão. Choudhury *et al.* (2018) define que uma RNA é formada por centenas de neurônios artificiais ou elementos de processamento, conectados com coeficientes (pesos), que constituem a estrutura neural, organizados em camadas.

Uma das principais características de uma rede neural artificial é habilidade de aprender conforme as entradas passadas. A RNA pode ser treinada tanto com aprendizado supervisionado, em que são fornecidos dados rotulados para a rede aprender a classificar novos dados, quanto com aprendizado não supervisionado, em que a rede consegue descobrir padrões por conta própria. Conforme Santos (2017), as redes neurais possuem a habilidade de receber várias entradas em simultâneo, e distribuí-las organizadamente, as informações armazenadas pela rede são compartilhadas por todas suas unidades de processamento.

Os neurônios são módulos simples, mas não lineares, conectados, que estão estruturados em camadas sequencias, onde a saída de uma camada é a entrada da próxima camada (Rodrigues

et al., 2021). Na Figura 3.2 é apresentado a estrutura de organização em camadas de uma rede neural artificial simples.

Figura 3.1 – Estrutura de uma rede neural artificial simples.



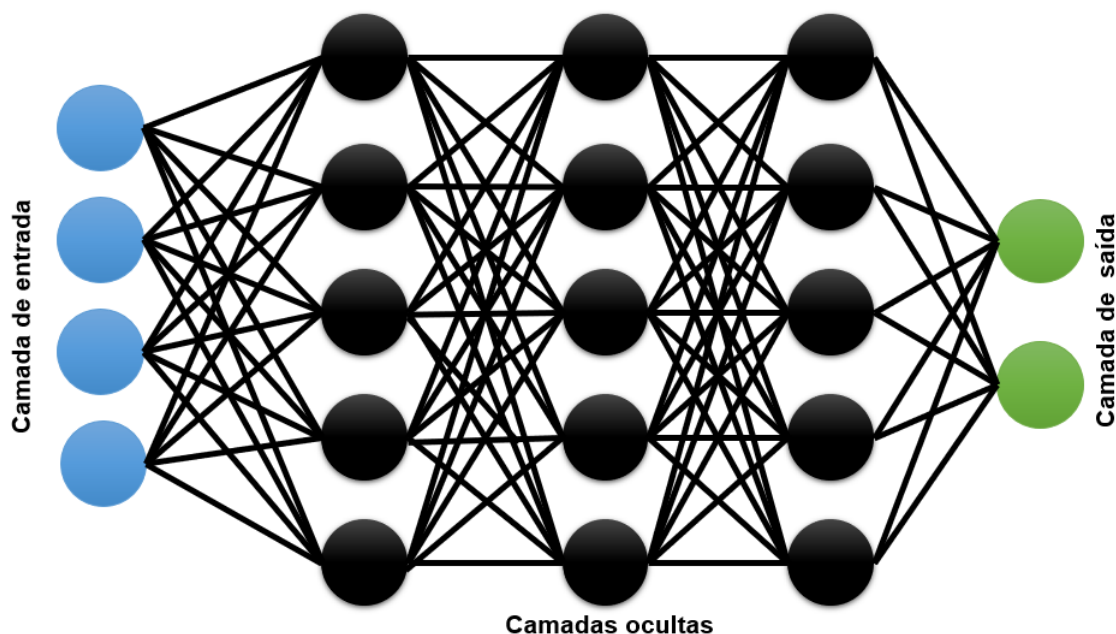
Fonte: Elaborada pelo autor.

A primeira camada é camada de entrada, é onde os dados brutos são recebidos pela rede. As camadas ocultas, também chamadas de intermediárias, são compostas por neurônios que transformam os dados em uma representação abstrata (LeCun *et al.*, 2015). A camada de saída geralmente é responsável pela classificação dos dados de entrada, mas nem sempre é o caso.

3.3.1 – Redes Neurais Convolucionais

As redes neurais convolucionais se enquadram nas técnicas de *Deep Learning*, uma subárea das RNAs, sendo amplamente utilizadas em tarefas de reconhecimento de padrões de imagens, por apresentarem como característica a aptidão superior de aprendizagem de padrões (Maior *et al.*, 2021). As CNNs são comumente utilizadas para realizar o processamento em imagens, para isso, se recorre a operações de convolução para abstrair informações das camadas subsequentes (Schmidhuber, 2015). As camadas ocultas de uma CNN executam principalmente dois tipos diferentes de funções: convolução e agrupamento (Bhadoria *et al.*, 2021). Na Figura 3.2 é apresentado a estrutura de organização em camadas de uma rede neural profunda (Deep Learning).

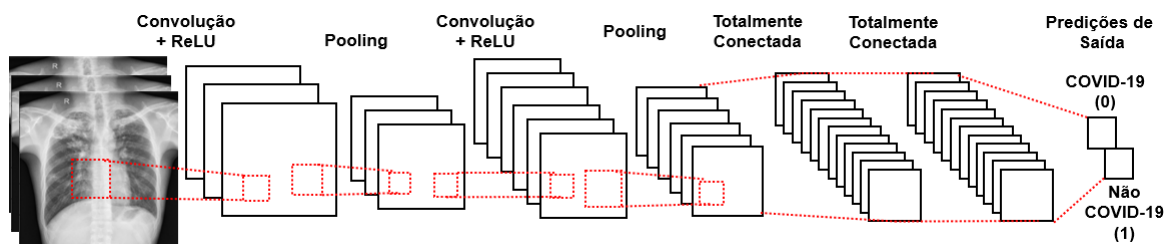
Figura 3.2 – Estrutura de uma rede neural profunda simplificada.



Fonte: Elaborada pelo autor.

Uma CNN é normalmente subdividida em camada convolucional, camada de subamostragem (*pooling*) e camada totalmente conectada (Yu *et al.*, 2021). A organização das redes neurais convolucionais acontece por estágios, cada estágio pode ter uma ou mais camadas de convolução em sequência, que segue normalmente uma camada de sub-mostragem e opcionalmente por uma camada de normalização. Na Figura 3.3 é apresentado a estrutura de uma CNN.

Figura 3.3 – Estrutura de uma CNN.



Fonte: Araújo *et al.* (2017) (com adaptações).

Primeiro, as imagens de entrada são pré-processadas com uma normalização padrão. Os dados então fluem em pares de camadas convolucionais com camadas de *pooling*, onde a extração de características e a redução de redundância são executadas. As características simples se reúnem gradualmente de maneira eficiente. Depois, todas as características são parcialmente combinadas e as características resultantes contam como parte da configuração

da classe do marcador. Por fim, essas características principais são enviados para a camada totalmente conectada, que fornece uma estimativa da classificação (Liu, 2018).

As CNNs têm sido constantemente utilizadas em aplicações que envolvem imagens médicas (Han *et al.*, 2022), (Kathamuthu *et al.*, 2023), (Huang & Wu, 2022), (Hajabdollahi *et al.*, 2020), (Wang *et al.*, 2022), entre outros. Essas imagens possuem um alto nível de complexidade, difícil de compreender (Yu *et al.*, 2021). Devido a tais características, as utilizações de técnicas de aprendizado profundo favorecem a uma melhor compreensão dessas imagens. A utilização dessas técnicas, implica na otimização de tarefas e menor custo de tempo. As CNNs, normalmente realizam tarefas de classificação, detecção e segmentação (Yu *et al.*, 2021).

Camada de convolução

A camada de convolução consiste em extrair um conjunto de características da camada anterior. Segundo Santos (2017), essa camada é constituída por vários neurônios, cada um responsável por aplicar um filtro em uma parte específica da imagem. Um filtro é apenas uma matriz de valores, chamados pesos, treinados para detectar características específicas. O objetivo do filtro é realizar a operação de convolução, sendo um produto elemento-a-elemento e soma entre duas matrizes, equação 3.1 (Abirami & Chitra, 2020). Os valores gerados pela operação de convolução passam por uma função de ativação, normalmente a função Unidade Linear Retificada (*Rectified Linear Unit* — ReLU) (Araújo *et al.*, 2017).

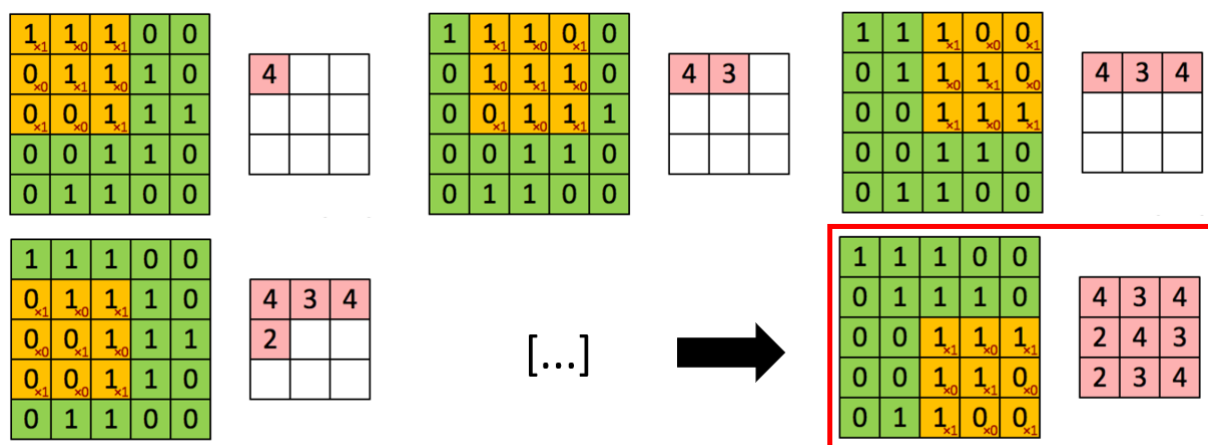
$$A(i, j) = (I * F)(i, j) = \sum_m \sum_n I(m, n)F(i - m, j - n) \quad (3.1)$$

Na Figura 3.4 é ilustrada a aplicação da operação de convolução em uma imagem com dimensões $5 \times 5 \times 1$ com um filtro 3×3 .

A aplicação dos filtros permite realizar a extração de características da entrada, a partir da extração obtêm-se características de baixo nível nas primeiras camadas, e a cada camada novas características de alto nível são extraídas, as informações extraídas são conhecidas como mapas de características (*features*), cada filtro resulta em um mapa de características.

Os filtros percorrem todos os dados da imagem em largura, altura e profundidade que desliza por toda a imagem, aplicando a operação de convolução. Conforme o treinamento da rede, ocorre um ajuste nos filtros, que torna possível identificar características que os lotes de entradas possuem em comum como cores, bordas, arestas, entre outras. À medida que as

Figura 3.4 – Operação de convolução.



Fonte: Elaborada pelo autor.

entradas vão sendo passadas para a rede, os filtros conseguem aprender características mais complexas (Rodrigues, 2018). Isso resulta no reconhecimento de mesmas características ou até mesmo o objeto em várias imagens em locais diferentes (Singh *et al.*, 2021).

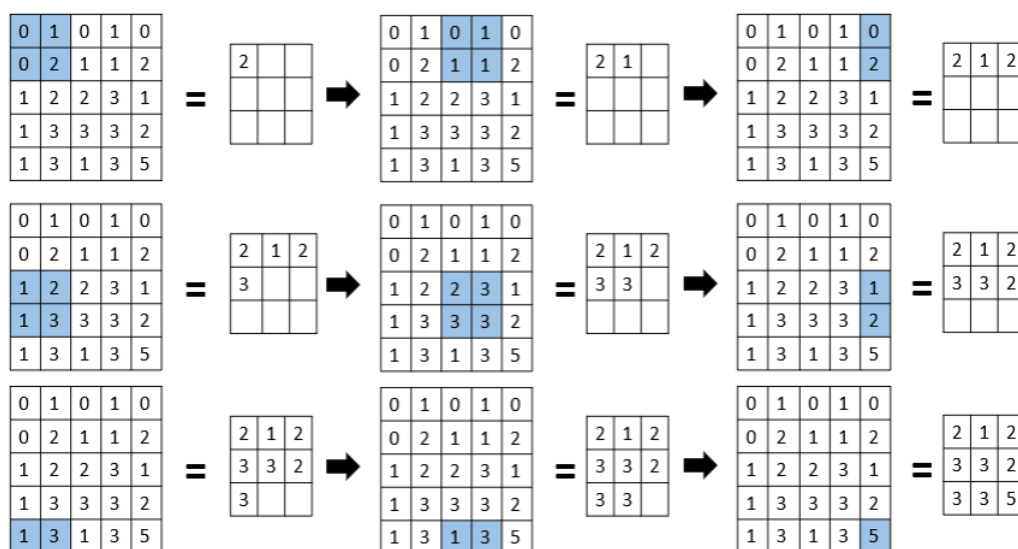
Pooling

Após a camada de convolução é utilizado a camada *pooling*. A camada de *pooling* é normalmente utilizada para construir o mapa de recursos agrupando as informações advindas das camadas anteriores. Esta camada visa reduzir os dados de entrada. Com a redução dos dados de entrada, também há a redução do custo computacional. Além disso, a operação de *pooling* contribui na redução do *overfitting*. *Overfitting* é a memorização dos dados utilizados gerando bons resultados na classificação, mas quando utilizado outro conjunto de dados sua generalização acaba se tornando ineficiente.

Nessa camada é destacado ainda mais as características, nesta etapa é utilizando o método de *max pooling* que é o mais comum, é utilizado um passo para o filtro deslizar, a cada passo é selecionado o *pixel* com valor maior, criando uma matriz de *pixels* com valores maiores. Desta forma são descartados os valores desprezíveis, gerando uma invariância a pequenas mudanças e distorções de locais. Na Figura 3.5 é ilustrado a operação de um *max pooling*.

Camadas totalmente conectadas

Camadas totalmente conectadas são as camadas finais na estrutura da CNN que podem ser uma ou mais camadas e colocadas após uma sequência de camadas de convolução e

Figura 3.5 – Funcionamento do *max pooling* com um filtro 2×2 .

Fonte: Elaborada pelo autor.

pooling (Ranjbar *et al.*, 2020). As camadas totalmente conectadas estão localizadas no final da rede. Após as características serem extraídas, as camadas totalmente conectadas utilizam essas características para produzir a resposta de classificação da rede. A camada totalmente conectada também é chamada camada densa. Todos neurônios da camada anterior está conectado com todos da próxima camada. Os dados nessa camada são processados e produzem os mapas de características convertidos em um vetor de características, que executa o raciocínio de alto nível no modelo desenvolvido (Abirami & Chitra, 2020).

Na camada totalmente conectada, acontece a propagação do sinal por meio da multiplicação ponto a ponto e o uso de uma função de ativação. A imagem de entrada é classificada conforme as suas características extraídas, as camadas totalmente conectadas utilizam dessas características para ter a saída da classificação da rede. A classificação acontece pela probabilidade da imagem pertencer a uma das classes treinadas (Araújo *et al.*, 2017). A atualização e otimização dos pesos do algoritmo são efetuadas através do algoritmo de retropropagação do gradiente do erro (Abirami & Chitra, 2020).

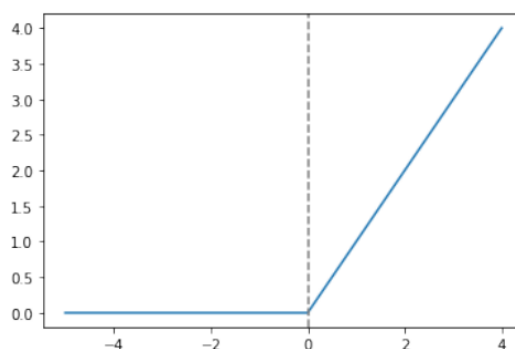
Funções de ativações

Na criação de modelos de redes neurais, as funções de ativação são elementos fundamentais, fornecem a não-linearidade do modelo. Decidindo se um neurônio deve ser ativado ou não conforme a relevância da informação.

A ReLU é a função de ativação bastante utilizada, é normalmente seguida por operações adicionais, como camadas de agrupamento, camadas totalmente conectadas e camadas de normalização. A retro propagação é usada para distribuição de erros e ajuste de peso (Ren *et al.*, 2021). A função ReLU gera valores entre 0 e infinito. Quando os valores são negativos o retorno é 0, quando o valor é positivo o retorno é o próprio valor (Ceccon, 2020). Dessa forma, para ativar o neurônio é preciso que o valor seja maior que zero, isso o faz ser esparsos e eficiente (Academy, 2019). A função ReLU é definida como mostra a Eq. 3.2. Essa função é representada graficamente na Figura 3.6.

$$ReLu(x) = \max\{0, x\} \quad (3.2)$$

Figura 3.6 – Função de ativação ReLu.

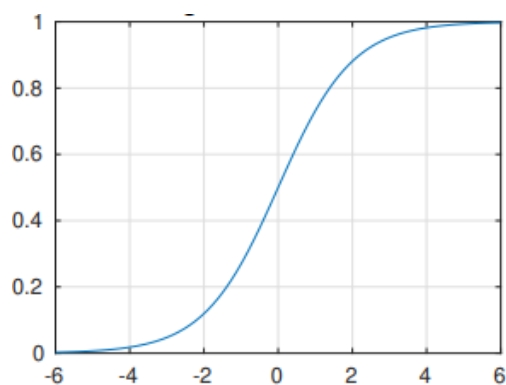


Fonte: (Goldbarg, 2021).

A função de ativação *sigmoid* é uma função de ativação não linear amplamente utilizada. Ela reduz um número real para a faixa entre 0 e 1. Se a entrada for grande, o resultado será próximo a 1, enquanto se a entrada for grande e negativa, o resultado será próximo a zero. Essa função permite explicar a intensidade da ativação dos neurônios quando estimulados, onde 0 significa nenhuma ativação e 1 representa uma ativação completa (Zhang, 2021). A função é definida conforme a Eq. 3.3. Essa função é representada graficamente na Figura 3.7.

$$\sigma(x) = \frac{1}{1 + e^{-x}} \quad (3.3)$$

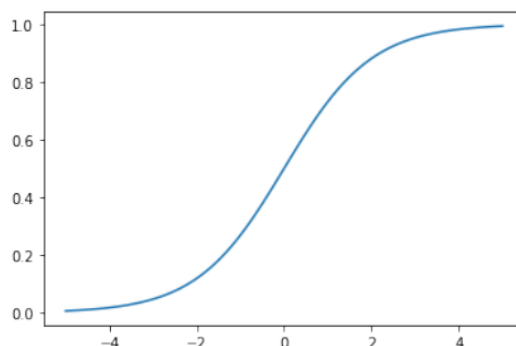
A função de ativação *softmax* também conhecida como função exponencial normalizada, é muito utilizada em problemas de classificação e em inúmeras tarefas, é aplicada na camada de saída do classificador para determinar a qual classe pertence cada entrada. É realizada uma transformação da saída para valores entre 0 e 1 para cada classe, dividido pela soma das saídas, resultando na probabilidade de a informação de entrada estar em uma determinada

Figura 3.7 – Função de ativação *sigmoid*.

Fonte: (Zhang, 2021).

classe (Academy, 2019). A função é definida conforme a Eq. 3.4. Essa função é representada graficamente na Figura 3.8.

$$\sigma(Z)_j = \frac{e^{z_j}}{\sum_{k=1}^K e^{z_k}} \quad \text{para } j = 1, \dots, K. \quad (3.4)$$

Figura 3.8 – Função de ativação *softmax*.

Fonte: (Goldbarg, 2021).

Aumento de dados

O aumento de dados (*data augmentation*) é uma técnica para aumentar as amostras do conjunto de dados de treinamento de maneira artificial a partir de dados existentes. É amplamente utilizada quando o conjunto de dados é pequeno, entretanto, também podem ser utilizados mesmo em conjunto de dados grande. Para gerar novas imagens artificiais são aplicadas técnicas como translação, rotação, modificação da perspectiva, ruídos, achatamento, alongamento, entre outros, aplicado nos dados de entrada na etapa de treinamento. Além disso, o aumento de dados ajuda na redução do *overfitting*.

Adam

O otimizador Adam é um algoritmo usado em modelos de redes neurais para treinar mais rapidamente. Ele combina um método chamado gradiente descendente estocástico com uma técnica que ajusta automaticamente a taxa de aprendizado com base na velocidade e na direção no qual os gradientes estão mudando (Kingma & Ba, 2014).

Dropout

O *dropout* é uma técnica de regularização usada em redes neurais para reduzir o *overfitting*. O *dropout* funciona desligando aleatoriamente um número de unidades de saída em uma camada durante o treinamento. Isso força a rede a aprender representações redundantes e reduz a dependência de características específicas, melhorando assim a capacidade de generalização do modelo (Srivastava *et al.*, 2014).

Flatten

O *flatten* é uma camada usada em redes neurais convolucionais usada para converter a saída das camadas convolucionais em um vetor unidimensional. Essa camada é usada para passar a saída de uma camada convolucional para uma camada densa para a rede poder aprender padrões mais complexos.

3.4 – Segmentação de imagens

A segmentação de imagens consiste em particionar a imagem. Essa subdivisão visa extrair regiões ou objetos conforme o objetivo que se deseja alcançar. Zhao *et al.* (2021) define a segmentação de imagens como o processo de dividir matematicamente uma imagem em regiões com propriedades semelhantes que não se cruzam. É o processo de extrair várias regiões homogêneas de uma imagem com base em suas características como escala de cinza, textura e cor, etc. A segmentação tem um papel fundamental no processamento e análise de imagens, quanto melhor a segmentação, melhor a análise de imagens. Entretanto, Guo & Ashour (2019) define a segmentação de imagens como uma tarefa desafiadora e complexa afetada por vários aspectos, incluindo ruído, baixo contraste, iluminação e irregularidade dos limites do objeto.

A segmentação pode acontecer manualmente, imagem por imagem, como também utilizando algoritmos que conduza o processo automaticamente. Muitos métodos de

segmentação têm sido propostos, a segmentação tem se tornado essencial em diversos campos como reconhecimento de faces, sistemas de controle de tráfego, reconhecimento de impressão digital, visão computacional e imagens médicas (Zhao *et al.*, 2021).

Nas imagens médicas, a segmentação é um dos principais métodos quando se trata processamento de imagens. É aplicada em tarefas como segmentação de lesão de pele (Iranpoor *et al.*, 2020), (Ghofrani *et al.*, 2020), segmentação de tumores cerebrais (Lin *et al.*, 2021), segmentação de lesões mamárias (Liang *et al.*, 2019), entre outros.

Diversas categorias de imagens médicas, como, microscopia, dermatoscopia, radiografia, ultrassom, tomografia computadorizada, ressonância magnética e tomografia por emissão de pósitrons estão sendo implementados por pesquisadores por meio de algoritmos de processamento de imagens médicas (Guo & Ashour, 2019).

Conforme Sharma & Aggarwal (2010) a segmentação consiste em subdividir os objetos em uma imagem. Se tratando de imagens médicas, Sharma & Aggarwal (2010) ressalta que essa categoria de imagens possui padrões de características complexas, tornando uma tarefa complicada. Alguns fatores podem afetar a saída do algoritmo de segmentação, alguns deles são:

- ▣▣▣▣ efeito de volume parcial;
- ▣▣▣▣ intensidade inhomogeneidade;
- ▣▣▣▣ presença de artefatos;
- ▣▣▣▣ proximidade ao nível de cinza de diferentes tecidos moles.

Segundo Wang *et al.* (2021) em imagens médicas a segmentação é fundamental para extrair evidências confiáveis no diagnóstico, análise de doenças, tratamento clínico e pesquisa médica e cirurgia auxiliar. Com a segmentação de imagens médicas, é possível extrair a região de interesse (*Defining Region of Interest — ROI*) com base nas características. Algumas características segmentadas em imagens médicas, são, órgãos, tumores, nódulos, inflamações, manchas, entre outras. A segmentação é um processo essencial na utilização dessa natureza de imagens.

No decorrer dos últimos anos foram desenvolvidos métodos em aprendizado profundo que estão sendo amplamente utilizado no campo da segmentação de imagens, incluindo redes neurais recorrentes (*Recurrent Neural Network — RNN*), redes neurais convolucionais

(CNN) e redes totalmente convolucionais (*Fully Convolutional Networks* — FCN) (Wang *et al.*, 2021). As FNCs são formadas por camadas conectadas localmente, como camadas de convolução, *pooling* e *upsampling*. As FNC têm como características principais os caminhos de *downsampling* e *upsampling*. O *downsampling* captura informações contextuais, enquanto o *upsampling* recupera informações espaciais. As conexões de salto entre camadas são executadas para recuperar informações espaciais de baixa granularidade, potencialmente perdidas nas camadas de *pooling* e *downsampling* (Renard *et al.*, 2020).

Uma arquitetura amplamente utilizada para segmentação de imagens é a U-Net. Atualmente há diversas variações dessa arquitetura. Uma dessas variações é a MultiResUNet que será a rede utilizada nesse trabalho para realizar a segmentação de imagens. Para isso, é preciso explicar a ideia por trás da U-Net. A U-Net é uma arquitetura baseada em FNC. O que difere a U-Net da FNC é que cada escala de *downsampling* está vinculada à escala de *upsampling* correspondente com um operador de concatenação, levando a uma melhor segmentação por conta das informações que a escala de *upsampling* possui que tem da escala de *downsampling* correspondente e da escala de *upsampling* inferior (Renard *et al.*, 2020).

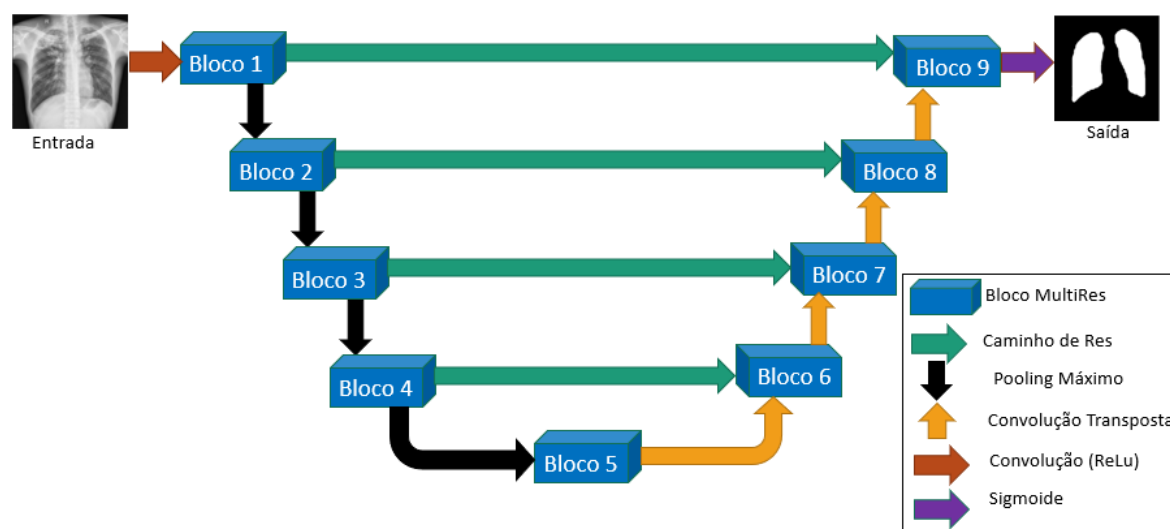
A U-Net possui esse nome por conta do formato de sua arquitetura em forma de U. É uma das arquiteturas mais tradicionais utilizadas na área de segmentação de imagens médicas (Zhang *et al.*, 2020b) (Su *et al.*, 2021). Ela realiza duas tarefas principais, a convolução e convolução transposta. A convolução transposta é um processo para aumentar a resolução espacial da entrada por *upsampling* do *kernel*. O lado da frente do U contém camadas de convolução para *downsampling* e o lado de trás possui camadas de convolução transpostas para *upsampling* (Kusakunniran *et al.*, 2021).

MultiResUNet é uma versão melhorada baseada na UNet proposta por (Ibtehaz & Rahman, 2020). O MultiResUNet também substitui todas as conexões de salto da UNet por caminhos de Res. Quando comparado ao UNet, a arquitetura MultiResUNet possuem uma melhor precisão ao ser utilizada na segmentação de imagens complexas (Chen *et al.*, 2021). A MultiResUNet foi superior que a U-Net em cinco conjuntos de dados (Ibtehaz & Rahman, 2020). Sendo quatro conjuntos de imagens 2D e um conjunto 3D.

A arquitetura MultiResUNet se divide em duas partes: bloco MultiRes e caminho de Res. Na Figura 3.9 é apresentada a arquitetura da MultiResUNet. Nessa arquitetura, a imagem de entrada passa inicialmente por uma convolução. Em seguida, o primeiro bloco MultiRes é iniciado. Do bloco 1 até o bloco 5, acontece a operação de *pooling* após cada bloco. Além

disso, os blocos estão ligados a outros blocos por meio dos caminhos de Res, representados pela seta verde na Figura 3.9, com exceção do bloco 5, que não possui essa ligação com outro bloco. A partir do bloco 5 até o bloco 9, ocorre a operação de convolução transposta em cada bloco. No MultiResUNet, todas as camadas convolucionais, exceto a final, utilizam a função de ativação ReLU (LeCun *et al.*, 2015). Enquanto a camada convolucional final usa uma função de ativação sigmoide (Ibtehaz & Rahman, 2020). Após toda a execução, é obtida a máscara da imagem passada como entrada. O bloco MultiRes e os caminhos de Res são melhor ilustrados nas Figuras 3.10 e 3.11, respectivamente.

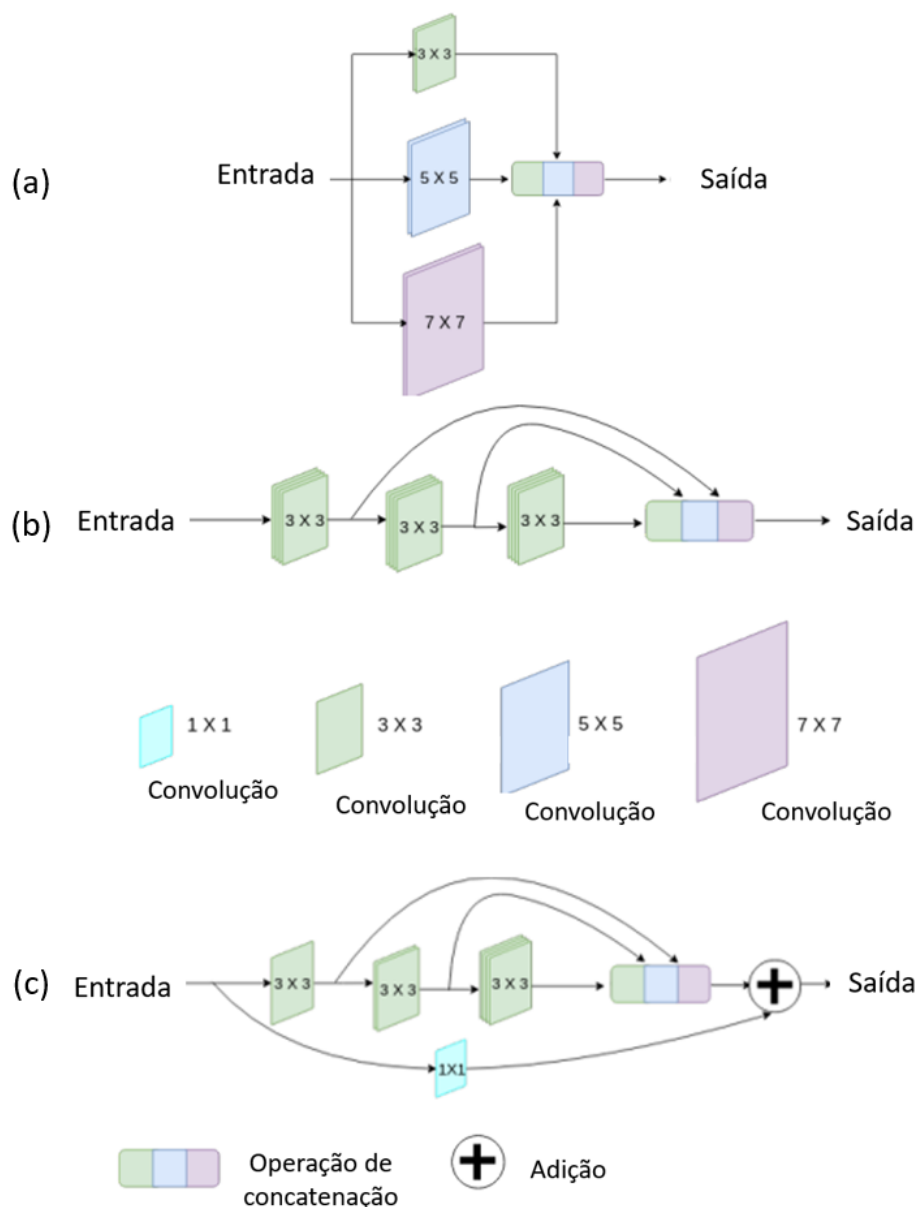
Figura 3.9 – Arquitetura MultiResUNet.



Fonte: (Ibtehaz & Rahman, 2020) (com adaptações).

Conforme Ibtehaz & Rahman (2020) o bloco MultiRes possui três partes: (a) começa com um bloco simples do tipo Inception. Conforme Szegedy *et al.* (2015a) o bloco inception permite usar vários tipos de tamanho de filtro, em vez de ficar restrito a um único tamanho de filtro, em um único bloco de imagem, que então é concatenado e passado para a próxima camada. Foram utilizados filtros convolucionais 3×3 , 5×5 e 7×7 em paralelo, concatenando os mapas de características gerados, permitindo reconciliar características espaciais de diferentes tamanhos de contexto; (b) em vez de usar filtros 3×3 , 5×5 e 7×7 em paralelo, os filtros 5×5 e 7×7 maiores e mais caros são fatorados como uma sucessão de filtros 3×3 ; (c) ilustra que o bloco MultiRes é estabelecido, aumentando o número de filtros nas três camadas sucessivas gradualmente e adicionando uma conexão residual(salto), com filtros 1×1 para conservar as dimensões. o bloco MultiResUNet é ilustrado na Figura 3.10.

Figura 3.10 – Características de um bloco MultiRes.



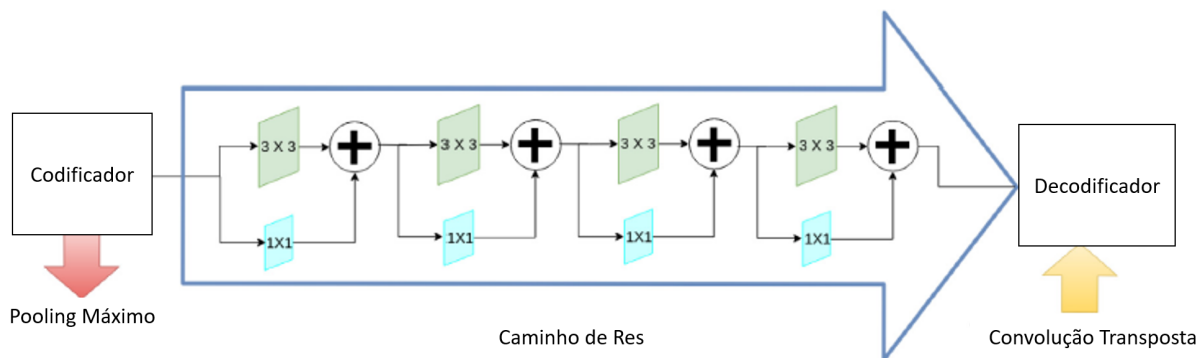
Fonte: (Ibtehaz & Rahman, 2020) (com adaptações).

Os caminhos de Res constituem uma cadeia de camadas convolucionais com conexões residuais, de modo a diminuir a discrepância entre os níveis dos mapas de recursos do codificador e decodificador, mesclados, durante o processo de segmentação (Ibtehaz & Rahman, 2020). O caminho Res é melhor ilustrado na Figura 3.11.

Na MultiResUnet, as características obtidas do codificador não são concatenadas com as características obtidas do decodificador, mas primeiro passam por uma sequência de camadas de convolução 3×3 com conexões residuais e, em seguida, concatenados com o recurso do decodificador (Lou *et al.*, 2019). Segundo DiGregorio *et al.* (2021) as unidades de codificação

se encontram entre as camadas de *pooling* máximo, enquanto as unidades de decodificação se encontram entre as camadas convolucionais transpostas.

Figura 3.11 – Caminho de Res com filtros 1×1 e 3×3 .



Fonte: (Ibtehaz & Rahman, 2020) (com adaptações).

3.4.1 – Segmentação da infecção COVID-19

Métodos de *deep learning* estão sendo propostos visando classificar imagens conforme a patologias, como também busca-se identificar em quais partes do pulmão a infecção COVID-19 está presente.

A segmentação da infecção por COVID-19 em imagens de radiografias de tórax pode ser utilizada para a detecção automática da patologia em uma região pulmonar específica, auxiliando os médicos na identificação da infecção nos pulmões. A identificação da infecção por COVID-19 nos pulmões é útil tanto para o diagnóstico médico quanto para avaliar o grau de deterioração do pulmão causado pela doença.

A segmentação automatizada da região patológica e o diagnóstico de COVID-19 com base em imagens de radiografias de tórax são importantes para um exame mais rápido, onde a segmentação da infecção é uma etapa essencial para avaliação e quantificação (Zhang *et al.*, 2020a).

3.5 – Classificação das imagens de radiografias

Grande parte do sucesso, nas aplicações envolvendo as técnicas de DL, principalmente as CNNs, está relacionado a utilização de grandes conjuntos de treinamento. Entretanto, há tarefas para as quais, esses conjuntos são pequenos ou limitados, permitindo a aplicação dos

conceitos de aprendizagem por transferência. Em outras palavras, modelos pré-treinados em grandes conjuntos de dados podem ser utilizados para aplicações em tarefas semelhantes.

Grande parte do sucesso, nas aplicações envolvendo as técnicas de DL, em especial as CNNs, é a utilização de modelos pré-treinados em grandes conjuntos de dados. Através da aprendizagem por transferência, é possível aproveitar o conhecimento adquirido pelo modelo em tarefas anteriores e aplicá-lo em novas tarefas semelhantes, resultando em um desempenho superior em comparação com o treinamento a partir do zero.

Neste trabalho, os modelos utilizados para realizar as classificações foram: VGG16, VGG19, InceptionV3, DenseNet201, ResNet50, Xception, InceptionResNetV2 e NASNetLarge. Esses modelos vêm da biblioteca Keras que foram pré-treinados no conjunto de dados do ImageNet, conjunto este que possui milhões de imagens e mais de 1000 classes distintas. Para isso é aplicando o conceito de aprendizagem por transferência, utilizando pesos pré-treinados do conjunto de dados ImageNet que é realizando um ajuste fino. Internamente são ativadas com função *ReLU* nas camadas intermediária e *sigmoide* ou *softmax* nas camadas totalmente conectadas, que extraíram recursos que permitiram a classificação.

3.5.1 – VGG16 e VGG19

A VGG é uma arquitetura que utiliza camadas convolucionais de kernel 3×3 empilhadas com profundidade crescente. Após cada pilha de convolução é aplicado o *pooling* máximo (Simonyan & Zisserman, 2014). As VGGs possuem três camadas totalmente conectadas (*Full connected*) às duas primeiras com 4096 neurônios e a última realiza a classificação, com a função *softmax*, que distribui probabilidades que caracterizam as classes. São aplicadas em todas as camadas ocultas a função de ativação ReLu (Silva, 2018). A regularização *dropout* ocorre nas camadas totalmente conectadas (Rezende *et al.*, 2019). A diferença entre a VGG16 e VGG19 diz respeito ao número de camadas, pois a VGG19 possui três camadas convolucionais adicionais, resultando em mais parâmetros. Esses modelos como os outros que foram utilizados neste trabalho estão disponíveis pré-treinados na biblioteca do Keras.

3.5.2 – InceptionV3

A InceptionV3 é uma rede neural convolucional arquitetada por Szegedy *et al.* (2015b) que foi treinada com milhões de imagens para classificar 1000 classes. A arquitetura da

InceptionV3 foi desenvolvida visando melhorar a eficiência computacional e a precisão de classificação em comparação com modelos anteriores, como a InceptionV2 (Szegedy *et al.*, 2014).

Conforme Szegedy *et al.* (2015b) a InceptionV3 consiste em uma série de camadas de convolução, normalização de lote e pooling, intercaladas com camadas de ativação e regularização. A arquitetura da InceptionV3 é composta por duas partes principais: uma série de camadas de "base" que extraem características de baixa e média frequência da imagem de entrada, e uma série de camadas de "topo" que extraem características de alta frequência. As camadas de base são formadas por uma sequência de camadas de convolução e pooling, enquanto as camadas de topo são formadas por uma sequência de camadas de "inception" que combinam diferentes filtros de tamanhos diferentes em paralelo.

A InceptionV3 é uma arquitetura de 48 camadas de profundidade. O foco é em problemas de classificação de imagens Silva (2018). Essa arquitetura possui menor custo computacional, menor número de parâmetros, quando comparada com as VGGs (Szegedy *et al.*, 2015b).

3.5.3 – DenseNet201

A DenseNet201 é uma CNN que possui um padrão de conectividade simples, onde todas as camadas são conectadas diretamente entre si com mapas de características de tamanhos correspondentes, garantindo entre as camadas o fluxo máximo de informações. Cada camada retém entradas adicionais das camadas antecedentes e transita seus respectivos mapas de características para camadas subsequentes. As características são concatenadas antes de serem enviadas para uma camada. Esta rede tem a vantagem de atenuar o problema do esvaecimento do gradiente, além de fortalecer a propagação de características, incentivar o reuso de características e, assim, diminuir a quantidade de parâmetros (Huang *et al.*, 2017).

3.5.4 – ResNet50

A ResNet50 é uma rede do tipo residual de 50 camadas. Conforme Silva (2018) a ResNet50 possui uma estratégia de “atalhos”, a ideia por trás dos atalhos é pular pares de grupos de camadas convolucionais, a rede conforme fica mais profunda, não morrerá de esvaecimento de gradientes. Outra característica importante desta rede é seu grande foco em normalização de lotes. Esta rede detém camadas que se ajustam a um mapeamento residual

explicitamente, ao invés de aguardar que as camadas empilhadas se regulem diretamente a um desejado mapeamento subjacente.

3.5.5 – Xception

É uma CNN inspirada na Inception, onde os módulos Inception foram substituídos por convoluções separáveis em profundidade. A Xception superou a InceptionV3 no conjunto de dados da ImageNet para qual a InceptionV3 foi projetada. A Xception tem o mesmo número de parâmetros que a Inception, entretanto, o ganho da Xception se dá pelo uso mais eficiente dos parâmetros do modelo (Chollet, 2017a).

A base de extração de características da rede Xception é composta por 36 camadas convolucionais. As camadas possuem uma estrutura com 14 módulos, sendo que todos têm em suas proximidades as conexões residuais lineares, com exceção do primeiro e último módulo (Chollet, 2017a).

3.5.6 – InceptionResNetV2

A InceptionResNetV2 é uma CNN que combina as propriedades da Inception e da conexão residual do ResNet, contendo 162 camadas de rede profundas (Albahli & Albattah, 2020). No bloco Inception-Resnet, filtros convolucionais de vários tamanhos são combinados com conexões residuais, essas conexões reduzem o custo de tempo no treinamento e evita o problema de degradação causado por estruturas profundas (Elhamraoui, 2020). As conexões residuais mais profundas permitem atalhos no modelo, o qual pode permitir modelos mais profundos, levando a um melhor desempenho (Szegedy *et al.*, 2017).

3.5.7 – NASNetLarge

A NASNet é uma família de modelos projetados automaticamente, aprendendo as arquiteturas de modelos diretamente no conjunto de dados de interesse. A NASNetLarge é uma CNN treinada em mais de um milhão de imagens do banco de dados ImageNet. Esta rede pode classificar mais de 100 categorias. A rede aprendeu representações de recursos avançados para uma ampla gama de imagens. Mas, é exigido um alto poder computacional (Center, 2019).

3.6 – Aprendizado por transferência

Grande parte do sucesso, nas aplicações envolvendo as técnicas de DL, principalmente as CNNs, está relacionado a utilização de grandes conjuntos de treinamento. O processamento de grandes volumes de dados requer máquinas com grande capacidade computacional, além de um longo tempo para o treinamento. Entretanto, esses recursos podem não estar disponíveis, tornando o processo de treinamento inviável (Rodrigues *et al.*, 2021).

A aprendizagem por transferência é um método de aprendizagem profunda que permite reutilizar o conhecimento aprendido em uma tarefa para realizar outra (Afridi *et al.*, 2018). É possível utilizar os parâmetros de um modelo pré-treinado no conjunto de dados do ImageNet, que contém mais de 1000 classes de imagens. Em outras palavras, modelos pré-treinados em grandes conjuntos de dados podem ser aplicados em tarefas semelhantes com maior agilidade e precisão, por aproveitarem os pesos pré-treinados de um conjunto de dados muito maior, resultando em menor tempo de convergência.

Durante a execução do aprendizado de transferência, é possível treinar todas as camadas da CNN ou somente algumas delas. Em algumas aplicações, a CNN é mantida e apenas a camada de classificação final, também chamada de camada totalmente conectada ou MLP, é re-treinada com os dados específicos da nova tarefa. Isso permite que os pesos da CNN já pré-treinados em grandes conjuntos de dados, como o ImageNet, sejam reutilizados e ajustados para a nova tarefa, resultando em uma convergência mais rápida e melhor desempenho. Já em outras aplicações, é possível re-treinar a CNN parcialmente, ou seja, manter alguns dos pesos pré-treinados e ajustar outros para a nova tarefa, dependendo das necessidades específicas da aplicação. (Chauhan *et al.*, 2021).

3.7 – Explicações agnósticas de modelos interpretáveis locais (LIME)

Explicações agnósticas de modelos interpretáveis locais (Local Interpretable Model-Agnostic Explanations — LIME) é uma técnica para explicação dos modelos de caixa preta, esse método foi criado por (Ribeiro *et al.*, 2016). O LIME gera explicações com base em componentes interpretáveis que são representações dos dados subjacentes compreendidos por pessoas, podendo diferir das entradas do modelo. As características mais importantes do

modelo complexo original são então identificadas e usadas para explicar a previsão do modelo para um exemplo específico. A interpretação local é gerada apenas para o exemplo de entrada específico e não para todo o conjunto de dados (Guidotti *et al.*, 2018).

Conforme Ribeiro *et al.* (2016) o LIME é uma abordagem modelo-agnóstica que pode ser aplicada a qualquer modelo de aprendizado de máquina sem exigir acesso interno ao modelo ou sua implementação. O LIME opera selecionando uma instância de entrada e usando uma função de similaridade para identificar as instâncias de treinamento mais semelhantes a ela. O LIME então ajusta um modelo linear interpretável localmente nas instâncias semelhantes para explicar como o modelo original toma sua decisão. Essa aproximação linear pode ser usada para explicar o resultado da previsão do modelo original para a instância de entrada.

É possível determinar se há presença ou não de componentes interpretáveis por meio do LIME. Dessa forma, pode ser aplicado para segmentação de píxel para classificar imagens mapeando os mesmos (Wörner & Spitzer, 2019).

Materiais e Metodologia

Este trabalho propõe um método em duas etapas para a segmentação de imagens de radiografia do tórax. A primeira etapa envolve a segmentação dos pulmões nas imagens para delimitar a região pulmonar, enquanto a segunda etapa envolve a segmentação da infecção por COVID-19 (área patológica). A segmentação pulmonar e a segmentação da infecção por COVID-19 são processos independentes. Para avaliar a eficácia, foram realizados quatro diferentes testes no treinamento da MultiResUnet para ambas as etapas. No primeiro teste, utilizou-se o filtro CLAHE e não houve aumento de dados. No segundo teste, não foi utilizado o filtro CLAHE nem houve aumento de dados. No terceiro teste, utilizou-se o filtro CLAHE e houve aumento de dados. No quarto teste, não foi utilizado o filtro CLAHE, mas houve aumento de dados.

Na terceira etapa do trabalho, utilizamos as imagens segmentadas na primeira etapa para classificar em duas perspectivas: binária (COVID-19 ou Não COVID-19) e multiclases (COVID-19, Normal ou Não COVID-19). Para avaliar o desempenho e selecionar o melhor modelo de classificação, comparamos o resultado de oito redes neurais convolucionais diferentes.

Na quarta etapa, o melhor modelo de classificação obtido na terceira etapa foi selecionado e utilizado para segmentar a infecção COVID-19 com a técnica LIME. Essa técnica permite a visualização das áreas da imagem que mais contribuíram para a classificação do modelo.

4.1 – Base de Dados

Nesta seção serão brevemente descritas às duas bases de dados utilizadas neste trabalho, sendo elas, a COVID-QU-Ex Datase (Tahir, 2021) e a QaTa-COV19 Dataset (Degerli, 2021). As bases de dados estão disponíveis na plataforma online Kaggle.

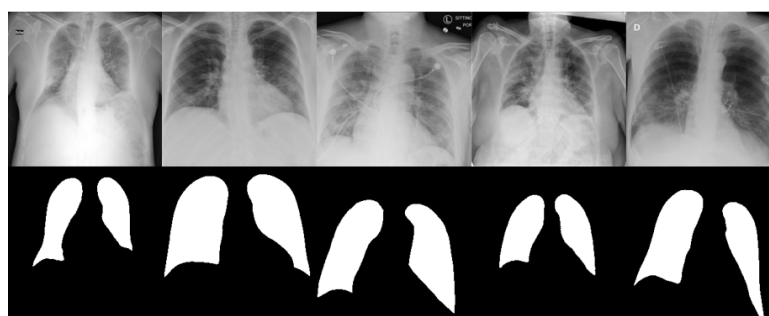
A primeira base de dados será utilizada na etapa de segmentação do pulmão, permitindo a obtenção de imagens com o pulmão segmentado, que serão posteriormente utilizadas na etapa de classificação. Já a segunda base de dados é utilizada na segunda etapa, será empregada para a segmentação da infecção por COVID-19 em imagens de radiografia do tórax, possibilitando a identificação da doença.

4.1.1 – COVID-QU-Ex Dataset

A base de dados COVID-QU-Ex, criada pelos pesquisadores da Universidade do Qatar e Tampere. A base contém 33.920 imagens de RT divididas em três categorias: 11.956 imagens com COVID-19, 11.263 imagens com infecções de Não COVID-19 (pneumonia viral e bacteriana) e 10.701 imagens normais. O conjunto de dados COVID-QU-Ex é composto por imagens de radiografia do tórax coletadas de vários repositórios e estudos (Tahir *et al.*, 2021a), (Tahir *et al.*, 2021b), (Rahman *et al.*, 2021), (Degerli *et al.*, 2021) e (Chowdhury *et al.*, 2020).

A utilização de várias fontes para coleta de imagens garante que o conjunto de dados COVID-QU-Ex seja diverso e representativo de uma ampla gama de casos. As imagens desse conjunto têm as correspondentes máscaras de segmentação do pulmão. As Figuras 4.1, 4.2 e 4.3 apresentam exemplos de imagens das classes utilizadas e suas respectivas máscaras pulmonares que estão disponíveis na COVID-QU-Ex.

Figura 4.1 – Amostras de radiografias da classe COVID-19 com as respectivas máscaras da base COVID-QU-Ex Dataset.

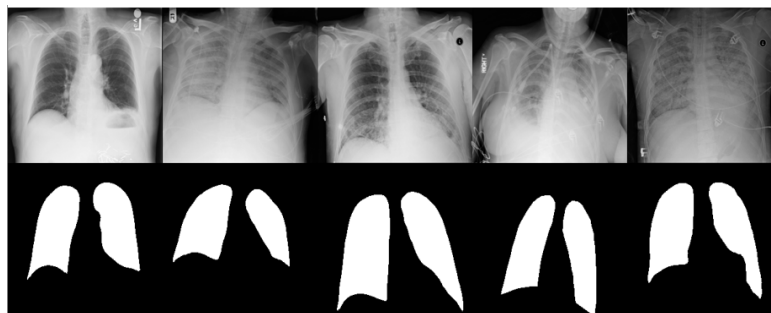


Fonte: Elaborada pelo autor.

4.1.2 – QaTa-COV19 Dataset

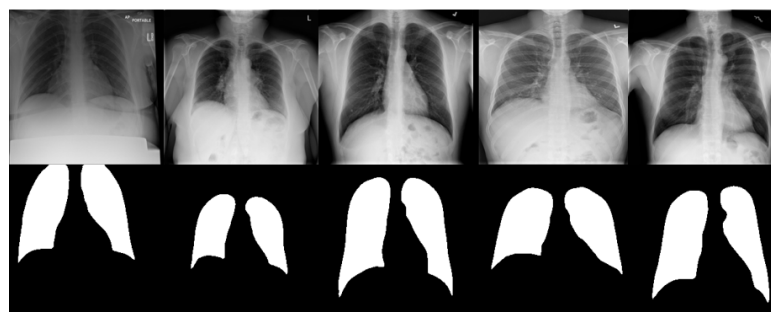
Os pesquisadores da Universidade do Qatar University e da Universidade de Tampere criaram um conjunto de dados de imagens de COVID-19 denominado QaTa-COV19. O QaTa-COV19 é um conjunto de dados que inclui imagens de radiografias de tórax. A base contém

Figura 4.2 – Amostras de radiografias da classe Não COVID-19 com as respectivas máscaras da base COVID-QU-Ex Dataset.



Fonte: Elaborada pelo autor.

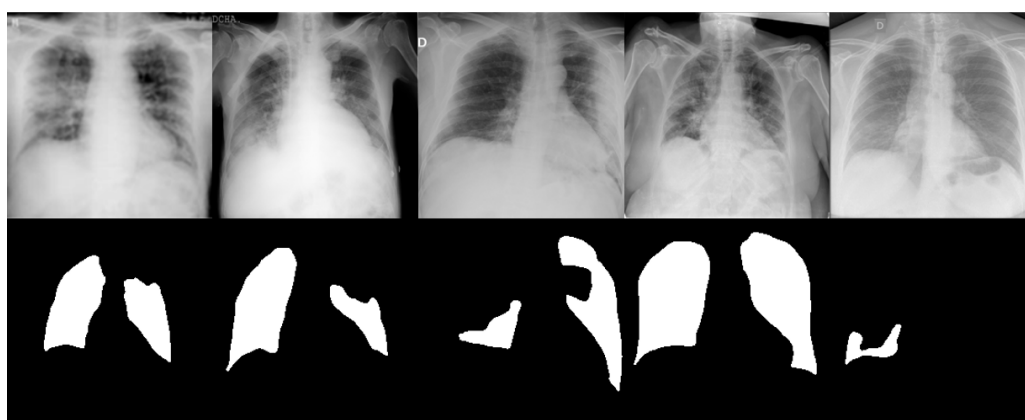
Figura 4.3 – Amostras de radiografias da classe Normal com as respectivas máscaras da base COVID-QU-Ex Dataset.



Fonte: Elaborada pelo autor.

2951 imagens de COVID-19 que possui suas máscaras de segmentação da região patológica pela COVID-19. Essas imagens podem ser usadas para realizar tarefas de segmentação da região infectada pelo vírus. Na Figura 4.4 são apresentadas alguns exemplos de imagens com suas respectivas máscaras.

Figura 4.4 – Amostras de radiografias de COVID-19 com máscaras da base QaTa-COV19 Dataset.



Fonte: Elaborada pelo autor.

Para criar esse conjunto de imagens contendo a segmentação da infecção, o grupo utilizou máscaras desenhadas manualmente e depois submeteram ao processo de aprendizado

de máquina para automatizar o processo e melhorar a qualidade da máscara conforme os refinamentos. Depois, as máscaras criadas pelo algoritmo foram comparadas por médicos para validar a precisão até que as máscaras obtidas sejam boas o suficiente. As melhores máscaras, foram aplicadas aos algoritmos para gerar novas máscaras de forma automática e obtendo um conjunto de dados maior (Degerli *et al.*, 2021). Isso reduziu o esforço de especialistas.

4.2 – Pré-processamento de dados utilizados na segmentação de imagens

O pré-processamento de dados transforma imagens brutas em um formato desejado para análise. Isso melhora a qualidade dos dados e facilita a extração de informações. O pré-processamento pode abranger redimensionamento, normalização, transformação de RGB para escala de cinza, entre outros pré-processamentos.

As imagens do banco COVID-QU-Ex possuem uma resolução de 256×256 *pixels*. Inicialmente, foi realizado um redimensionamento para 224×224 *pixels*. No entanto, devido ao gargalo resultante, foi possível utilizar apenas um pequeno número de imagens. Para contornar essa limitação, foi decidido realizar um novo redimensionamento para 128×128 *pixels*, permitindo o carregamento de mais de 5 mil imagens. As imagens do banco Qa-Ta COV19 também foram redimensionadas de 224×224 *pixels* para 128×128 *pixels*.

As imagens foram convertidas para escala de cinza e normalizadas, e em seguida, foi aplicado a equalização de histograma adaptativo limitada por contraste (*Contrast Limited Adaptive Histogram Equalization* — CLAHE) (Chamseddine *et al.*, 2022).

O filtro CLAHE visa aumentar o contraste dos pequenos ladrilhos da imagem e associar os ladrilhos adjacentes por interpolação bilinear. Este processo remove as bordas induzidas artificialmente. (Chamseddine *et al.*, 2022).

Os parâmetros utilizados foram o *clipLimit* e o *tileGridSize*, são parâmetros da função *createCLAHE* da biblioteca OpenCV. O *clipLimit* especifica o limite de contraste na equalização de histograma e afeta a intensidade do contraste aplicado à imagem. O valor utilizado foi 2.0. O *tileGridSize* especifica o tamanho do *grid* de blocos usados na equalização de histograma. O valor utilizado foi o padrão (8,8).

4.3 – Divisão dos conjuntos de dados nas etapas de segmentação

Na etapa segmentação pulmonar (1ª Etapa), foram utilizadas 2200 imagens de cada classe, sendo elas: Normal, COVID-19 e Não COVID-19. As imagens foram divididas em dois grupos: o conjunto de treinamento, composto por 5280 (80%) das imagens, e o conjunto de teste, composto por 1320 (20%) das imagens.

Na etapa segmentação da infecção COVID-19 (2ª Etapa), utiliza-se apenas imagens de COVID-19 que possui máscara correspondente da infecção COVID-19, totalizando 2950 imagens. O conjunto de imagens foi dividido em dois grupos, sendo 2360 (80%) das imagens destinadas ao conjunto de treinamento e 590 (20%) destinadas ao conjunto de teste.

4.3.1 – Conjunto de validação

Durante o processo de treinamento da MultiResUNet, é utilizada a técnica de validação cruzada. O conjunto de treinamento é dividido em cinco partes, ou *folds*, são gerados índices para acessar os dados em cada *fold*. A partir disso, é criado o conjunto de validação, utilizado para avaliar o desempenho do modelo durante o treinamento. A cada iteração do loop, o modelo é treinado em quatro *folds* do conjunto de treinamento e validado no quinto *fold*, utilizado como conjunto de validação. Esse processo é repetido cinco vezes, de forma que cada *fold* é utilizado como conjunto de validação pelo menos uma vez. O conjunto de teste, por sua vez, é mantido separado e não é utilizado durante o processo de treinamento, sendo reservado para avaliar o desempenho final do modelo em dados não vistos durante o treinamento.

4.4 – Aumento do conjunto de dados de treinamento para segmentação

Na tarefa de segmentação, a utilização de aumento de dados é uma técnica que pode melhorar a capacidade do modelo em generalizar para novos exemplos, evitando problema de *overfitting*, mesmo quando existe uma quantidade maior de imagens disponíveis. Ao aplicar o aumento de dados, é possível aumentar a diversidade da base de dados de treinamento.

Para aumentar o conjunto de dados, foi utilizada a biblioteca disponível no Python chamada *Imgaug*. Essa biblioteca permite aplicar uma série de transformações aleatórias em imagens e máscaras, gerando novas imagens e máscaras a partir das originais. As transformações utilizadas foram: eliminação aleatória (*dropout*), nitidez (*sharpening*), transformação afim (*affine transformation*), transformação elástica (*elastic transformation*), ajuste linear de contraste (*linear contrast*) e desfoque gaussiano (*gaussian blur*). Cada uma dessas transformações possui seus próprios parâmetros, como, por exemplo, o intervalo de probabilidade de *dropout*, o intervalo de intensidade de *sharpening*, entre outros. Todos esses parâmetros são especificados por intervalos, permitindo que cada transformação seja aplicada de forma aleatória dentro desses intervalos. Os aumentos de dados utilizados está listado na Tabela 4.1.

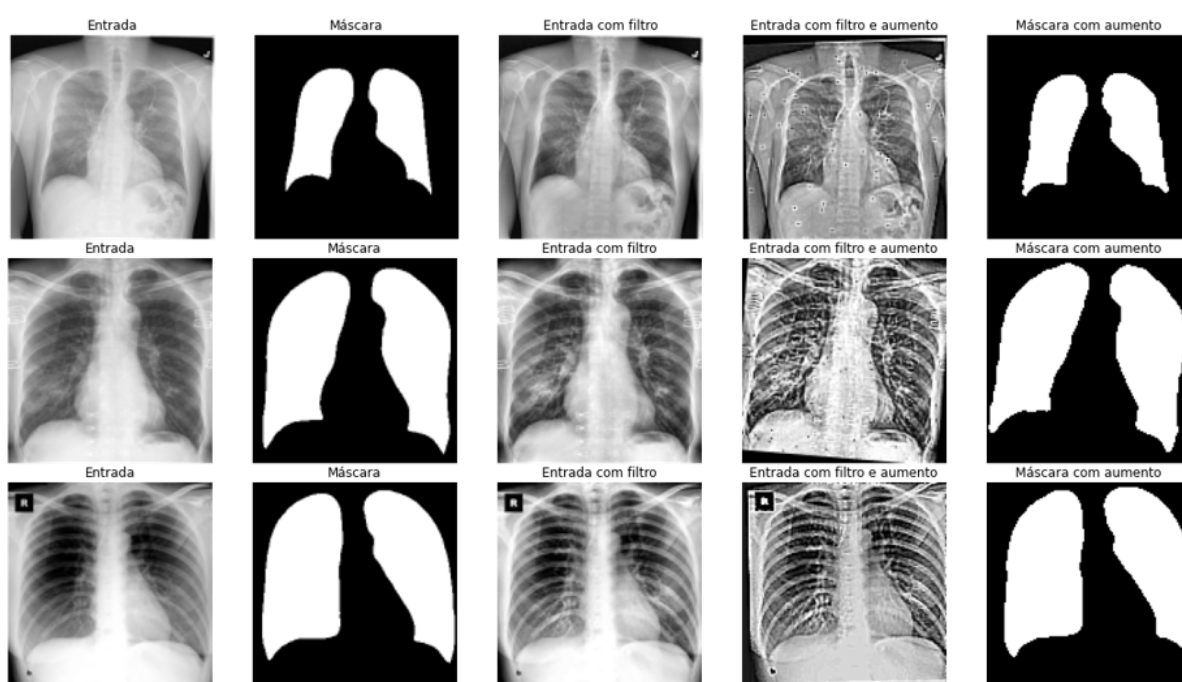
Tabela 4.1 – Parâmetros utilizados no aumento de dados.

Nome	Parâmetros
<i>Dropout</i>	(0,0, 0,004)
<i>Sharpen</i>	(0,0, 1,0)
<i>Translate Percent</i>	x: (-0,05, 0,05), y": (-0,05, 0,05)
<i>Rotate</i>	(-10, 10)
<i>Elastic Transformation</i>	alpha=(0,1), sigma=(0,03)
<i>Linear Contrast</i>	(0,75, 1,5)
<i>Gaussian Blur</i>	sigma=(0, 0,5)

Esses valores são ajustados pelo usuário com base na análise dos dados originais e na escolha do grau de variação desejado. Os valores especificados para cada transformação são os intervalos de variação das suas respectivas configurações. O *Dropout* vai remover *pixels* aleatoriamente com uma probabilidade entre 0% e 0,4%; o *Sharpen* vai aumentar o contraste das bordas da imagem entre 0 e 100%; o *Affine* vai realizar rotações entre -10 e 10 graus, translações aleatórias entre -5% e 5% em ambas as dimensões e escalas aleatórias entre 90% e 110%; o *Elastic Transformation* vai deformar a imagem de forma elástica com um fator de 0,1 e uma variação de 0,03; o *Linear Contrast* ajustará o contraste da imagem entre 75% e 150%; e o *Gaussian Blur* adicionará borrão gaussiano com uma variância aleatória entre 0 e 0,5. Os valores desses aumentos de dados são definidos de forma aleatória conforme as faixas determinadas acima. Foram geradas duas novas imagens com aumento para cada imagem original na segmentação pulmonar (1ª etapa), totalizando 10.560 imagens no conjunto de treinamento. Já na segmentação da infecção (2ª etapa), foram geradas três novas imagens, resultando em um total de 7.080 imagens no conjunto de treinamento.

Essas transformações são aplicadas de forma aleatória e em ordem aleatória à imagem original, gerando novas imagens que formaram o conjunto de dados de treinamento. Mas, não necessariamente todos os aumentos são aplicadas em uma mesma imagem. Isso garante que o conjunto de treinamento seja ampliado de forma diversificada, o que pode auxiliar o modelo a generalizar melhor para dados futuros. As máscaras das imagens originais foram aplicadas às novas imagens geradas para manter a consistência. Na Figura 4.5 é ilustrado exemplos da aplicação do filtro CLAHE e posteriormente o aumento de dados.

Figura 4.5 – Amostras de imagens com aplicação do filtro CLAHE e aumento de dados.



Fonte: Elaborada pelo autor.

A Figura 4.5 representa uma série de cinco imagens dispostas lado a lado, cada uma mostrando uma variação diferente da imagem original. A entrada é a imagem original sem nenhum processamento, ao lado, sua máscara correspondente. A entrada com filtro é a imagem original com um filtro CLAHE aplicado, que aumenta o contraste da imagem. A entrada com filtro e aumento, representa a imagem original com a utilização do filtro CLAHE e também algumas variações introduzidas usando técnicas de aumento de dados. Isso aumenta a diversidade do conjunto de dados disponível para treinar com a MultiResUNet. A máscara com aumento, representa que as máscaras originais receberam aumentos de dados para manter a consistência.

4.5 – Segmentações utilizando a MultiResUNet

Nesta seção, são descritos os passos para realizar a segmentação utilizando a MultiResUNet, seja para a segmentação pulmonar ou da infecção COVID-19. Foram aplicados os mesmos passos para realizar ambas as segmentações, o que diferem são os conjuntos de dados utilizados.

A segmentação da área de interesse das imagens, visa manter as áreas dos pulmões e remover bordas e detalhes decorrentes da presença de ossos e órgãos que normalmente são visíveis em exames de radiografias de tórax.

Para isso segmentar as imagens de radiografias de tórax, foi utilizada uma rede neural artificial denominada MultiResUNet, projetada especificamente para a segmentação de imagens médicas propostas por (Ibtehaz & Rahman, 2020). Em seu trabalho, Ibtehaz & Rahman (2020) testou a MultiResUNet em 5 diferentes conjuntos de imagens médicas, sendo eles: imagens de Microscopia de Fluorescência; imagens de Microscopia Eletrônica; imagens de Dermoscopia; imagens de Endoscopia e imagens de Ressonância Magnética. A MultiResUNet alcançou resultados superiores em todos esses conjuntos na medida de IoU, quando comparado a U-Net.

Para realizar o treinamento da MultiResUNet foram definidos os valores dos parâmetros de forma empírica. Os parâmetros e seus valores são apresentados na Tabela 4.2.

Tabela 4.2 – Configurações do treinamento da MultiResUNet.

Variáveis	Valores
Dimensões das imagens	128 × 128 <i>pixels</i>
Tamanho do batch	32
Função de perda	<i>Binary Crossentropy</i>
Otimizador	Adam
Épocas	50

As dimensões das imagens indicam o tamanho das imagens de entrada para a rede neural. Nesse caso, as imagens são de 128 × 128 *pixels*. O tamanho do *batch* é o número de imagens processadas pelo modelo em uma única iteração do treinamento.

A função de perda é a métrica utilizada para avaliar o quão bem o modelo está se saindo. A função de perda utilizada foi *Binary Crossentropy*, sendo a mesma função de perda utilizada no trabalho original da MultiResUNet (Ibtehaz & Rahman, 2020).

A atualização dos pesos da rede neural durante o treinamento é realizado pelo otimizador. O otimizador utilizado foi o Adam. O Adam ajuda a ajustar os parâmetros do modelo para que ele possa encontrar a melhor solução possível para o problema em questão.

As épocas são o número de vezes que o modelo é treinado com o conjunto completo de dados (Chollet, 2017b).

Para realizar os testes nas etapas de segmentações, seja segmentação pulmonar (Etapa 1^a) ou segmentação da infecção COVID-19 (Etapa 2^a), foi definido 4 diferentes testes para verificar qual tipo de metodologia será mais indicado para realizar a segmentação. A Tabela 4.3 consiste em quatro linhas, cada uma representando um teste diferente. As colunas mostram as configurações utilizadas para cada teste, incluindo se o filtro CLAHE foi utilizado ou não, e se houve aumento de dados. Essa divisão permite comparar facilmente as configurações utilizadas em cada teste e avaliar como cada uma delas pode ter afetado o desempenho da tarefa de segmentação pulmonar e segmentação de infecção.

Tabela 4.3 – Divisão dos 4 testes realizados para realizar a segmentação.

Teste	Filtro CLAHE	Aumento de dados
1	Sim	Não
2	Não	Não
3	Sim	Sim
4	Não	Sim

No primeiro teste, o filtro CLAHE foi utilizado para ajustar o contraste da imagem, sem aumento de dados. No segundo teste, a imagem não foi submetida ao processo de ajuste de contraste pelo filtro CLAHE sendo mantido a quantidade original de dados. No terceiro teste, a imagem foi submetida tanto ao ajuste de contraste, através do filtro CLAHE, quanto a um processo de aumento de dados. Por fim, no quarto teste, a imagem não foi submetida ao ajuste de contraste, porém houve um processo de aumento de dados.

4.5.1 – Treinamento da MultiResUNet

No treinamento da MultiResUNet para a segmentação pulmonar, foram usadas conjuntamente três classes: COVID-19, Normal e Não COVID-19. Cada imagem de cada classe possui uma máscara correspondente da região pulmonar de referência. Já no treinamento da MultiResUNet para a segmentação da infecção por COVID-19, são utilizadas apenas imagens da classe COVID-19, sendo que cada uma possui uma máscara de referência correspondente à região infectada pela doença.

A principal diferença entre a segmentação pulmonar e a segmentação da infecção por COVID-19 é que a primeira visa identificar e segmentar toda a região pulmonar, independentemente de estar infectada ou não pela COVID-19, enquanto a segunda visa detectar

e segmentar apenas a área afetada pela doença. Dessa forma, a segmentação pulmonar pode ser útil para análises mais abrangentes dos pulmões, enquanto a segmentação da infecção por COVID-19 é mais específica e direcionada para a detecção da doença em si.

O processo de treinamento consiste no uso da validação cruzada. A validação cruzada é um método estatístico utilizado para avaliar e comparar algoritmos de aprendizado. Ele consiste em dividir os dados em dois segmentos: um para treinar o modelo e outro para validá-lo (Refaeilzadeh *et al.*, 2009).

A forma básica da validação cruzada é a k-fold, em que os dados são particionados em k dobras de tamanho igual ou quase igual. Durante k iterações de treinamento e validação, uma dobra diferente é mantida para validação enquanto as k-1 dobras restantes são usadas para aprendizado (Refaeilzadeh *et al.*, 2009).

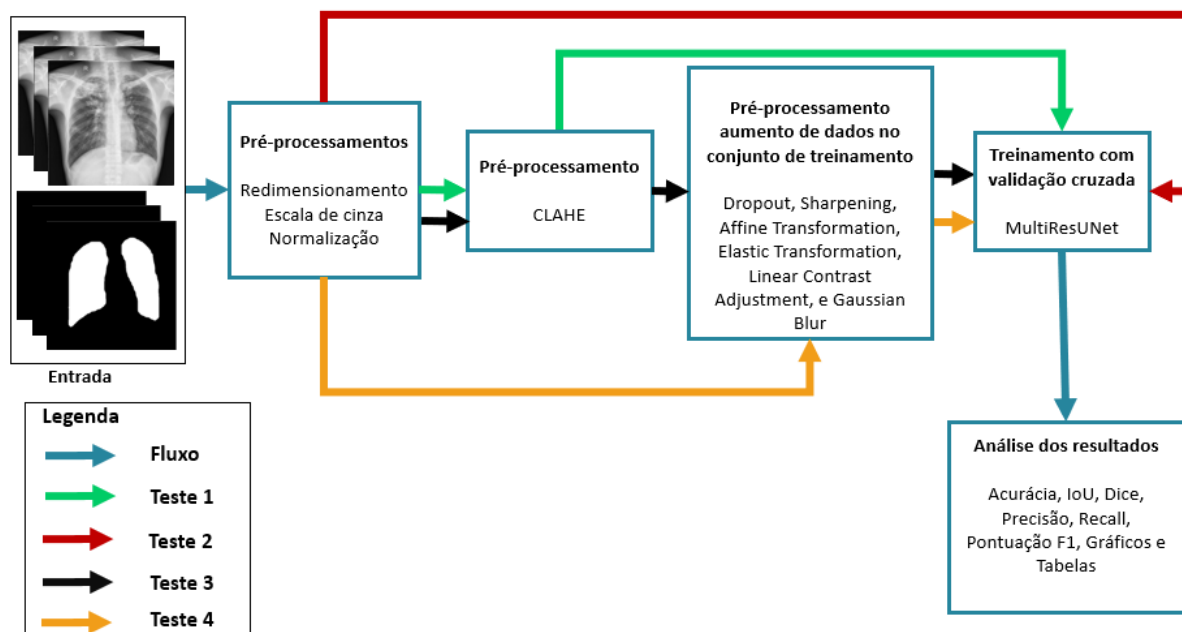
Durante o treinamento, o modelo é avaliado usando o conjunto de validação, o conjunto de treino foi dividido em 5 *folds*, cada *fold* será usado como conjunto de validação uma vez e o restante dos *folds* será usado como conjunto de treino. Isso significa que cerca de 20% do conjunto de treino será usado como conjunto de validação em cada iteração do loop.

Após cada ciclo de validação cruzada com k = 5 iterações, as medidas de desempenho, como acurácia, IoU, Dice, precisão, recall e pontuação F1, são registradas com base no conjunto de teste. Essas medidas são apresentadas na seção 4.8. As médias e desvio padrão dessas medidas são então calculados, permitindo avaliar a capacidade do modelo de generalização, ou seja, sua eficácia em dados novos e desconhecidos. Essas medidas de avaliação são usadas para comparar as previsões do modelo com os dados de referência e avaliar seu desempenho geral. Os passos realizados são ilustrados na Figura 4.6 que exibi como é feito cada passo em cada tipo teste.

A Figura 4.6 apresenta os passos necessários para o treinamento do modelo MultiResUNet obtido em quatro testes diferentes. Cada teste foi identificado com uma seta de cor distinta, sendo a seta azul para os passos básicos e essenciais, seta verde para o Teste 1, seta vermelha para o Teste 2, seta preta para o Teste 3 e seta laranja para o Teste 4.

No Teste 1, identificado pela seta verde, o filtro CLAHE é aplicado, mas não há aumento de dados. Já no Teste 2, identificado pela seta vermelha, não há aplicação do filtro CLAHE e nem aumento de dados. No Teste 3, identificado pela seta preta, há a aplicação do filtro CLAHE e o aumento de dados. Por fim, no Teste 4, identificado pela seta laranja, não há aplicação do filtro CLAHE, mas há aumento de dados. Em todos os testes, os dados são pré-processados, o

Figura 4.6 – Passos para obter o modelo a partir da MultiResUNet em cada teste.



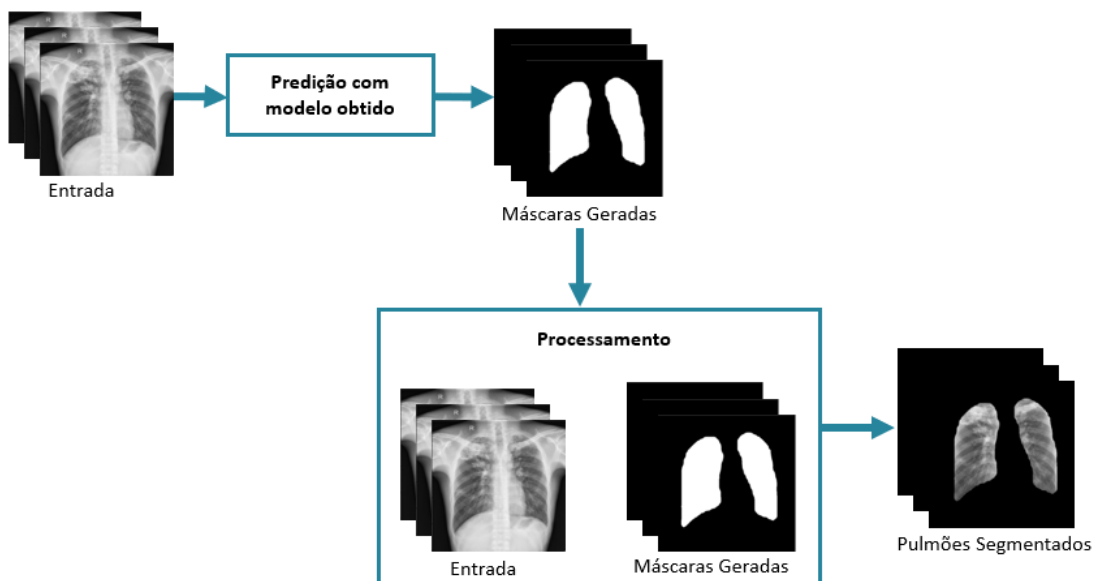
Fonte: Elaborada pelo autor.

modelo é treinado e avaliado com métricas, sendo realizada a análise de resultados para verificar o desempenho do modelo.

O modelo é usado no processo de predição, processo o qual realiza a segmentação pulmonar nas imagens, onde, serão criadas imagens com apenas o pulmão segmentado utilizando como entrada posteriormente na fase de classificação da COVID-19, Não COVID-19 e Normal. Esse processo é ilustrado na Figura 4.7 apresenta o processo de segmentação de pulmões a partir de imagens de entrada utilizando o modelo MultiResUNet treinado anteriormente.

O processo começa com a inserção das imagens de entrada. Essas imagens serão processadas pelo modelo previamente treinado para gerar máscaras correspondentes a cada imagem de entrada. Essas máscaras são então combinadas com as imagens de entrada e submetidas a um processamento adicional para obter a segmentação dos pulmões.

Esse processo resulta em imagens segmentadas dos pulmões, que podem ser utilizadas para uma variedade de finalidades, incluindo diagnósticos médicos e análises clínicas. A utilização do modelo permite uma segmentação de imagens que não possui máscaras referentes, sendo possível criá-las a partir do modelo obtido.

Figura 4.7 – Passos para obter imagens segmentadas partir do modelo MultiResUNet obtido.

Fonte: Elaborada pelo autor.

4.6 – Classificação das imagens de radiografias de tórax

A classificação das imagens (3ª Etapa) consiste em treinar um modelo de aprendizado de máquina para reconhecer padrões nas imagens e classificá-las corretamente em uma das duas classes (binária) ou três, ou mais classes (multiclasses). Para isso, é necessário utilizar técnicas de aprendizado de máquina, como redes neurais convolucionais (CNNs), que conseguem identificar padrões em imagens e extrair características relevantes para a classificação.

4.6.1 – Pré-processamento de dados utilizados na classificação de imagens

O processo de classificação de imagens, utilizou as imagens do conjunto de dados COVID-QU-Ex, que contém as classes de COVID-19, Não COVID-19 e Normal. Foram utilizadas 2450 imagens por classe, totalizando 7350 imagens. Foi realizado a segmentação pulmonar com a MultiResUNet (1ª Etapa) que separou a região pulmonar. Essa segmentação é importante para melhorar a precisão da classificação, uma vez que a presença de outras estruturas na imagem pode interferir na classificação. As imagens foram carregadas na resolução de 128×128 pixels. Além disso, as imagens foram normalizadas por escala de recursos, cada pixel de uma imagem é dividido pelo valor máximo que ele pode ter (255, no caso de imagens RGB).

Classificação Binária

Na classificação binária, foram utilizadas as classes COVID-19 e Não COVID-19 na classificação binária, sem considerar a classe normal como uma opção de classificação de Não COVID-19. Foram usadas 4900 imagens, sendo 3920 (80%) para treinamento e 980 (20%) para teste. A classificação binária serve para determinar se uma imagem é positiva ou negativa para uma doença específica.

Classificação Multiclasse

Na classificação multiclasse, foram utilizadas as três classes: COVID-19, Normal e Não COVID-19. Foram utilizadas todas as 7350 imagens segmentadas. a mesma divisão do passo anterior foi aplicada, 5880 (80%) das imagens foram utilizadas para treinamento e 1470 (20%) para teste. Nesse tipo de classificação, o objetivo é determinar em qual das três classes a imagem se enquadra. Os testes realizados são apresentados na Tabela 4.4.

Tabela 4.4 – Divisão dos testes realizados.

Teste	Classificação	Imagem de entrada	Qtd Treino	Qtd Teste
1	COVID-19 e Não COVID-19	Segmentada	3920	980
2	COVID-19, Normal e Não COVID-19	Segmentada	5880	1470

4.6.2 – CNNs utilizadas

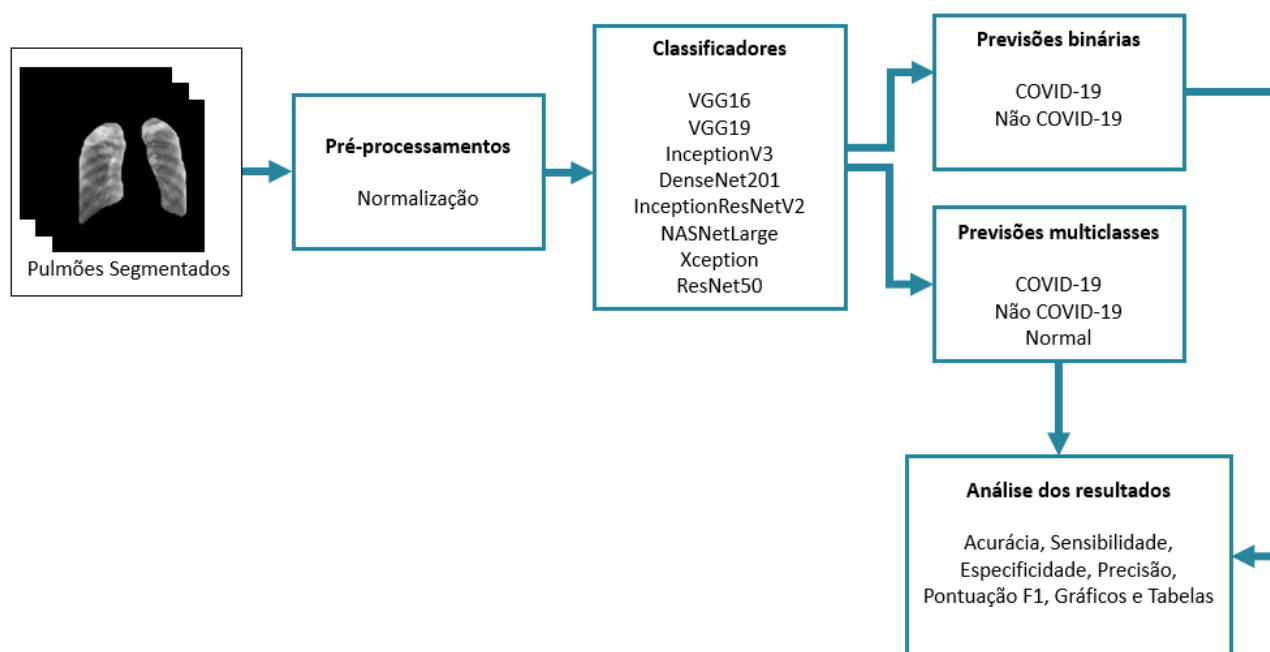
As CNNs utilizadas são: VGG16, VGG19, InceptionV3, DenseNet201, InceptionResNetV2, NASNetLarge, Xception e ResNet50. São redes neurais convolucionais amplamente utilizadas em tarefas de classificação de imagens.

A Figura 4.8 apresenta os passos para classificar as imagens segmentadas a partir dos classificadores.

4.6.3 – Construção do modelo CNN

As redes utilizadas foram a VGG16, VGG19, InceptionV3, DenseNet201, ResNet50, Xception, InceptionResNetV2 e NASNetLarge. Foi utilizado o conceito de transferência de aprendizado, que aproveita modelos pré-treinados no conjunto de dados ImageNet. Após inicializar as camadas da rede com os pesos pré-treinados da rede ImageNet, as camadas são congeladas, exceto as camadas adicionadas para o novo conjunto de dados. Essas camadas incluem uma camada de *pooling* global médio, responsável por reduzir a dimensionalidade do

Figura 4.8 – Passos para classificar as imagens.



Fonte: Elaborada pelo autor.

tensor de saída da rede pré-treinada, uma camada *flatten*, que transforma o tensor de saída em um vetor 1D, e camadas densas totalmente conectadas com ativação ReLU e *dropout*. Essas camadas densas aplicam operações lineares e não lineares para produzir a saída final da rede. São compostas por uma camada com 64 unidades e outra com 3 unidades (para classificação em três classes) ou 2 unidades (para classificação em duas classes), ambas com *dropout* de 0,5, que ajudam a prevenir o *overfitting*. Por fim, é adicionada a camada de saída com ativação sigmoide para duas classes ou *softmax* para três classes.

O modelo foi compilado utilizando Adam como otimizador com uma taxa de aprendizado de 0,0001, o número de épocas igual a 100, tamanho do *batch* igual a 128. Esses valores foram definidos de forma empírica. A função de perda utilizada foi *binary crossentropy* para a classificação binária e *categorical crossentropy* para classificação multiclases.

4.6.4 – Treinamento das CNNs para classificação de imagens

Para treinar os modelos, utilizou-se o *ImageDataGenerator* que é uma classe da biblioteca Keras. Essa instância foi utilizada para aumentar o conjunto de dados de treinamento. Foi configurado para aplicar uma rotação aleatória de até 15 graus no sentido horário ou anti-horário em cada imagem de entrada e preencher os *pixels* de borda com os *pixels* mais próximos.

O tamanho do lote utilizado foi de 128 imagens, o que significa que o gerador de dados produzirá lotes de 128 imagens por vez. No processo de validação cruzada com $k=5$, o conjunto de dados de treinamento foi dividido em cinco *folds*. Em cada iteração, um *fold* foi usado como conjunto de validação e os outros quatro foram combinados em um subconjunto de treinamento parcial. O gerador de dados foi criado com base no subconjunto de treinamento parcial em cada iteração.

As métricas de acurácia, sensibilidade, especificidade, precisão, sensibilidade e F1 foram calculadas conforme o conjunto de teste e os valores foram armazenados. Ao final de todas as iterações, a média e o desvio padrão das métricas acumuladas foram calculadas.

4.7 – Segmentação da infecção causada pela COVID-19 usando o LIME

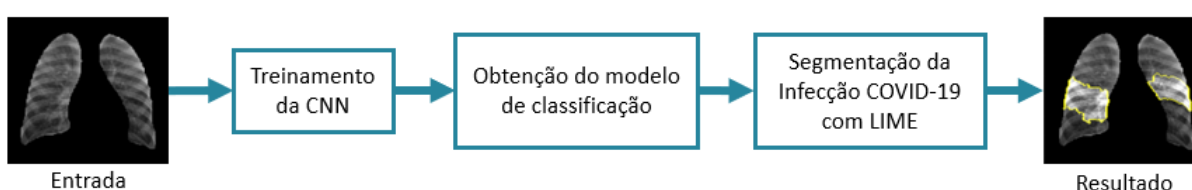
O LIME é uma técnica utilizada para explicar as previsões de modelos de aprendizado de máquina de forma local e interpretável (Ribeiro *et al.*, 2016). Em outras palavras, ele nos ajuda a entender como um modelo de aprendizado de máquina chegou a uma determinada previsão para uma entrada específica. Isso é importante porque muitas vezes os modelos de aprendizado de máquina podem ser considerados caixas-pretas, ou seja, é difícil entender como eles chegam a suas previsões.

O LIME funciona identificando quais características da entrada (imagens de radiografia do tórax) foram mais importantes para a previsão do modelo de classificação obtido. Ele faz isso criando uma explicação local, ou seja, uma explicação para uma entrada específica. Para isso, o LIME utiliza uma função de similaridade para encontrar outras entradas semelhantes à entrada específica e ajusta um modelo linear interpretável localmente nessas entradas para explicar como o modelo original chegou a sua previsão. Essa explicação linear pode ser usada para explicar o resultado da previsão do modelo original para a entrada específica (Ribeiro *et al.*, 2016).

Neste trabalho, o LIME é utilizado para identificar as regiões patológicas causadas pela infecção de COVID-19 nas imagens de raios-X do tórax dos pacientes. Ele faz isso segmentando a imagem com base nas áreas que mais influenciaram a decisão do modelo de classificação obtido. Isso permite uma identificação mais clara das áreas que tiveram maior impacto na decisão do modelo, portanto, pode ajudar a identificar as regiões patológicas.

Dessa forma, ao aplicar o LIME no modelo de classificação treinado com um conjunto de imagens de treinamento, é possível gerar segmentações interpretáveis das imagens de teste e identificar as áreas que mais influenciaram a decisão do modelo de classificação, servindo como segmentação. Isso pode servir para identificar as regiões patológicas causadas pela infecção de COVID-19, portanto, pode auxiliar no diagnóstico e tratamento da doença. O uso da técnica de segmentação LIME é ilustrado na Figura 4.9, onde são apresentados os passos para a obtenção da segmentação da infecção da COVID-19 utilizando o LIME.

Figura 4.9 – Segmentação da infecção COVID-19 com o LIME.



Fonte: Elaborada pelo autor.

4.8 – Medidas de desempenho para avaliar os modelos

Para avaliar o desempenho dos modelos desenvolvidos para segmentação e classificação de imagens médicas, foram aplicadas medidas específicas. Para a segmentação pulmonar e segmentação da infecção COVID-19, foram utilizadas medidas como acurácia, Intersecção sobre União (*Intersection-Over-Union* — IoU) também conhecida como Índice de Jaccard, coeficiente de similaridade de dados (*Dice Coefficient*), precisão, recall e pontuação F1.

Na classificação de imagens médicas, foram aplicadas as medidas de acurácia, sensibilidade, especificidade, precisão e pontuação F1. Se tratando de imagens médicas, as medidas de desempenho de sensibilidade e especificidade são critérios utilizados para mensurar a classificação dessas imagens (Wang *et al.*, 2019). Na segmentação quanto na classificação, os valores de Verdadeiro Positivo (VP), Verdadeiro Negativo (VN), Falso Positivo (FP) e Falso Negativo (FN) são usados para calcular as medidas para avaliar a qualidade dos resultados. Na segmentação, esses valores são definidas para cada *pixel* da imagem, enquanto na classificação elas se referem ao número de amostras de cada classe. No processo de segmentação, as classes positivas e negativas correspondem às regiões do pulmão e do fundo, respectivamente, enquanto na classificação elas correspondem às diferentes categorias a serem identificadas na imagem médica. O cálculo dessas medidas permite avaliar o desempenho do modelo em relação à

capacidade de detectar as áreas de interesse, bem como a capacidade de distinguir diferentes categorias de imagem com precisão.

IoU

O IoU é um método amplamente utilizado na segmentação semântica, o valor varia entre 0 e 1, onde valores mais próximos de 1 indicam uma maior similaridade entre os conjuntos. Ele é considerado sensível às diferenças entre os elementos dos conjuntos, o que significa que pequenas mudanças nos conjuntos podem ter um impacto significativo no valor de IoU. Esse método permite quantificar a precisão da segmentação entre a máscara de referência e a saída do modelo (Jordan, 2018).

Esta medida verifica o quanto a máscara de saída é semelhante à máscara correta. O IoU é definida como o tamanho da interseção dividido pelo tamanho da união dos conjuntos de amostra (DeepAI, 2020). O IoU é definido como:

$$IoU = \frac{|A \cap B|}{|A \cup B|}, \quad (4.1)$$

em que

- ▣ A e B são os conjuntos a serem comparados;
- ▣ $|A \cap B|$ representa a interseção entre a máscara de referência e a máscara obtida, ou seja, a quantidade de elementos em comum entre as duas;
- ▣ $|A \cup B|$ representa a união entre a máscara de referência e a máscara obtida, ou seja, a quantidade total de elementos de A e B, incluindo as repetições.

Coefficiente Dice

O coeficiente Dice (Dice) é um método similar ao IoU para validação de imagens de segmentação geradas por inteligência artificial. O coeficiente Dice varia de 0 a 1, onde valores mais próximos de 1 indicam uma maior similaridade entre os conjuntos. Ele é considerado menos sensível às diferenças entre os elementos dos conjuntos do que o IoU, o que significa que pequenas mudanças nos conjuntos podem ter um impacto menor no valor do coeficiente Dice. O coeficiente Dice é definido como:

$$Dice = \frac{2|A \cap B|}{|A| + |B|}, \quad (4.2)$$

em que

- ▣ A e B são os conjuntos a serem comparados;
- ▣ $|A \cap B|$ representa a interseção entre a máscara de referência e a máscara obtida, ou seja, a quantidade de elementos em comum entre as duas;
- ▣ $|A|$ representa o número total de elementos na máscara de referência;
- ▣ $|B|$ representa o número total de elementos na máscara obtida.

Acurácia

A acurácia é uma medida comum na segmentação e classificação de imagens. Na segmentação, é a proporção de *pixels* corretamente identificados em comparação com a imagem de referência. Na classificação, é a porcentagem de previsões corretas em relação ao total de exemplos. A acurácia é definido como:

$$\text{Acurácia} = \frac{VP + VN}{VP + VN + FP + FN}, \quad (4.3)$$

Precisão

A precisão em segmentação de imagens mede a capacidade do modelo de identificar corretamente os *pixels* positivos. Na classificação de imagens, a precisão é uma medida que expressa a porcentagem de previsões corretas de um modelo em relação ao total de previsões positivas feitas, podendo ser utilizada para avaliar a capacidade do modelo em evitar falso positivos. A precisão é definida como:

$$\text{Precisão} = \frac{VP}{VP + FP}, \quad (4.4)$$

Recall

Recall, também conhecido como sensibilidade, é a medida para medir a capacidade do modelo de detectar corretamente todas as regiões de interesse na imagem, indicando quantos *pixels* da região de interesse foram corretamente identificados. Na classificação, a sensibilidade é a proporção de identificar corretamente os casos positivos, ou seja, a capacidade do modelo em detectar a presença de uma determinada classe. Quanto maior for

o valor resultante da sensibilidade, maior o grau de confiabilidade das imagens classificadas corretamente pelo modelo, ou seja, menor a taxa de diagnósticos errados. Sendo a mais relevante das medidas de desempenho no contexto de classificação de imagens médicas, seguida pela especificidade (Wang *et al.*, 2019). O recall é definido:

$$\text{Recall} = \frac{VP}{VP + FN} \quad (4.5)$$

Pontuação F1

A pontuação F1 (F1) é uma métrica que combina as métricas de precisão e recall para avaliar a qualidade da segmentação em comparação com a região de interesse na imagem de referência. Na classificação de imagens, a pontuação F1 (F1) é uma medida que combina precisão e sensibilidade em uma única medida, permitindo avaliar o desempenho do modelo de forma mais completa. A fórmula para o cálculo da pontuação F1 é dada por:

$$F1 = 2 \cdot \frac{\text{Precisão} \cdot \text{Recall}}{\text{Precisão} + \text{Recall}} \quad (4.6)$$

Especificidade

A especificidade se trata da capacidade identificar caso não positivo corretamente. Quanto maior for a especificidade, maior será a quantidade de imagens de casos não positivos classificados corretamente, ou seja, menor será a taxa de diagnósticos classificados de forma incorreta (Wang *et al.*, 2019). A especificidade é definida:

$$\text{Especificidade} = \frac{VN}{VN + FP} \quad (4.7)$$

Sobre essas medidas, quando se trata de segmentação, o IoU e o coeficiente Dice são mais importantes na avaliação da qualidade da segmentação de imagens porque são mais adequados para avaliar a similaridade entre duas regiões. Mas demais medidas fornecem informações diferentes sobre o desempenho do modelo, o que pode servir para auxiliar na avaliação o desempenho do modelo.

Matriz de Confusão

A matriz de confusão é uma forma de interpretação visual de quão bem os modelos classificou o conjunto de dados de validação individualmente. As matrizes de confusão consistem em uma classe real, que representa as classes das imagens que neste trabalho são, COVID-19, Não COVID-19 e Normal, enquanto, a classe prevista representa as classes classificadas. A matriz de confusão consiste em um método para verificar e avaliar a classificação, quantificando os valores de verdadeiro positivo, verdadeiro negativo, falso positivo e falso negativo.

Os resultados são apresentados em quatro partes: segmentação pulmonar com a MultiResUNet (Seção 5.1), Classificação binária das imagens de radiografias de tórax (COVID-19 e Não COVID-19) (Seção 5.2), Classificação multiclases das imagens de radiografias de tórax (COVID-19, Normal e Não COVID-19) (Seção 5.3) e Segmentação da infecção causada pela COVID-19 (Seção 5.4).

5.1 – Segmentação pulmonar com a MultiResUNet

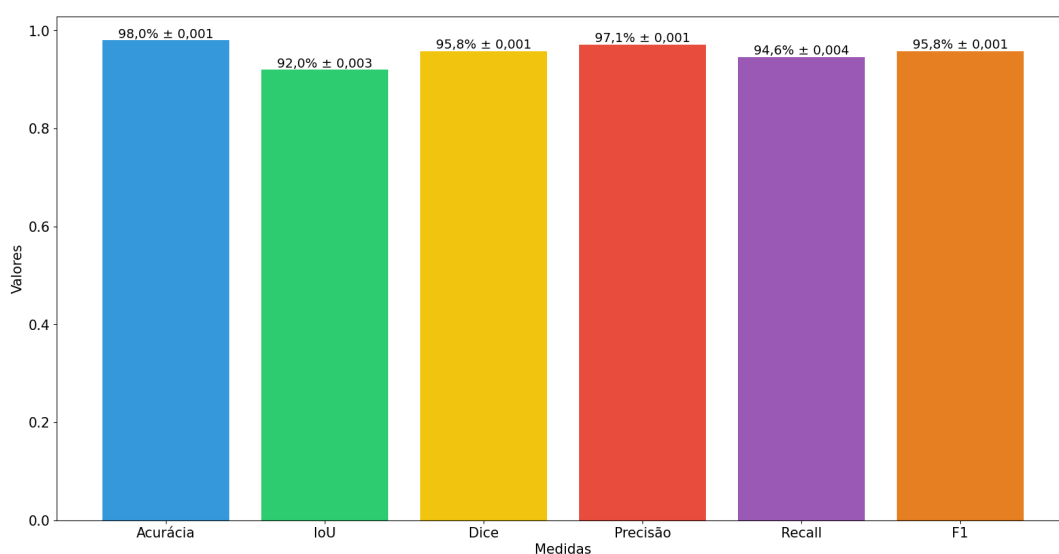
Nesta subseção, são apresentados os resultados alcançados a partir do modelo MultiResUNet. O modelo obtido consegue gerar novas máscaras a partir de imagens desconhecidas. Para realizar o treinamento da rede foi, utilizado imagens dos pulmões e suas respectivas máscaras. Foi usado a técnica de validação cruzada $k - fold$ conhecida por validação cruzada. Nesta técnica é dividido o conjunto de dados em um determinado número de *folds* (também chamados de “partições”). Cada *fold* é composto de dados de treinamento e validação. Os dados de treinamento são usados para treinar o modelo, enquanto os dados de validação são usados para validar o processo de treinamento. Dessa forma, o modelo é treinado e validado, cada vez com um conjunto diferente de dados, isso fornece uma estimativa mais precisa do desempenho do modelo no mundo real.

A avaliação do modelo após cada *fold* da MultiResUNET foi realizada utilizando um conjunto de imagens de teste, que não foram utilizadas durante o treinamento, ou seja, é um conjunto não visto pela rede em nenhuma etapa anterior. O desempenho final do modelo foi avaliado comparando as saídas produzidas pelo modelo com as máscaras consideradas de referência. A precisão do modelo foi calculada através da porcentagem de píxeis corretamente classificados, utilizando o IoU e o Dice. Além disso, diversas medidas foram utilizadas. As medidas de desempenho são calculadas com base nas previsões do modelo e nas saídas reais

do conjunto de teste. As medidas como todo incluem: acurácia, Índice de Jaccard (IoU), coeficiente Dice (Dice), precisão, recall e pontuação f1 (F1).

Para realizar a segmentação, foram testados 4 testes diferentes, na Figura 4.6 observa-se cada teste. Em todos os testes, foram realizados com 5 *fold* e ao final calculado a média de cada medida de desempenho bem como o seu desvio padrão. No primeiro teste, foi utilizado a aplicação do filtro CLAHE, mas não foi utilizado o aumento de dados. A Figura 5.1 apresenta as médias obtidas pelo Teste 1.

Figura 5.1 – Médias das medidas de avaliação na segmentação pulmonar (Teste 1).

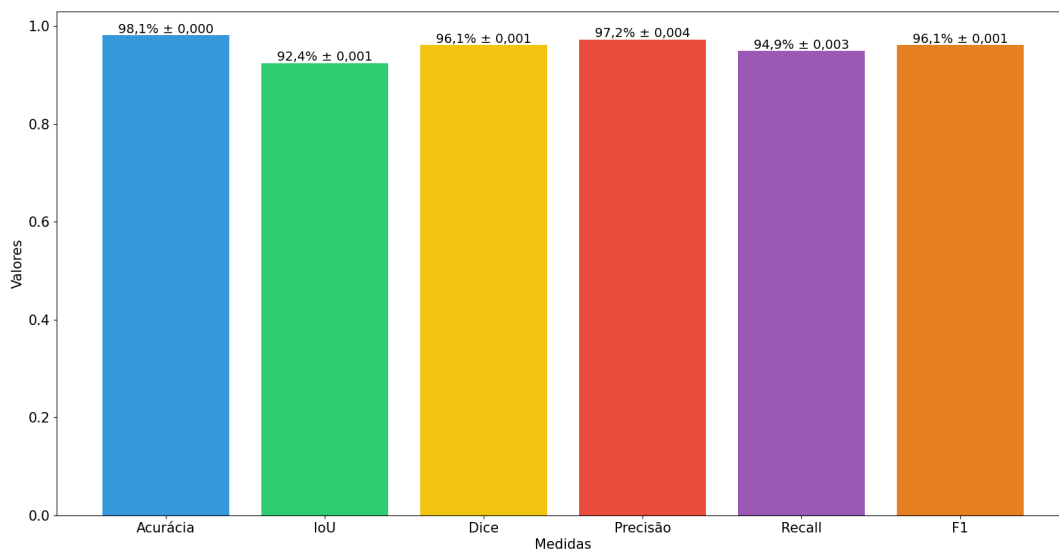


Fonte: Elaborada pelo autor.

A acurácia do modelo foi de 98%, com uma variação de $\pm 0,001$. O valor do IoU foi de 92%, com uma variação de $\pm 0,003$. O valor de Dice foi de 95,8%, com uma variação de $\pm 0,001$. A precisão do modelo foi de 97,1%, com uma variação de $\pm 0,001$. O valor do recall foi de 94,6%, com uma variação de $\pm 0,004$. Por fim, o valor do F1 foi de 95,8%, com uma variação de $\pm 0,001$. Os resultados do Teste 1 apontam para um desempenho elevado do modelo de aprendizado de máquina na tarefa de segmentação.

O Teste 2 foi realizado visando avaliar o desempenho do modelo, mas diferentemente do Teste 1, não foi utilizado o filtro CLAHE para pré-processamento das imagens e também não foi realizada a técnica de aumento de dados. A Figura 5.2 ilustra as médias obtidas pelo Teste 2.

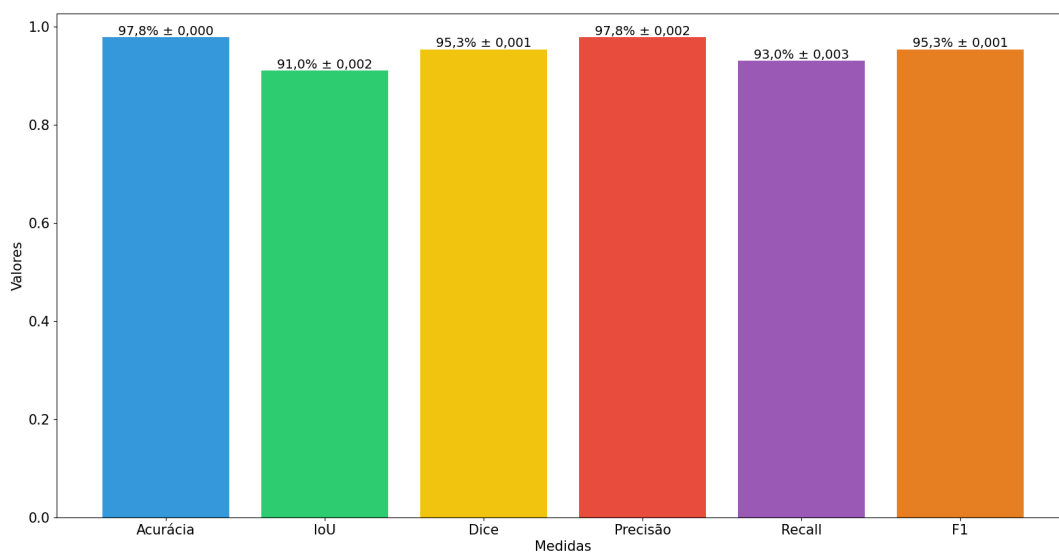
A acurácia do modelo foi de 98,1%, com uma variação de $\pm 0,0$. O valor do IoU foi de 92,4%, com uma variação de $\pm 0,001$. O valor de Dice foi de 96,1%, com uma variação de $\pm 0,001$. O valor da precisão do modelo foi de 97,2%, com uma variação de $\pm 0,004$. O valor

Figura 5.2 – Médias das medidas de avaliação na segmentação pulmonar (Teste 2).

Fonte: Elaborada pelo autor.

do recall foi de 94,9%, com uma variação de $\pm 0,003$. O valor do F1 foi de 96,1%, com uma variação de $\pm 0,001$. Os resultados do Teste 2 apontam para um desempenho superior quando comparado ao Teste 1.

No Teste 3, foram utilizadas duas técnicas de pré-processamento, o filtro CLAHE e o aumento de dados. Isso permitiu aumentar a quantidade de dados de treinamento. A Figura 5.3 apresenta as médias obtidas pelo Teste 3.

Figura 5.3 – Médias das medidas de avaliação na segmentação pulmonar (Teste 3).

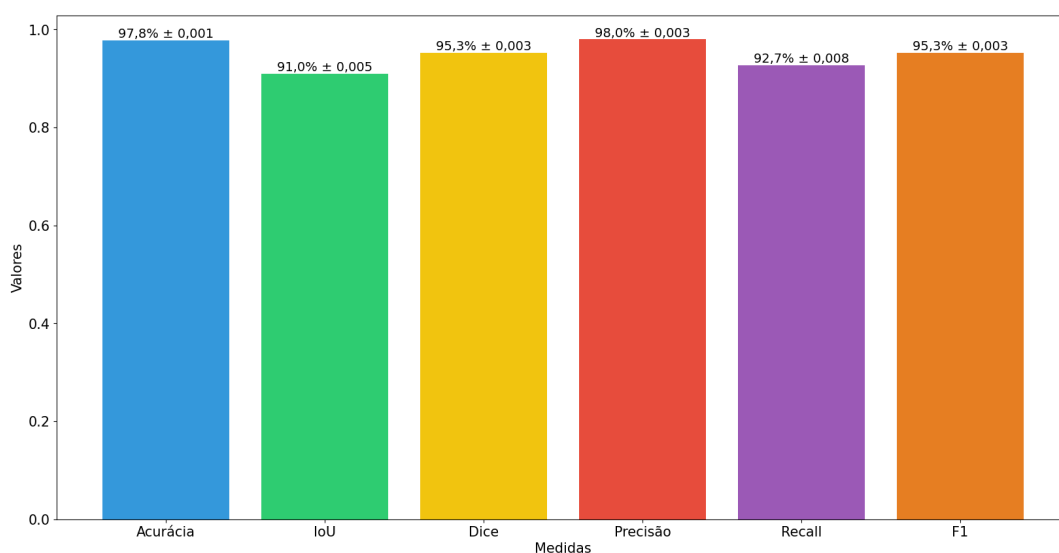
Fonte: Elaborada pelo autor.

A acurácia do modelo foi de 97,8%, com uma variação de $\pm 0,0$. O valor do IoU foi de 91%, com uma variação de $\pm 0,002$. O valor de Dice foi de 95,3%, com uma variação de \pm

0,001. O valor da precisão do modelo foi de 97,8%, com uma variação de $\pm 0,002$. O valor do recall foi de 93%, com uma variação de $\pm 0,003$. O valor do F1 foi de 95,3%, com uma variação de $\pm 0,001$. Os resultados do Teste 3 indicam que o uso de técnicas de pré-processamento, como filtro CLAHE e aumento de dados, não apresentou resultados melhores que os testes anteriores.

O Teste 4 foi realizado com o aumento de dados no conjunto de treinamento, permitindo aumentar a quantidade de dados de treinamento disponíveis para o modelo, mas sem o uso do filtro CLAHE para pré-processamento das imagens. A Figura 5.4 apresenta as médias obtidas pelo Teste 4.

Figura 5.4 – Médias das medidas de avaliação na segmentação pulmonar (Teste 4).



Fonte: Elaborada pelo autor.

A acurácia do modelo foi de 97,8%, com uma variação de $\pm 0,001$. O valor do IoU foi de 91%, com uma variação de $\pm 0,005$. O valor de Dice foi de 95,3%, com uma variação de $\pm 0,003$. O valor da precisão do modelo foi de 98%, com uma variação de $\pm 0,003$. O valor do recall foi de 92,7%, com uma variação de $\pm 0,008$. O valor do F1 foi de 95,3%, com uma variação de $\pm 0,003$.

As Tabelas 5.1 e 5.2 apresentam os resultados dos 4 testes realizados. Analisando esses resultados, o Teste 2 alcançou os melhores resultados no IoU, Dice, acurácia, recall e F1. A única medida que não alcançou o melhor resultado, foi na precisão, no qual o Teste 4 alcançou o melhor valor nessa medida.

O modelo utilizado no Teste 2 alcançou os melhores resultados em termos de acurácia, IoU, Dice, recall e F1. Por essa razão, esse modelo foi escolhido para realizar uma análise de segmentação. A medida de IoU foi selecionada para esta análise, já que é amplamente utilizada

Tabela 5.1 – Resultados dos testes realizados em segmentação pulmonar nas medidas de IoU e Dice.

	IoU	Dice
Teste 1	92% ± 0,003	95,8% ± 0,001
Teste 2	92,4% ± 0,001	96,1% ± 0,001
Teste 3	91% ± 0,002	95,3% ± 0,001
Teste 4	91% ± 0,005	95,3% ± 0,003

Tabela 5.2 – Resultados dos testes realizados em segmentação pulmonar nas demais medidas.

	Acurácia	Precisão	Recall	F1
Teste 1	98% ± 0,001	97,1% ± 0,001	94,6% ± 0,004	95,8% ± 0,001
Teste 2	98,1% ± 0,0	97,2% ± 0,004	94,9% ± 0,003	96,1% ± 0,001
Teste 3	97,8% ± 0,0	97,8% ± 0,002	93% ± 0,003	95,3% ± 0,001
Teste 4	97,8% ± 0,001	98% ± 0,003	92,7% ± 0,008	95,3% ± 0,003

quando se trata de segmentações de imagens. Foram definidas 10 faixas de valores obtidos usando a medida de IoU, como apresentado na Tabela. 5.3.

Tabela 5.3 – Resultados segmentação da infecção por faixas usando a medida IoU.

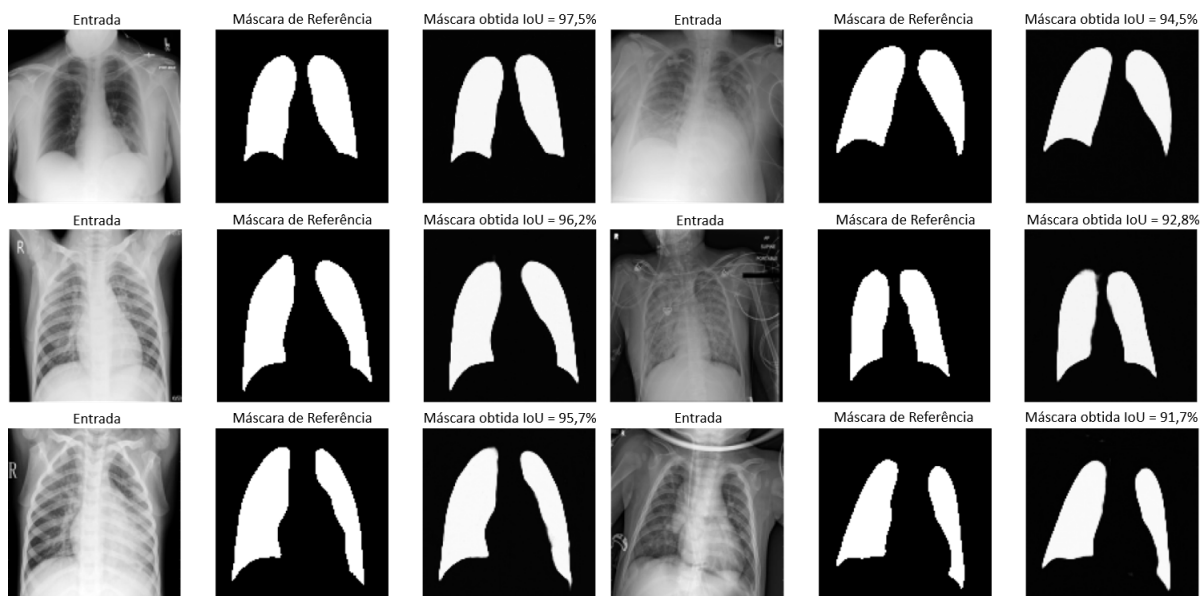
Faixa de IoU	Número de segmentações
0% — 9,9%	0
10% — 19,9%	0
20% — 29,9%	0
30% — 39,9%	0
40% — 49,9%	0
50% — 59,9%	0
60% — 69,9%	6
70% — 79,9%	28
80% — 89,9%	242
90% — 99,9%	1044
100%	0

Na Tabela 5.3 são apresentados as 1320 segmentações realizadas, das quais 1314 foram classificadas como boas ($\geq 70\%$), e 6 como ruins ($< 70\%$). Os resultados mostram que a maioria das segmentações boas (1044) têm valor de IOU entre 90% e 99,9%, seguido por 242 segmentações com valor de IOU entre 80% e 89,9% e 28 entre 70% e 79,9%. As segmentações ruins, por outro lado, têm valor de IOU entre 60% e 69,9%, com 6 ocorrências. Não houve nenhuma segmentação com valor de IOU abaixo de 60%.

As Figuras 5.5, 5.6, 5.7 e 5.8 apresentam seis exemplos, cada exemplo possui sua respectiva máscara de referência e máscara obtidas pela aplicação da técnica de segmentação de imagem com a MultiResUNet.

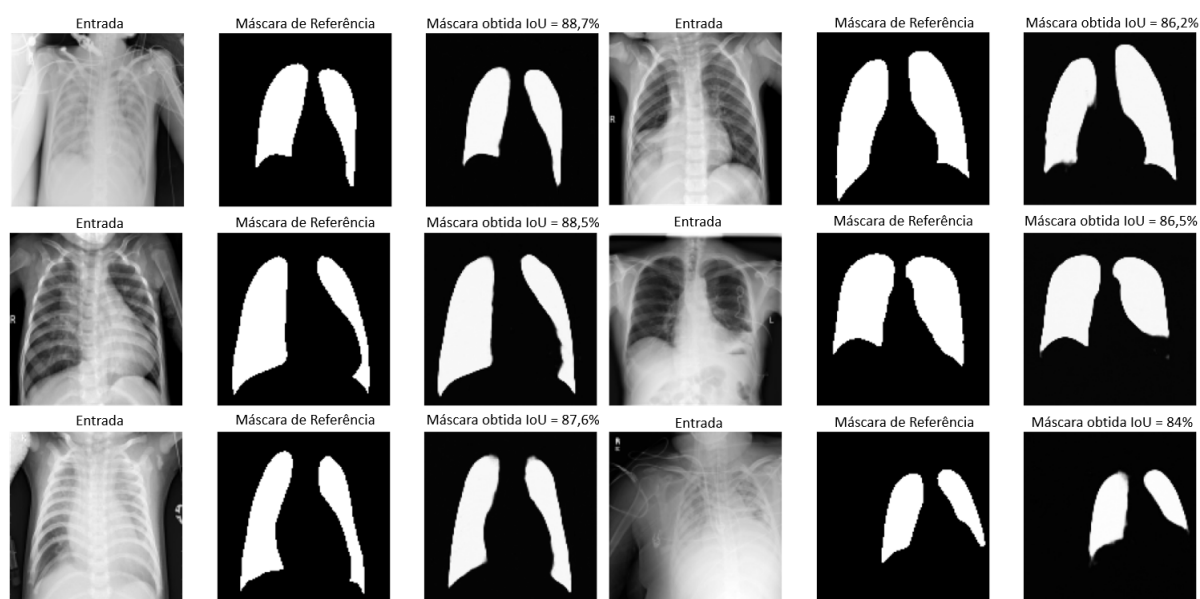
É possível observar que as máscaras obtidas se aproximam bastante da máscara de referência, mostrando a eficiência da técnica em identificar e separar os pulmões em imagens. É importante destacar que cada exemplo apresenta características diferentes, tanto na imagem

Figura 5.5 – Resultado da segmentação pulmonar na faixa de 90% a 99,9% de IoU.



Fonte: Elaborada pelo autor.

Figura 5.6 – Resultado da segmentação pulmonar na faixa de 80% a 89,9% de IoU.

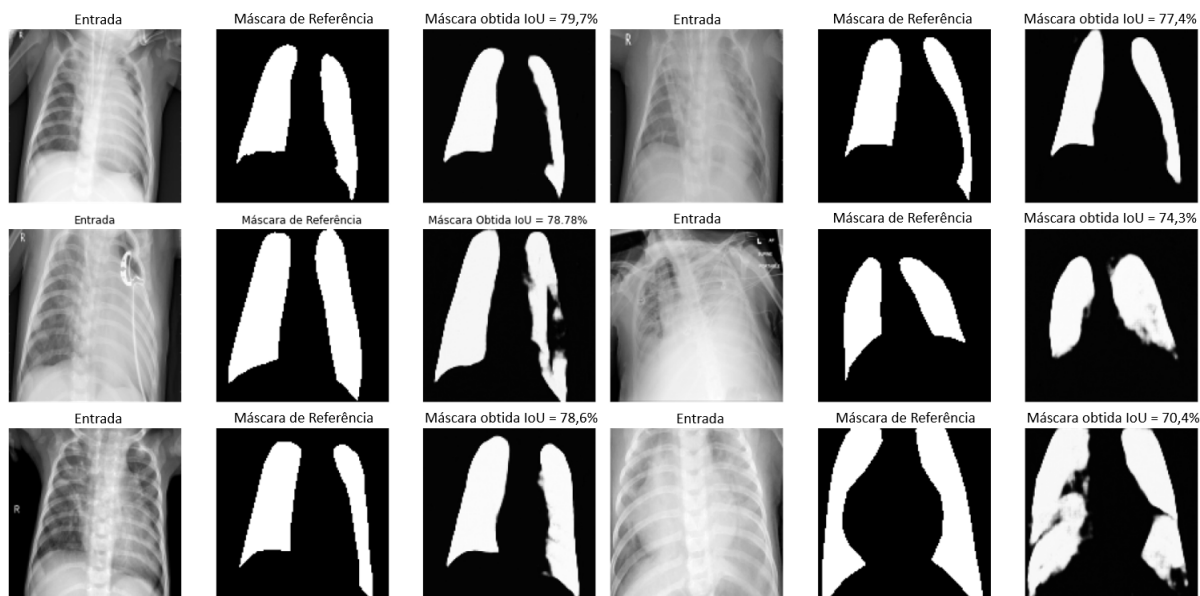


Fonte: Elaborada pelo autor.

original quanto nas máscaras de referência e obtidas pela rede. Isso mostra a versatilidade da rede em lidar com imagens de diferentes complexidades e conteúdos. Vale ressaltar que nas faixas de IoU de 70% e 60%, apresentaram-se problemas na segmentação, faltando partes das imagens ou definindo partes que não existiam como parte da máscara.

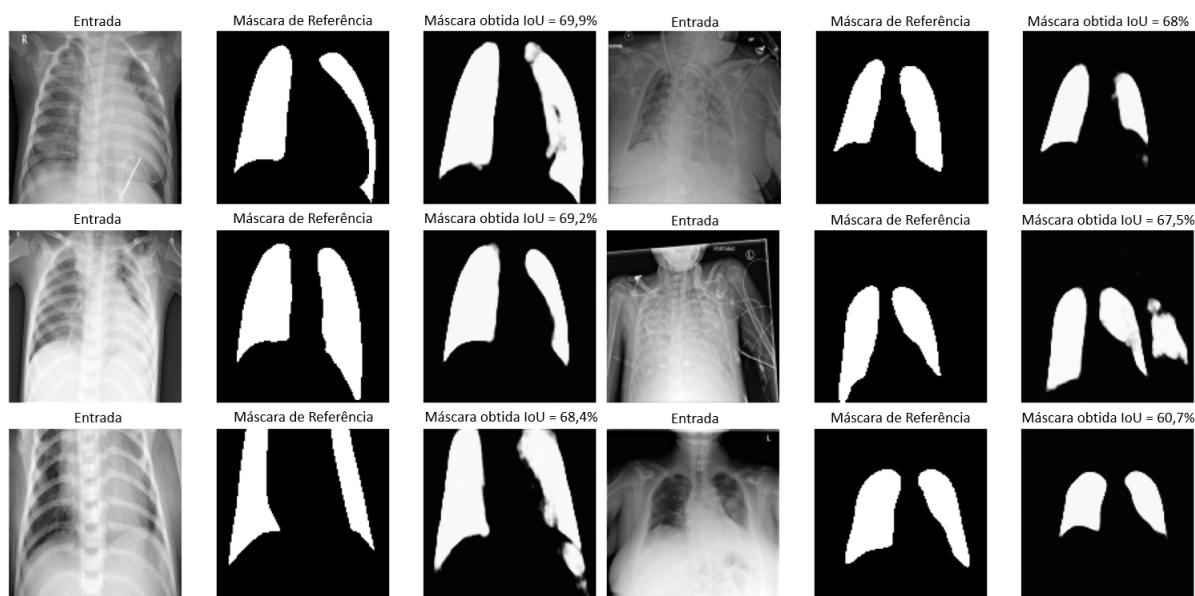
As imagens de COVID-19 e Não COVID-19 apresentam infecções, isso pode ter afetado a visibilidade do pulmão, e a qualidade das segmentações nas faixas de IoU 70% e 60% pode

Figura 5.7 – Resultado da segmentação pulmonar na faixa de 70% a 79,9% de IoU.



Fonte: Elaborada pelo autor.

Figura 5.8 – Resultado da segmentação pulmonar na faixa de 60% a 69,9% de IoU.

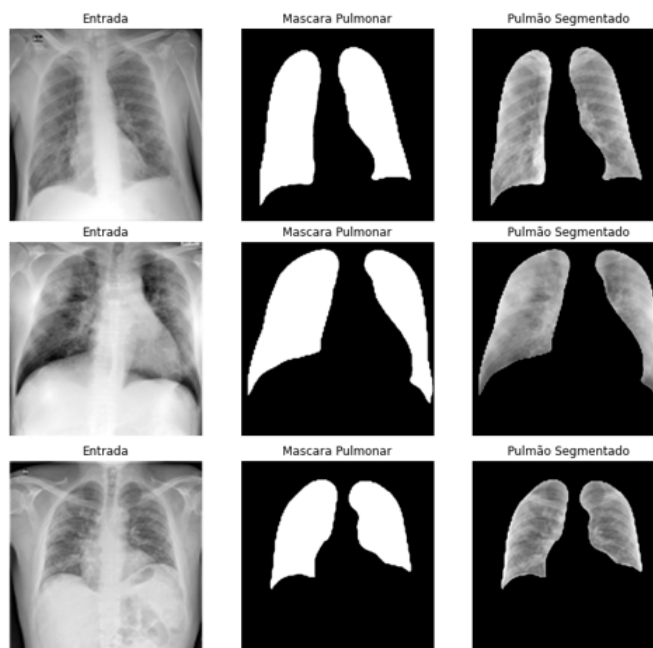


Fonte: Elaborada pelo autor.

não ser tão precisa por conta disso. Quando isso acontece, parte do pulmão acaba não sendo extraída, conforme pode-se observar nas Figuras 5.7 e 5.8.

Com o modelo obtido no Teste 2, as imagens de entrada são processadas e a região de interesse é extraída. A Figura 5.9 ilustra a imagem de entrada e a máscara obtida pela MultiResUNet. Além disso, mostra o resultado do processamento realizado para obter apenas os pulmões extraído. Esse processamento ajuda a remover partes da imagem que não são necessárias e que muitas vezes atrapalham na classificação, interferindo como ruído.

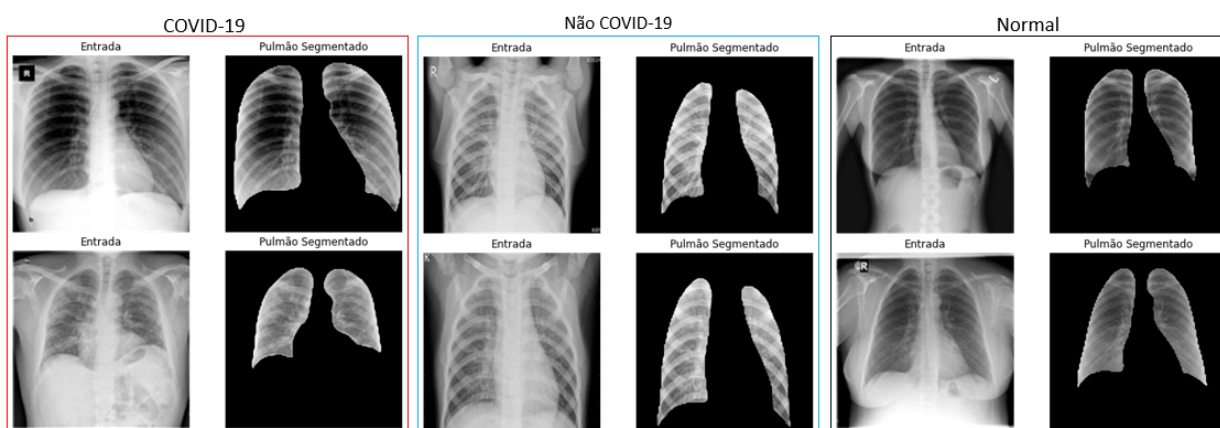
Figura 5.9 – Exemplos de segmentação pulmonar utilizando as máscaras geradas pelo modelo MultiResUnet.



Fonte: Elaborada pelo autor.

As imagens das classes COVID-19, Normal e Não COVID-19 foram segmentadas com o intuito de extrair os pulmões, conforme ilustrado na Figura 5.10. Os pulmões segmentados das três são utilizadas na etapa de classificação de imagens.

Figura 5.10 – Segmentação da região de interesse das três classes utilizadas para classificação de imagens.



Fonte: Elaborada pelo autor.

5.2 – Classificação binária das imagens de radiografias de tórax (COVID-19 e Não COVID-19)

Na etapa de classificação binária com imagens segmentadas, foram utilizados oito CNNs diferentes. Essas CNNs estão listadas Tabela 5.4. Cada modelo foi treinado utilizando validação cruzada. Para cada modelo, foram calculadas as medidas de desempenho: acurácia, sensibilidade, especificidade, precisão e F1. Cada medida é acompanhada pelo desvio padrão correspondente. Os resultados da classificação binária estão apresentados nas Tabelas 5.4 e 5.5.

Tabela 5.4 – CNNs utilizadas para classificação binária.

Número	Modelo	Tamanho do modelo obtido
1	VGG19	80,30 MB
2	VGG16	59,05 MB
3	InceptionV3	88,47 MB
4	DenseNet201	75,40 MB
5	ResNet50	95,32 MB
6	Xception	84,31 MB
7	InceptionResNetV2	219,48 MB
8	NASNetLarge	343,06 MB

Tabela 5.5 – Resultados das CNNs na classificação binária de imagens com o pulmão segmentado entre COVID-19 e Não COVID-19.

	Acurácia	Sensibilidade	Especificidade	Precisão	F1
1	84,7% ± 0,018	82,2% ± 0,070	87,3% ± 0,041	87,5% ± 0,026	84,6% ± 0,026
2	86,3% ± 0,021	81,6% ± 0,069	91,3% ± 0,031	91% ± 0,022	85,9% ± 0,028
3	79,2% ± 0,011	77,7% ± 0,025	80,9% ± 0,009	81,1% ± 0,006	79,4% ± 0,014
4	84,5% ± 0,008	86,4% ± 0,009	82,4% ± 0,014	83,9% ± 0,011	85,1% ± 0,007
5	74,2% ± 0,003	89,3% ± 0,004	58,3% ± 0,010	69,4% ± 0,004	78,1% ± 0,002
6	83,5% ± 0,005	84,4% ± 0,038	82,6% ± 0,044	83,9% ± 0,027	84% ± 0,006
7	84% ± 0,011	86% ± 0,068	81,9% ± 0,060	83,7% ± 0,033	84,6% ± 0,017
8	80% ± 0,003	80,4% ± 0,025	79,5% ± 0,022	80,7% ± 0,012	80,5% ± 0,007

Os resultados apresentados na tabela indicam que a maioria das arquiteturas obteve uma acurácia superior a 80%. A VGG16 se destacou com a melhor acurácia de 86,3% com uma variação de $\pm 0,021$, seguida por VGG19 com 84,7% com uma variação de $\pm 0,018$ e DenseNet201 com 84,5% com uma variação de $\pm 0,008$. Por outro lado, a ResNet50 apresentou a pior acurácia de 74,2% com uma variação de $\pm 0,003$.

Em relação à sensibilidade, a ResNet50 teve o melhor desempenho, com 89,3% com uma variação de $\pm 0,004$, enquanto a InceptionV3 apresentou o pior resultado, com 77,7% com uma variação de $\pm 0,025$, sendo o único modelo com valor abaixo dos 80%. O modelo com

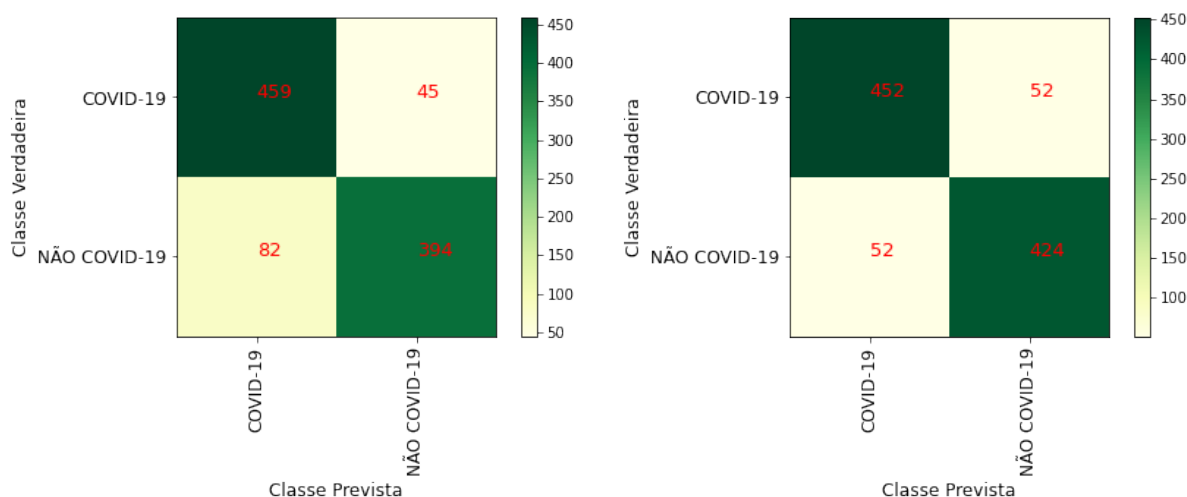
a melhor especificidade é VGG16 com 91,3% com uma variação de $\pm 0,031$, sendo o único modelo a obter um valor na faixa dos 90%, enquanto o modelo com a pior especificidade é ResNet50 com 58,3% com uma variação de $\pm 0,010$.

Em termos de precisão, a NASNetLarge apresentou o pior resultado, com 80,7% com uma variação de $\pm 0,012$, enquanto a VGG16 obteve a melhor precisão, com 91% com uma variação de $\pm 0,022$, sendo o único modelo a obter um valor na faixa dos 90%. Em relação ao F1, a VGG16 obteve a melhor pontuação, com 85,9% com uma variação de $\pm 0,028$, enquanto a ResNet50 obteve o pior F1, com 78,1% com uma variação de $\pm 0,002$.

No geral, a VGG16 se destacou como a arquitetura mais consistente, e confiável em termos de desempenho geral e medidas específicas, enquanto a ResNet50 teve um desempenho inferior em comparação com as outras arquiteturas testadas.

Foi utilizada a matriz de confusão para verificar se os modelos de classificação obtidos conseguiam distinguir as imagens de COVID-19 das imagens Não COVID-19. Como o treinamento foi realizado com base em 5 *fold*s, em cada *fold*, o melhor modelo obtido no treinamento era salvo, gerando 5 modelos para cada uma das 8 arquiteturas. Usando o conjunto de teste, foi analisado qual dos 5 modelos salvos de cada rede obteve a melhor acurácia e este modelo foi utilizado para gerar a matriz de confusão com base no conjunto de teste. As matrizes de cada modelo são apresentadas nas Figuras 5.11, 5.12, 5.13, 5.14.

Figura 5.11 – Matrizes de confusão classificação binária da VGG19 e VGG16.



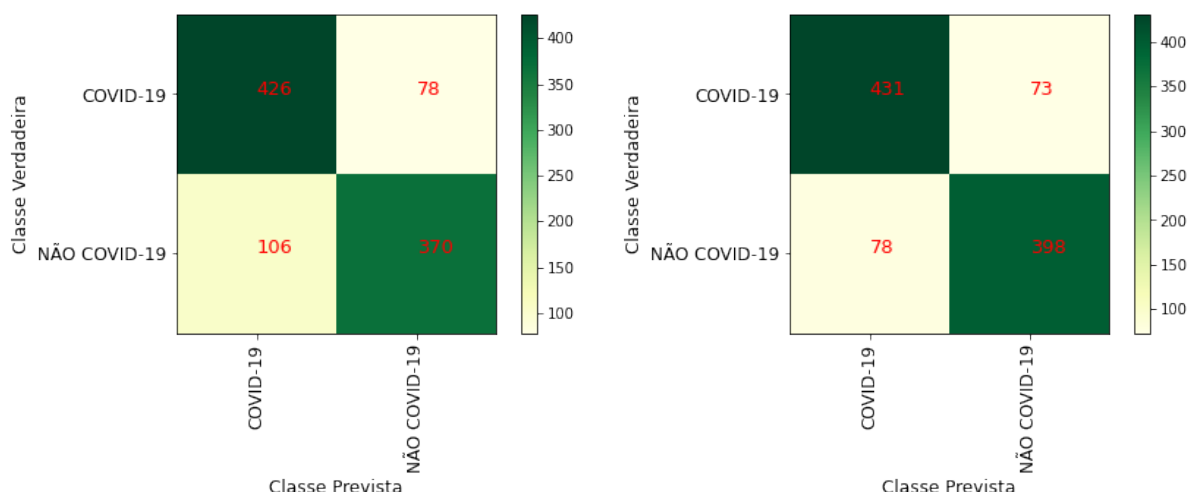
(a) Matriz de confusão VGG19

(b) Matriz de confusão VGG16

A matriz de confusão da VGG19 apresentou um desempenho com apenas 45 erros na classificação de COVID-19 e 82 erros na classificação de Não COVID-19. O modelo apresentou uma taxa de acerto de 91,1% para COVID-19 e 82,5% para Não COVID-19.

Com base na matriz de confusão da VGG16, pode-se observar que o modelo também apresenta um desempenho com 52 erros na classificação de COVID-19 e 52 erros na classificação de Não COVID-19. Isso indica que o modelo teve um equilíbrio melhor que a VGG19 entre a taxa de acerto para as duas classes. O modelo apresentou uma taxa de acerto de 89,7% para COVID-19 e 89,1% para Não COVID-19.

Figura 5.12 – Matrizes de confusão classificação binária da InceptionV3 e DenseNet201.



(c) Matriz de confusão InceptionV3

(d) Matriz de confusão DenseNet201

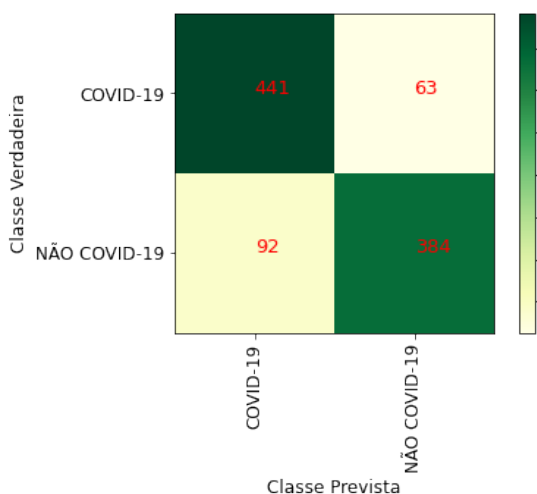
A matriz de confusão da InceptionV3 apresentou uma taxa de acerto de 84,8% para COVID-19 e 77,7% para não COVID-19, com 78 erros na classificação de COVID-19 e 106 erros na classificação de não COVID-19. Isso indica que o modelo teve um desempenho inferior em relação ao modelo VGG19 e VGG16 na classificação de COVID-19.

A matriz de confusão da DenseNet201 apresentou uma taxa de acerto de 86,2% para COVID-19 e 83,5% para Não COVID-19, com 73 erros na classificação de COVID-19 e 78 erros na classificação de Não COVID-19. Isso indica que o modelo teve um desempenho inferior em relação ao modelo VGG19 na classificação de COVID-19, mas superior na classe Não COVID-19, sendo inferior ao VGG16 em ambas as situações.

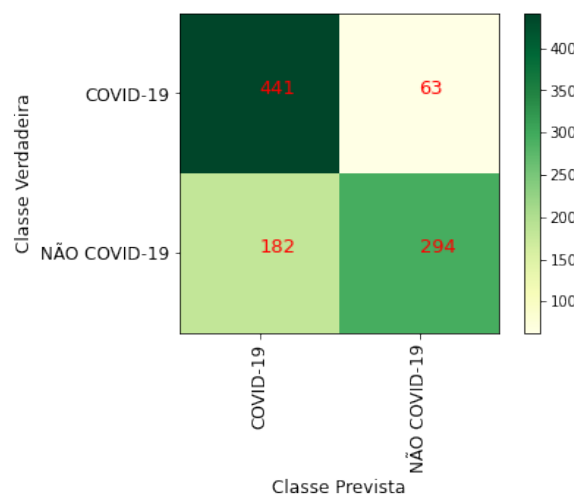
A matriz de confusão da Xception apresentou uma taxa de acerto de 88,2% para COVID-19 e 80,6% para Não COVID-19, com 63 erros na classificação de COVID-19 e 92 erros na classificação de Não COVID-19. Comparando com as outras arquiteturas avaliadas, a Xception apresentou um desempenho próximo ao da DenseNet201.

A matriz de confusão da ResNet50 apresentou uma taxa de acerto de 88,2% para COVID-19 e 61,8% para Não COVID-19, com 63 erros na classificação de COVID-19, assim como a Xception, e 182 erros na classificação de Não COVID-19. Isso indica que o modelo

Figura 5.13 – Matrizes de confusão classificação binária da Xception e ResNet50.



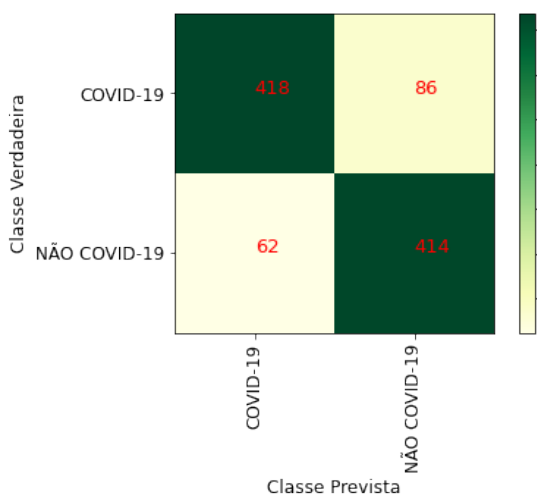
(e) Matriz de confusão Xception



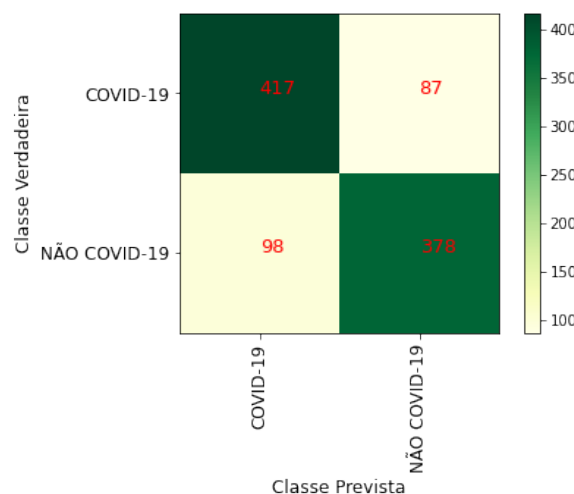
(f) Matriz de confusão ResNet50

apresentou muitos erros na classificação de Não COVID-19, sendo o que mais errou nessa classe.

Figura 5.14 – Matrizes de confusão classificação binária da InceptionResNetV2 e NASNetLarge.



(g) Matriz de confusão InceptionResNetV2



(h) Matriz de confusão NASNetLarge

A matriz de confusão da InceptionResNetV2 apresentou uma taxa de acerto de 83,7% para COVID-19 e 87% para Não COVID-19, com 86 erros na classificação de COVID-19 e 62 erros na classificação de Não COVID-19. Isso indica que teve mais erros na classificação de COVID-19 em comparação com outros modelos.

A matriz de confusão da NASNetLarge apresentou uma taxa de acerto de 83,4% para COVID-19 e 79,4% para Não COVID-19, com 87 erros na classificação de COVID-19 e 98

erros na classificação de Não COVID-19. Isso indica que o modelo apresentou um resultado similar ao InceptionResNetV2 na classe COVID-19.

Comparando, de forma geral, as matrizes de confusão dos modelos na classificação de COVID-19, podemos observar que há uma variação nos desempenhos dos diferentes modelos. Em geral, os modelos VGG19, VGG16 e DenseNet201 apresentaram um desempenho superior na classificação das imagens em comparação com os outros modelos. A Tabela 5.6 apresenta os resultados de acertos e erros das matrizes de confusão da classificação binária.

Tabela 5.6 – Tabela resumo das matrizes de confusão na classificação binária

Modelo	COVID-19		Não COVID-19	
	Acerto	Erro	Acerto	Erro
1	459 (91,1%)	45 (8,9%)	394 (82,5%)	82 (17,5%)
2	452 (89,7%)	52 (10,3%)	424 (89,1%)	52 (10,9%)
3	426 (84,8%)	78 (15,2%)	370 (77,7%)	106 (22,4%)
4	431 (86,2%)	73 (13,8%)	398 (83,5%)	78 (16,5%)
5	441 (88,2%)	63 (11,8%)	294 (61,8%)	182 (38,2%)
6	441 (88,2%)	63 (11,8%)	384 (80,6%)	92 (19,4%)
7	418 (83,7%)	86 (16,3%)	414 (87%)	62 (13,0%)
8	417 (83,4%)	87 (16,6%)	378 (79,4%)	98 (20,6%)

De forma geral, podemos observar que todos os modelos apresentaram uma taxa de acerto acima de 80% na detecção de imagens com COVID-19, indicando que a maioria das imagens foi corretamente classificadas como COVID-19. No entanto, a taxa de acerto para imagens de Não COVID-19 é menor do que a taxa de acerto para pacientes com COVID-19.

5.3 – Classificação multiclases das imagens radiografias de tórax (COVID-19, Normal e Não COVID-19)

Na etapa de classificação, foi realizada uma segunda análise que consistiu na classificação multiclases com imagens segmentadas. Ao contrário da análise anterior, que era uma classificação binária, foram utilizadas as classes COVID-19, Normal e Não COVID-19. Os mesmos modelos foram utilizados, e as mesmas medidas de desempenho foram aplicadas. Os resultados da classificação multiclases estão apresentados nas Tabelas 5.7 e 5.8.

A partir dos resultados apresentados, podemos observar que o modelo VGG16 obteve a maior acurácia entre os modelos avaliados, com uma média de 80,9% com uma variação de $\pm 0,010$. Isso sugere que o modelo VGG16 consegue prever corretamente a classe da maioria das imagens do conjunto de dados. Além disso, o modelo VGG16 também teve a

Tabela 5.7 – CNNs utilizadas para classificação multiclassadas.

Número	Modelo	Tamanho do modelo obtido
1	VGG19	80,30 MB
2	VGG16	59,05 MB
3	InceptionV3	88,47 MB
4	DenseNet201	75,40 MB
5	ResNet50	95,32 MB
6	Xception	84,31 MB
7	InceptionResNetV2	219,48 MB
8	NASNetLarge	343,06 MB

Tabela 5.8 – Resultados das CNNs com as imagens com o pulmão segmentado na classificação multiclassadas COVID-19, Normal e Não COVID-19.

	Acurácia	Sensibilidade	Especificidade	Precisão	F1
1	79,1% ± 0,005	84,3% ± 0,041	95% ± 0,018	95% ± 0,015	89,2% ± 0,017
2	80,9% ± 0,010	86,6% ± 0,032	94,6% ± 0,011	94,6% ± 0,008	90,4% ± 0,014
3	72,9% ± 0,007	72,9% ± 0,015	93,6% ± 0,018	91,8% ± 0,020	81,2% ± 0,007
4	75,2% ± 0,005	72,1% ± 0,011	95,1% ± 0,011	93,9% ± 0,012	81,6% ± 0,005
5	68,8% ± 0,0024	75,5% ± 0,016	85,7% ± 0,016	85,2% ± 0,011	80% ± 0,005
6	76,49% ± 0,008	81,50% ± 0,061	92,92% ± 0,039	92,57% ± 0,034	86,47% ± 0,021
7	77,6% ± 0,007	82,4% ± 0,023	95,2% ± 0,012	94,6% ± 0,012	88,1% ± 0,012
8	73% ± 0,004	75,2% ± 0,015	93,7% ± 0,012	92,3% ± 0,012	82,9% ± 0,007

maior especificidade entre os modelos avaliados, com uma média de 86,6% com uma variação de $\pm 0,032$.

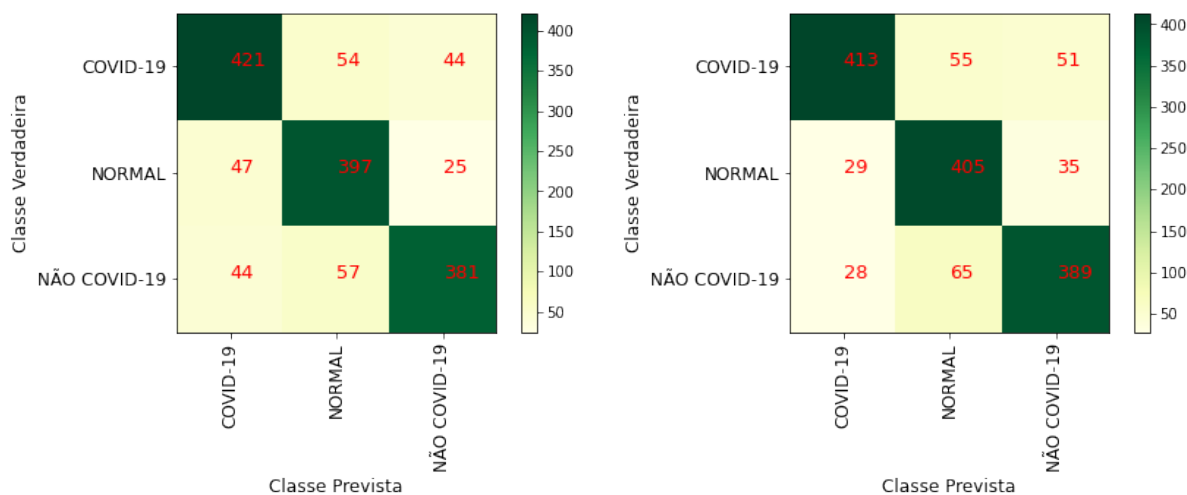
Outro modelo que se destacou foi o InceptionResNetV2, que obteve a maior especificidade entre os modelos avaliados, com uma média de 95,2% com uma variação de $\pm 0,012$. Em relação à precisão, o modelo VGG19 obteve o melhor resultado, com uma média de 95% com uma variação de $\pm 0,015$. O F1, foi maior para o modelo VGG16, com uma média de 90,4% com uma variação de $\pm 0,014$.

Quanto aos modelos com desempenho inferior, pode-se observar que o modelo ResNet50 obteve a menor acurácia, com uma média de 68,8% com uma variação de $\pm 0,0024$, e também a menor sensibilidade, com uma média de 75,5% com uma variação de $\pm 0,016$. Isso sugere que o modelo ResNet50 teve dificuldade em prever corretamente as imagens.

Em resumo, os resultados apresentados sugerem que o modelo VGG16 obteve o melhor desempenho geral na tarefa de classificação de imagens radiográficas de tórax para COVID-19, Normal e Não COVID-19. Esse modelo se destacou em medidas como acurácia, sensibilidade e F1. No entanto, outros modelos, como o InceptionResNetV2 e o VGG19, também apresentaram desempenho superior em medidas específicas, como especificidade e precisão, respectivamente.

Para analisar melhor os resultados, assim como foi realizado na classificação binária, na classificação multiclases também foram obtidas as matrizes de confusão. As matrizes de cada modelo são apresentadas nas Figuras 5.15, 5.16, 5.17, 5.18.

Figura 5.15 – Matrizes de confusão classificação multiclases da VGG19 e VGG16.



(a) Matriz de confusão VGG19

(b) Matriz de confusão VGG16

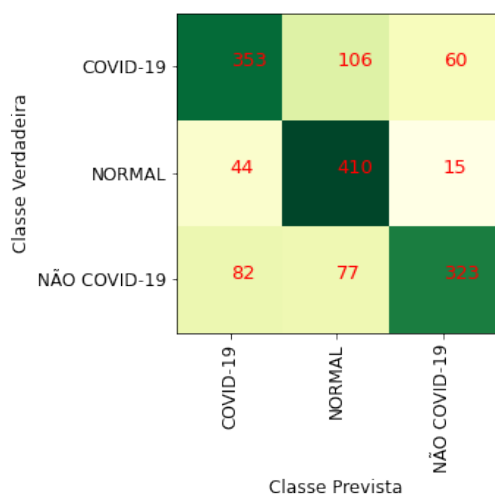
Na matriz de confusão da VGG19 obteve uma taxa de acerto de 81,1% para imagens de COVID-19, com 421 acertos e 98 erros; 84,5% para imagens normais, 397 acertos e 72 erros; e 79% para imagens Não COVID-19, com 381 acertos e 101 erros. Seu desempenho foi relativamente bom em comparação com outros modelos em todas as categorias.

VGG16 obteve uma taxa de acerto de 79,6% para imagens de COVID-19, com 413 acertos e 106 erros; 86,4% para imagens normais, 405 acertos e 64 erros; e 80,7% para imagens Não COVID-19, com 389 acertos e 93 erros. Seu desempenho foi relativamente bom em comparação com outros modelos na detecção de imagens Normal e Não COVID-19.

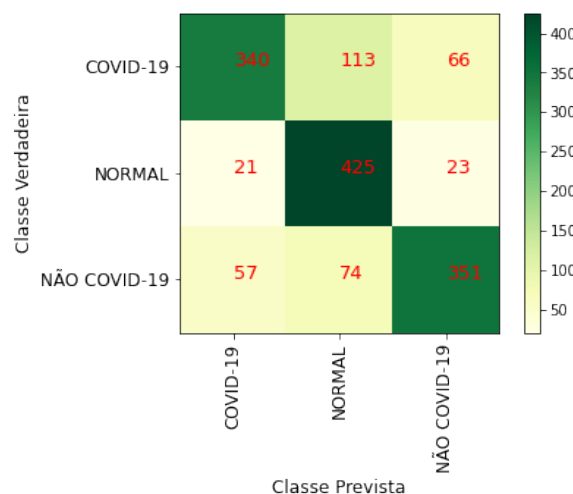
InceptionV3 obteve uma taxa de acerto de 68% para imagens de COVID-19, com 353 acertos e 166 erros; 87,5% para imagens normais, com 410 acertos e 59 erros; e 67,1% para imagens Não COVID-19, com 323 acertos e 159 erros. Seu desempenho foi relativamente baixo em comparação com outros modelos na detecção de imagens COVID-19 e Não COVID-19.

Ao analisar a matriz de confusão da rede DenseNet201 obteve uma taxa de acerto de 65,5% para imagens de COVID-19, com 340 acertos e 179 erros; 90,8% para imagens normais, 425 acertos e 44 erros; e 72,8% para imagens Não COVID-19, com 351 acertos e 131 erros. Seu desempenho foi relativamente baixo em comparação com outros modelos na detecção de imagens COVID-19, mas foi o melhor modelo na detecção de imagens Normal.

Figura 5.16 – Matrizes de confusão classificação multiclass da InceptionV3 e DenseNet201.

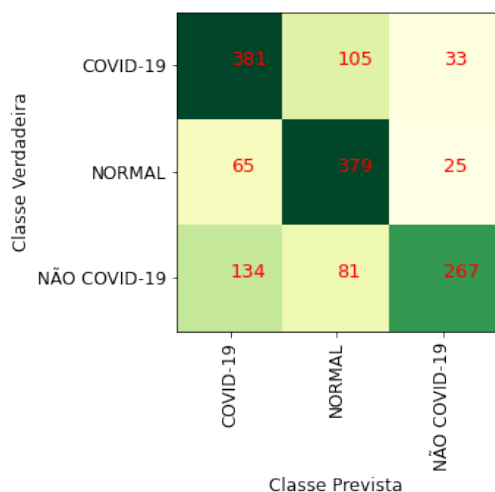


(c) Matriz de confusão InceptionV3

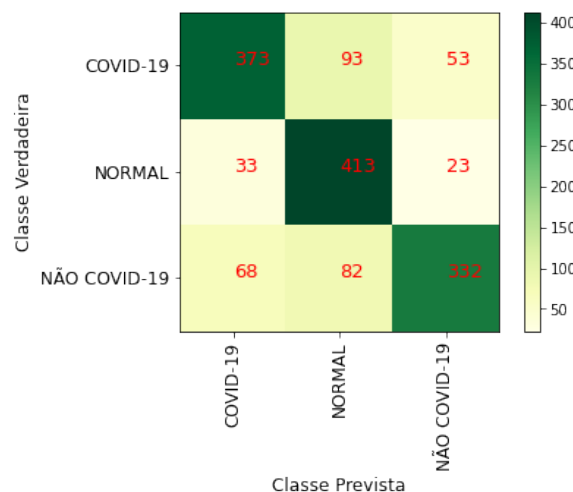


(d) Matriz de confusão DenseNet201

Figura 5.17 – Matrizes de confusão classificação multiclass da ResNet50 e Xception.



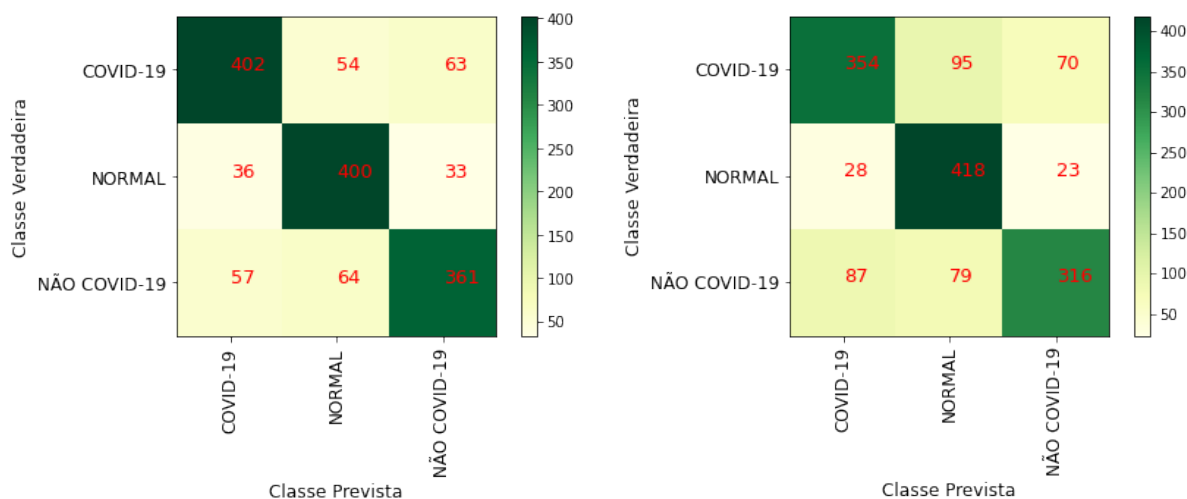
(f) Matriz de confusão ResNet50



(e) Matriz de confusão Xception

A matriz de confusão da rede ResNet50 obteve uma taxa de acerto de 73,4% para imagens de COVID-19, foram 381 acertos e 138 erros; 80,8% para imagens normais foram 379 acertos e 90 erros; e 55,4% para imagens Não COVID-19, foram 267 acertos e 215 erros. Seu desempenho na detecção de COVID-19 foi melhor do que a maioria dos outros modelos, mas foi relativamente baixo na detecção de imagens Não COVID-19.

A matriz de confusão da rede Xception obteve uma taxa de acerto de 71,9% para imagens de COVID-19, com 373 acertos e 146 erros; 88,1% para imagens normais, com 413 acertos e 56 erros; e 68,9% para imagens Não COVID-19, com 332 acertos e 150 erros. Seu desempenho foi relativamente bom em comparação com outros modelos na detecção de imagens Normal, mas foi o pior modelo na detecção de imagens Não COVID-19.

Figura 5.18 – Matrizes de confusão classificação multiclass da InceptionResNetV2 e NASNetLarge.

(g) Matriz de confusão InceptionResNetV2

(h) Matriz de confusão NASNetLarge

A InceptionResNetV2 obteve uma taxa de acerto de 77,4% para imagens de COVID-19, com 402 acertos e 117 erros; 85,3% para imagens normais, com 400 acertos e 69 erros; e 74,7% para imagens Não COVID-19, com 361 acertos e 121 erros. Seu desempenho foi relativamente bom em comparação com outros modelos na detecção de imagens COVID-19.

NASNetLarge obteve uma taxa de acerto de 68,2% para imagens de COVID-19, com 354 acertos e 165 erros; 89,1% para imagens de Normal, 418 acertos e 51 erros; e 65,6% para imagens Não COVID-19, com 316 acertos e 166 erros. Seu desempenho foi relativamente baixo em comparação com outros modelos na detecção de imagens COVID-19 e Não COVID-19, mas foi o segundo melhor modelo na detecção de imagens Normal.

A Tabela 5.9 apresenta os resultados de acerto e erro das matrizes de confusão da classificação multiclass.

Tabela 5.9 – Tabela resumo das matrizes de confusão na classificação multiclass.

Modelo	COVID-19		Normal		Não COVID-19	
	Acerto	Erro	Acerto	Erro	Acerto	Erro
1	421 (81,1%)	98 (18,9%)	397 (84,5%)	72 (15,5%)	381 (79,0%)	101 (21,0%)
2	413 (79,6%)	106 (20,4%)	405 (86,4%)	64 (13,6%)	389 (80,7%)	93 (19,3%)
3	353 (68,0%)	166 (32,0%)	410 (87,5%)	59 (12,5%)	323 (67,1%)	159 (33,0%)
4	340 (65,5%)	179 (34,5%)	425 (90,8%)	44 (9,2%)	351 (72,8%)	131 (27,2%)
5	381 (73,4%)	138 (26,6%)	379 (80,8%)	90 (19,2%)	267 (55,4%)	215 (44,6%)
6	373 (71,9%)	146 (28,1%)	413 (88,1%)	56 (11,9%)	332 (68,9%)	150 (31,1%)
7	402 (77,4%)	117 (22,6%)	400 (85,3%)	69 (14,7%)	361 (74,7%)	121 (25,3%)
8	354 (68,2%)	165 (31,8%)	418 (89,1%)	51 (10,9%)	316 (65,6%)	166 (34,4%)

Em geral, podemos perceber que os modelos tiveram um desempenho melhor na detecção de imagens normais e piores na detecção de imagens Não COVID-19. Em termos

de comparação entre os modelos, podemos observar que os modelos VGG19 e VGG16 tiveram um desempenho maior nas classes COVID-19 e Não COVID-19, respectivamente.

Em suma, os resultados indicam que os modelos de redes neurais convolucionais apresentaram um desempenho inferior na classificação multiclasse a partir de imagens de radiografias em comparação com a classificação binária.

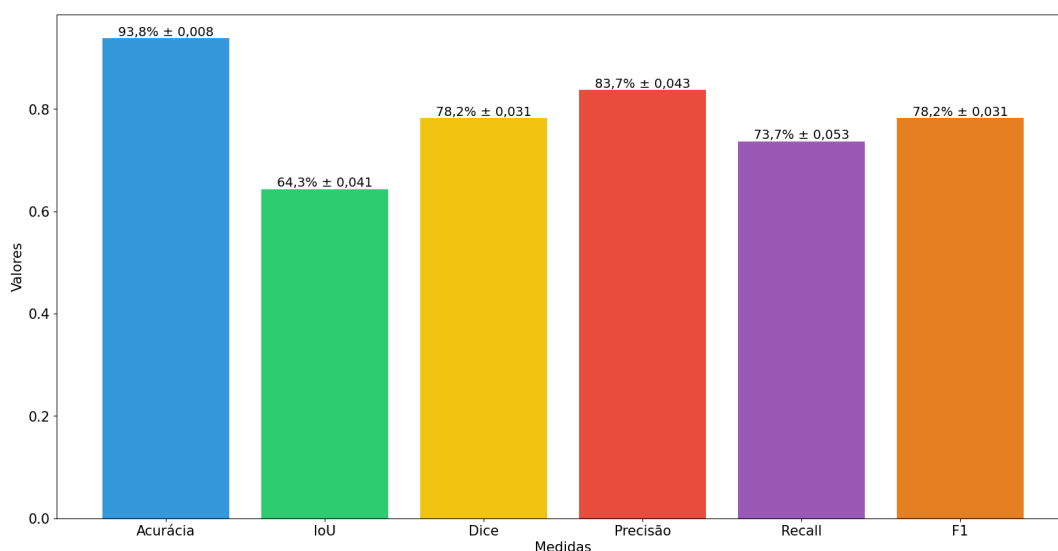
5.4 – Segmentação da infecção causada pela COVID-19

Nesta seção, é descrito os resultados obtidos na segmentação da infecção por COVID-19 em imagens de radiografia torácica usando a MultiResUNet. Também é apresentado os resultados obtidos utilizando a técnica LIME para segmentar a infecção por COVID-19 em imagens.

5.4.1 – Segmentação da Infecção COVID-19 com a MultiResUNET

Na segmentação da infecção por COVID-19 em radiografias de tórax foram realizados 4 diferentes testes, assim como foi realizado anteriormente na segmentação pulmonar. Na Figura 4.6 observa-se cada teste. Para cada teste, foram calculadas 6 medidas de desempenho após o treinamento de cada *fold*, usando o conjunto de teste para avaliar o modelo obtido. Vale ressaltar que esse conjunto não foi utilizado no treinamento. Na Figura 5.19 é ilustrado as medidas de desempenho obtidas.

Figura 5.19 – Médias das medidas de avaliação na segmentação da infecção com a MultiResUNet (Teste 1).

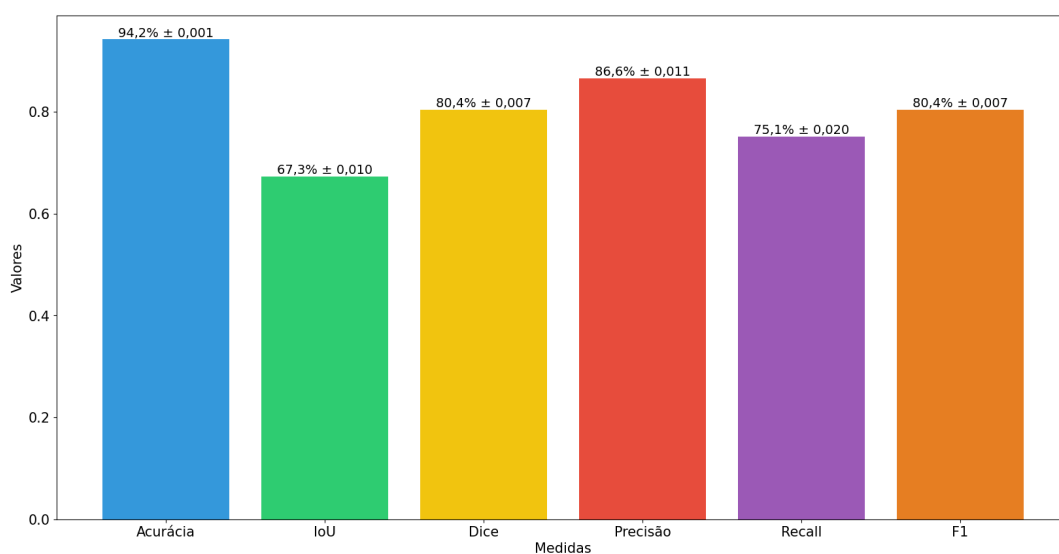


Fonte: Elaborada pelo autor.

No Teste 1 foi utilizado o filtro CLAHE, mas não foi usando o aumento de dados, os resultados mostram que o modelo de segmentação de imagens teve uma acurácia de 93,8% com uma variação de $\pm 0,008$, o valor do IoU foi de 64,3% com uma variação de $\pm 0,041$, o valor de Dice foi de 78,2% com uma variação de $\pm 0,031$. O valor da precisão foi de 83,7% com uma variação de $\pm 0,043$, o valor do recall foi de 73,7% com uma variação de $\pm 0,053$, o valor de F1 foi de 78,2% com uma variação de $\pm 0,031$.

O Teste 2 consisti na não utilização do filtro CLAHE e não se utilizou do aumento de dados. Os resultados são apresentados na Figura 5.20. O valor da acurácia foi de 94,2% com uma variação de $\pm 0,001$, o valor do IoU foi de 67,3% com uma variação de $\pm 0,010$, valor de Dice foi de 80,4% com uma variação de $\pm 0,007$. O valor da precisão foi de 86,6% com uma variação de $\pm 0,011$, o valor do recall foi de 75,1% com uma variação de $\pm 0,020$ e o valor do F1 foi de 80,4% com uma variação de $\pm 0,007$.

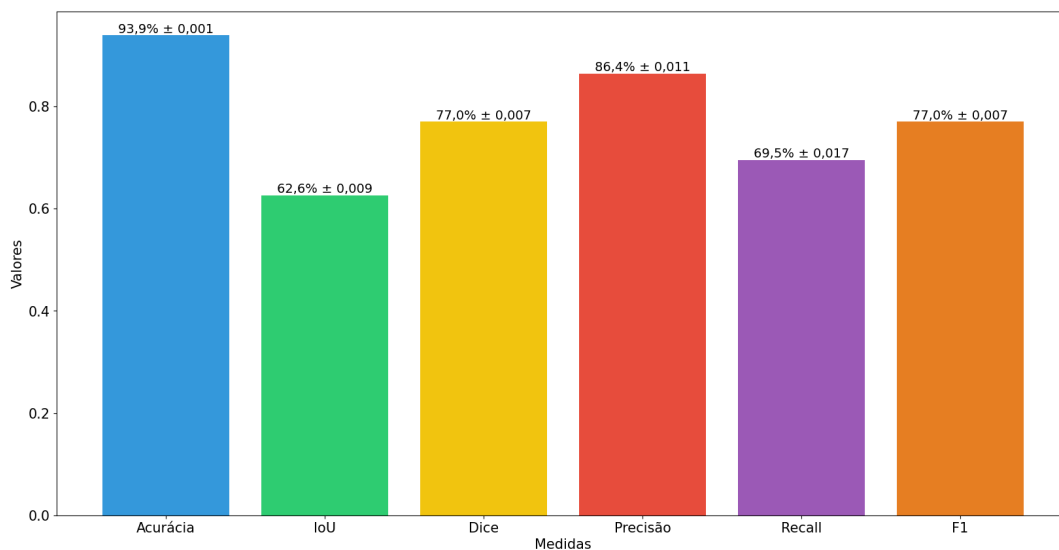
Figura 5.20 – Médias das medidas de avaliação na segmentação da infecção com a MultiResUNet (Teste 2).



Fonte: Elaborada pelo autor.

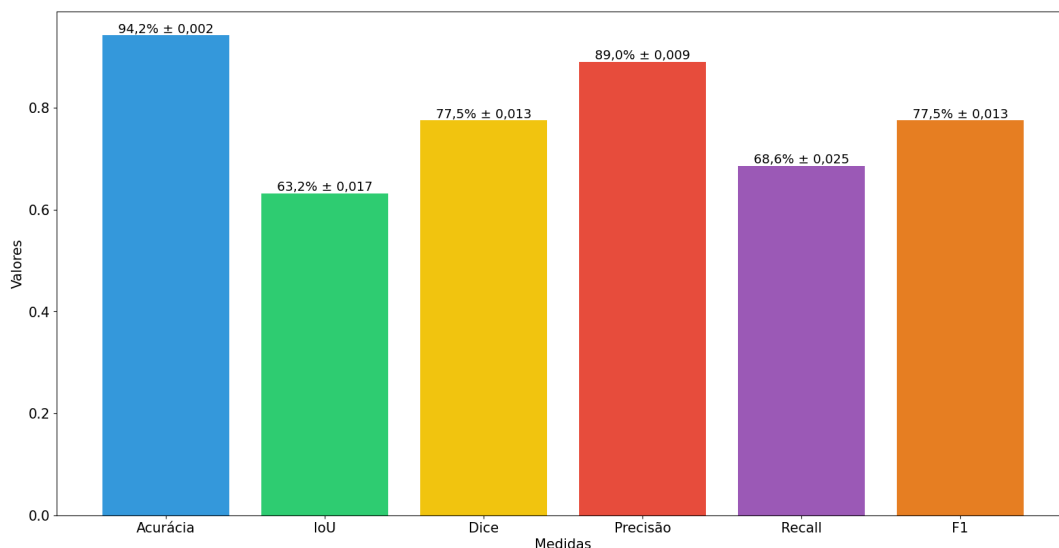
No Teste 3, foi aplicado o filtro CLAHE e também o aumento de dados. O resultado desse teste é ilustrado na Figura 5.21. A acurácia do modelo foi de 93,9% com uma variação de $\pm 0,001$, o valor do IoU foi de 62,6% com uma variação de $\pm 0,009$, o valor de Dice foi de 77% com uma variação de $\pm 0,007$. O valor da precisão foi de 86,3% com uma variação de $\pm 0,011$, o valor do recall foi de 69,5% com uma variação de $\pm 0,017$ e o valor do F1 foi de 77% com uma variação de $\pm 0,007$.

O Teste 4 não aplicou o filtro CLAHE, mas houve aumento de dados. Os resultados obtidos nesse teste foram: acurácia de 94,2% com uma variação de $\pm 0,002$, valor do IoU foi de

Figura 5.21 – Médias das medidas de avaliação na segmentação da infecção com a MultiResUNet (Teste 3).

Fonte: Elaborada pelo autor.

63,2% com uma variação de $\pm 0,017$, valor de Dice foi de 77,5% com uma variação de $\pm 0,013$. O valor da precisão foi de 89% com uma variação de $\pm 0,009$, o valor de recall foi de 68,6% com uma variação de $\pm 0,025$ e o valor de F1 foi de 77,5% com uma variação de $\pm 0,013$. Esse teste é apresentado na Figura 5.22.

Figura 5.22 – Médias das medidas de avaliação na segmentação da infecção com a MultiResUNet (Teste 4).

Fonte: Elaborada pelo autor.

Após a realização dos testes, foi possível identificar que o Teste 2 foi o que obteve os melhores resultados em relação à medida IoU, obteve 67,3% com uma variação de $\pm 0,010$. Além disso, o Teste 2 também obteve um bom desempenho na medida Dice, obteve 80,4% com uma variação de $\pm 0,007$, o que indica que foi o melhor teste em segmentar a região da infecção.

Os resultados obtidos na segmentação da infecção COVID-19 no Teste 2 foi inferior aos obtidos no estudo de Tahir *et al.* (2021a) usando a U-Net++ com codificador DenseNet121 que apresentou desempenho com valor do IoU de 83,05% e valor de Dice de 88,21%.

É importante destacar que o Teste 2 não utilizou o filtro CLAHE e nem o aumento de dados. Além disso, o Teste 2 também obteve bons resultados na precisão e recall, o que indica que o modelo está conseguindo classificar corretamente a maioria dos píxeis e está conseguindo detectar a maioria das áreas de interesse.

Outros testes que também obtiveram boas atuações foram o Teste 1 e o Teste 4, embora eles não tenham superado o Teste 2. O Teste 1 obteve 64,3% na medida IoU e 78,2% na medida Dice. Já o Teste 4 obteve 63,2% na medida IoU e 77,5% na medida Dice.

A Tabela 5.10, 5.11 apresenta todos esses resultados obtidos na segmentação da infecção de COVID-19 em imagens de radiografias de tórax usando a MultiResUNET. Cada linha representa um teste diferente descrito anteriormente, enquanto cada coluna apresenta uma média apresentadas nos gráficos anteriores, permitindo uma comparação entre os resultados de cada tarefa.

Tabela 5.10 – Resultados dos testes realizados em segmentação da infecção nas medidas de IoU e Dice.

	IoU	Dice
Teste 1	64% ± 0,041	78,2% ± 0,031
Teste 2	67,3% ± 0,010	80,4% ± 0,007
Teste 3	62,6% ± 0,009	77% ± 0,007
Teste 4	63,2% ± 0,017	77,5% ± 0,013

Tabela 5.11 – Resultados dos testes realizados em segmentação da infecção nas demais medidas.

	Acurácia	Precisão	Recall	F1
Teste 1	93,8% ± 0,008	83,7% ± 0,043	73,7% ± 0,053	78,2% ± 0,031
Teste 2	94,2% ± 0,001	86,6% ± 0,011	75,1% ± 0,020	80,4% ± 0,007
Teste 3	93,9% ± 0,001	86,4% ± 0,011	69,5% ± 0,017	77% ± 0,007
Teste 4	94,2% ± 0,002	89% ± 0,009	68,6% ± 0,025	77,5% ± 0,013

A partir do modelo obtido no Teste 2. Foi calculado o IoU para cada imagem, o IoU é a medida mais confiável quando se trata de segmentação de imagens. Foi verificado se o IoU é igual ou maior que um limiar específico (limiar = 70%) e caso seja, será classificado como segmentação boa, caso contrário classifica como segmentação ruim. Os resultados obtiveram 235 classificados como segmentação boa e 355 como segmentação ruim. Para detalhar mais esses resultados, foi verificado a faixa de resultados. Esses resultados em faixa são apresentados na Tabela 5.12

Tabela 5.12 – Resultados segmentação da infecção por faixas usando a medida IoU.

Faixa de IoU	Número de segmentações
0% — 9,9%	32
10% — 19,9%	24
20% — 29,9%	43
30% — 39,9%	51
40% — 49,9%	56
50% — 59,9%	66
60% — 69,9%	83
70% — 79,9%	94
80% — 89,9%	105
90% — 99,9%	36
100%	0

É possível observar que a maioria das segmentações boas estão na faixa de precisão entre 80% e 89,9%, com 105 ocorrências, seguida da faixa entre 70% e 79,9%, com 94 ocorrências. Isso significa que ainda há espaço para melhoria. Por outro lado, há um número significativo de segmentações ruins, especialmente na faixa entre 0% e 9,9%, com 32 ocorrências. Isso pode indicar a dificuldade de segmentar a área de infecção por COVID-19.

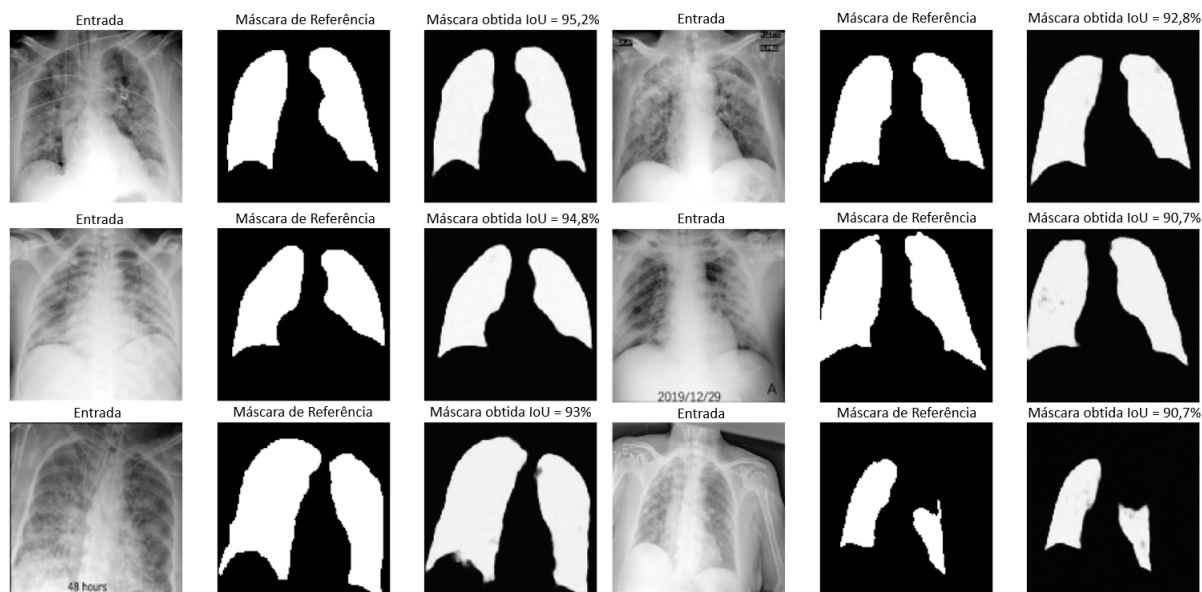
A seguir, são apresentadas figuras que ilustram as segmentações realizadas nas imagens de infecção por COVID-19. Cada Figura apresenta amostras de imagens originais, máscara de referência e máscara obtida com seu valor de IoU, fornecendo uma visão geral do desempenho das segmentações nesta faixa.

Na Figura 5.23 é possível verificar várias segmentações de imagens na faixa de 90% a 99,9% de IoU. É possível observar que a precisão dessas segmentações são altas, com as áreas corretamente segmentadas.

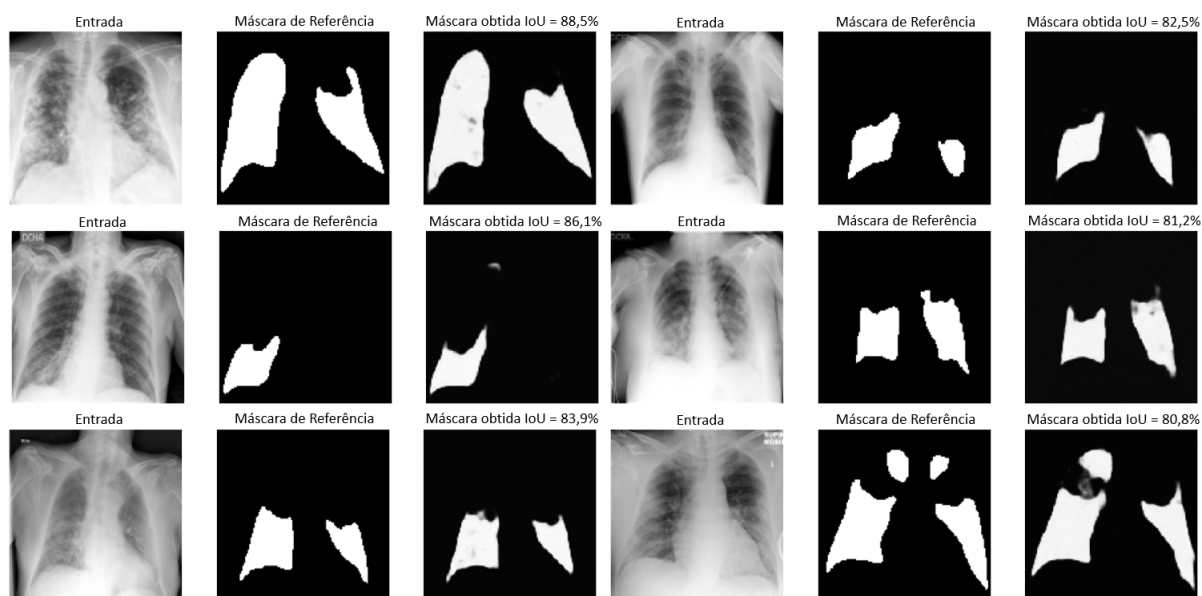
A Figura 5.24 exibe outras segmentações de imagens, mas agora na faixa de 80% a 89,9% de IoU. É possível observar que a precisão dessas segmentações é inferior em relação à Figura 5.23. Há poucas áreas na qual a segmentação não é tão precisa, mas elas são geralmente pequenas, afetando um pouco a qualidade geral da segmentação.

A Figura 5.25 apresenta outras segmentações de imagens, sendo na faixa de 70% a 79,9% de IoU. É possível observar que a maioria das segmentações possuem boas precisões. No entanto, há algumas áreas na qual a segmentação não é tão precisa, como algumas áreas não sendo completamente incluídas na segmentação.

A Figura 5.26 apresenta quando o limiar é entre 60% e 69,9% de IoU, é possível notar que a qualidade da segmentação cai um pouco, há algumas áreas onde a precisão pode ser melhorada. É possível observar que boa parte são segmentadas, no entanto, há algumas

Figura 5.23 – Resultado da segmentação da infecção com a MultiResUNet na faixa de 90% a 99% de IoU.

Fonte: Elaborada pelo autor.

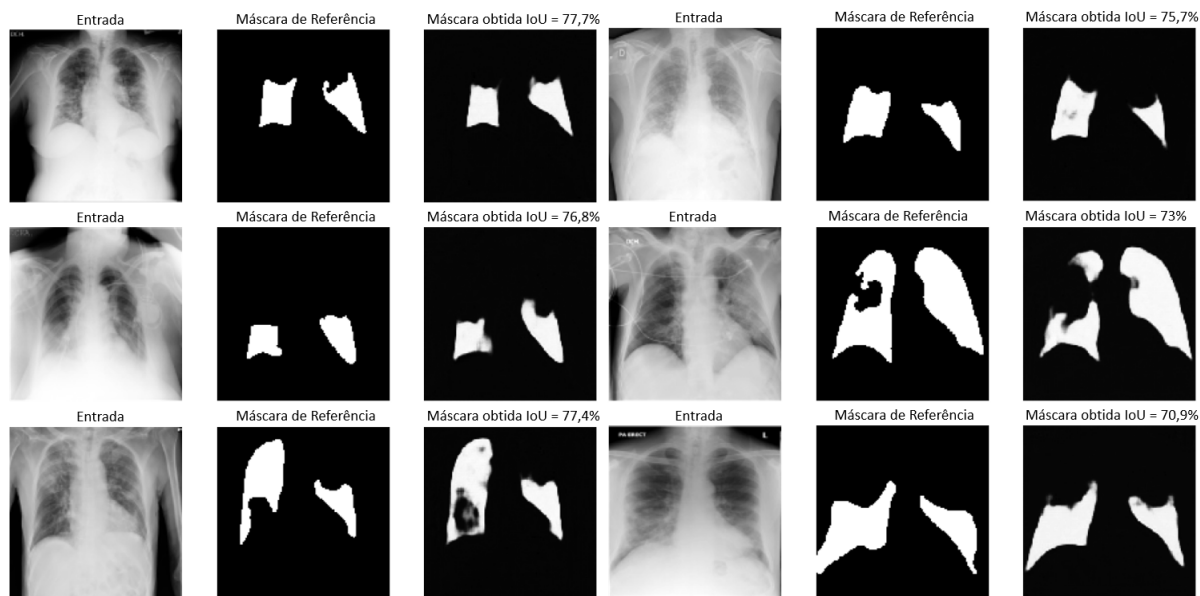
Figura 5.24 – Resultado da segmentação da infecção com a MultiResUNet na faixa de 80% a 89,9% de IoU.

Fonte: Elaborada pelo autor.

áreas que não foram completamente incluídas na segmentação. Isso pode afetar ligeiramente a qualidade geral da segmentação.

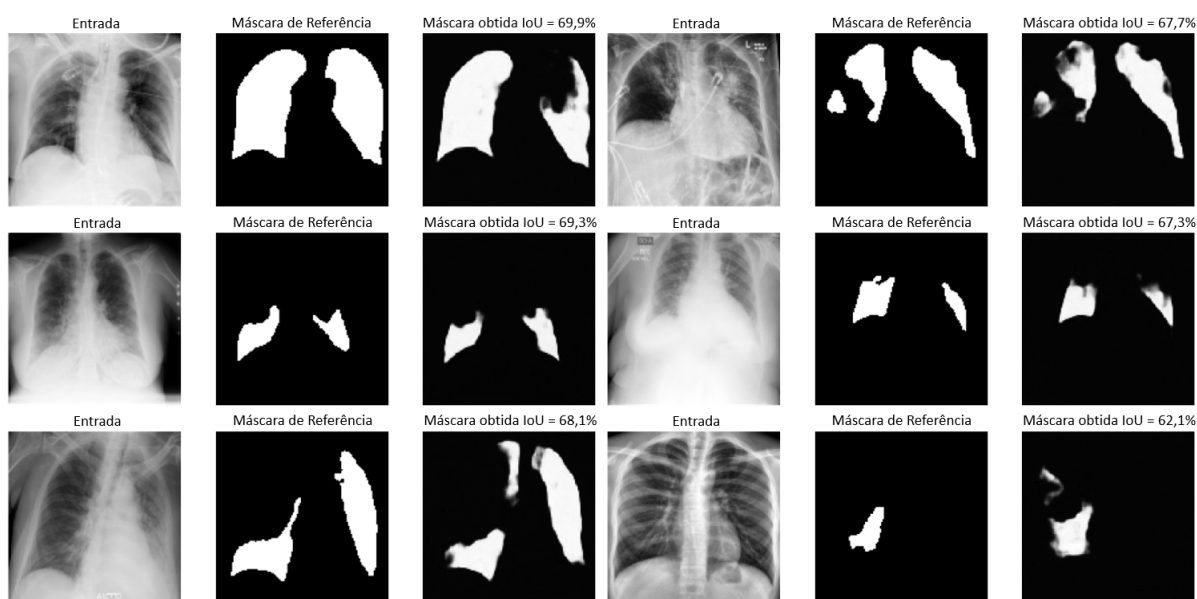
Quando o limiar é entre 50% e 59,9% de IoU, a qualidade da segmentação é intermediária. É possível observar que a maioria das áreas são identificadas e segmentadas corretamente, no entanto, há algumas áreas na qual a precisão pode ser melhorada. A quantidade de áreas não segmentadas corretamente é maior do que na faixa anterior, e isso pode afetar significativamente a qualidade geral da segmentação. Isso pode ser visto na Figura 5.27.

Figura 5.25 – Resultado da segmentação da infecção com a MultiResUNet na faixa de 70% a 79,9% de IoU.



Fonte: Elaborada pelo autor.

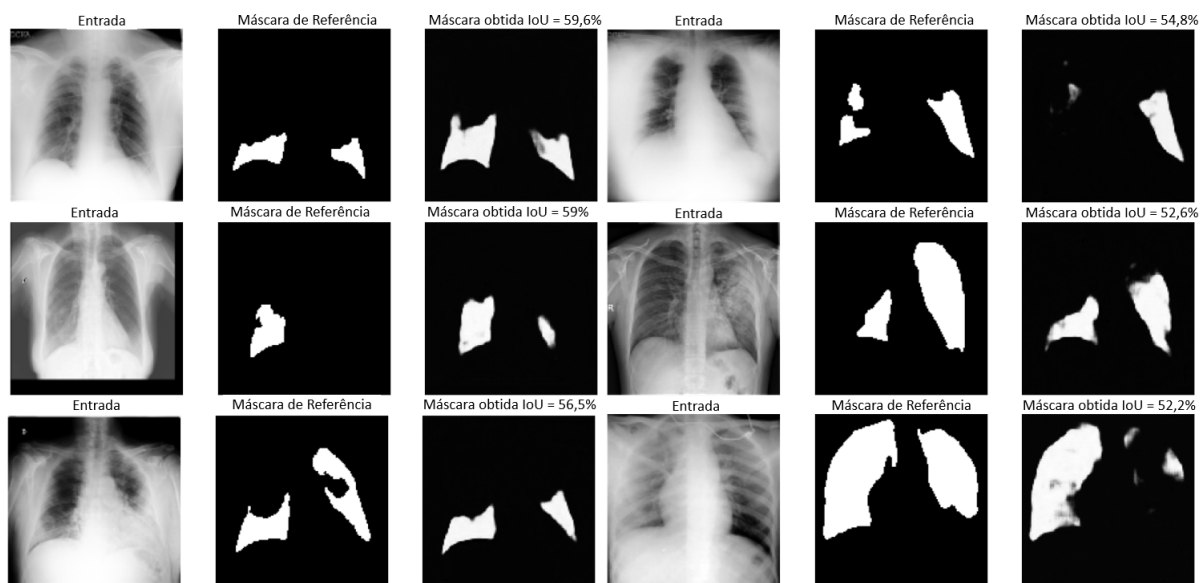
Figura 5.26 – Resultado da segmentação da infecção com a MultiResUNet na faixa de 60% a 69,9% de IoU.



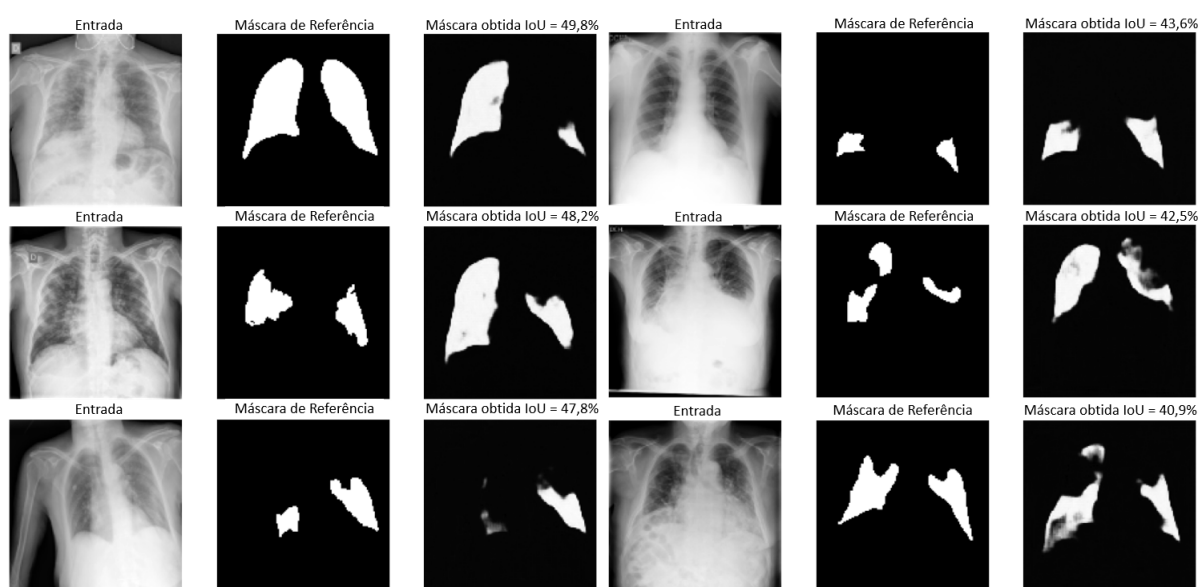
Fonte: Elaborada pelo autor.

Quando o limiar é entre 30% e 49,9% de IoU, a qualidade da segmentação é baixa. É possível observar que há muitas áreas não identificadas e não segmentadas corretamente, e a precisão geral é baixa. As áreas de interesse podem ser difíceis de serem identificadas devido à falta de precisão. A imagem pode ser considerada de baixa qualidade e pouco adequada para análise e interpretação, isso são observados nas Figuras 5.29, 5.28.

Quando o limiar é entre 20% e 29,9% de IoU, a qualidade da segmentação é muito baixa. Há muitas áreas não identificadas e não segmentadas corretamente. As áreas de interesse são

Figura 5.27 – Resultado da segmentação da infecção com a MultiResUNet na faixa de 50% a 59,9% de IoU.

Fonte: Elaborada pelo autor.

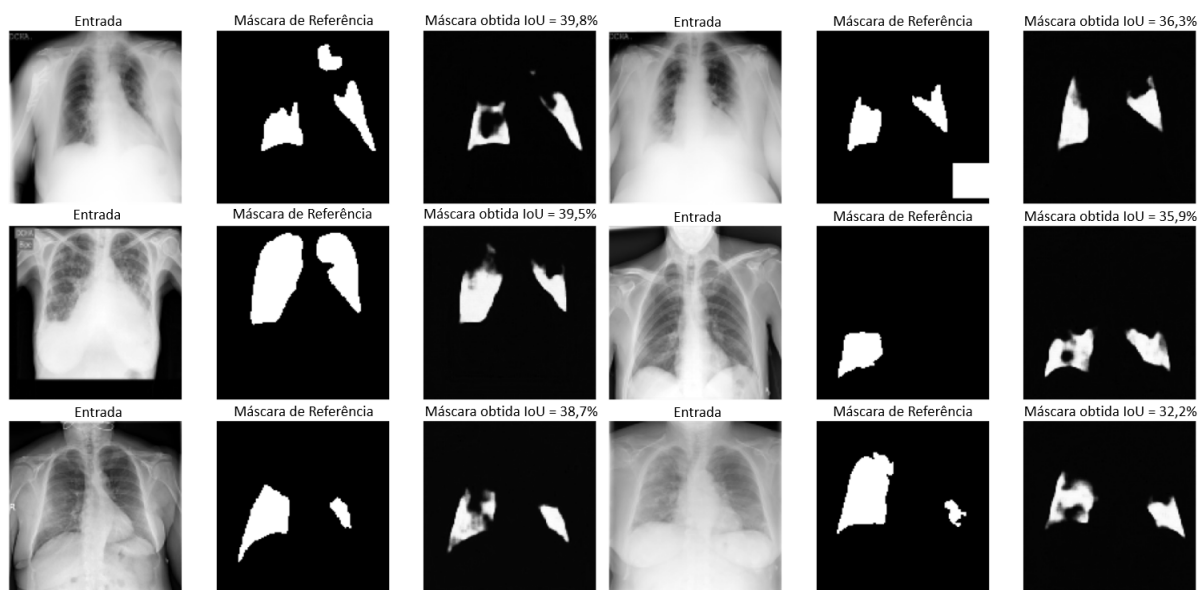
Figura 5.28 – Resultado da segmentação da infecção com a MultiResUNet na faixa de 40% a 49,9% de IoU.

Fonte: Elaborada pelo autor.

segmentadas em poucas partes, deixa desejar bastante, muito baixa qualidade e inadequada para uso. As imagens estão ilustradas na Figura 5.30.

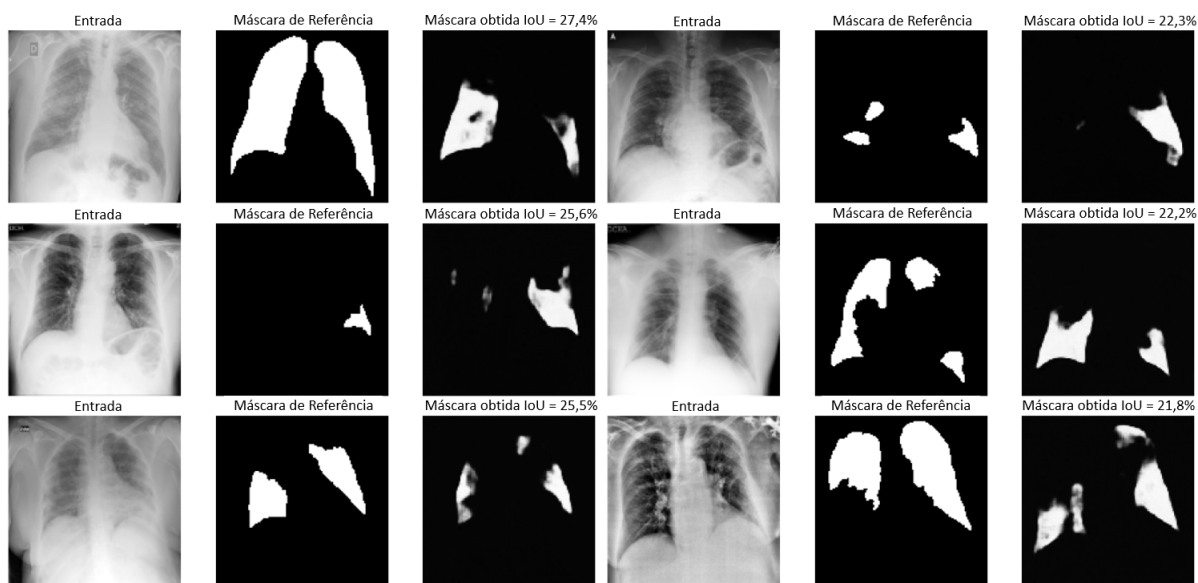
Quando o limiar é entre 10% e 19,9% de IoU, a qualidade da segmentação é ainda pior do que na faixa anterior, isso é ilustrado na Figura 5.31. É possível observar que a maioria das áreas não está sendo identificadas e segmentadas corretamente, e a precisão geral é muito baixa.

Figura 5.29 – Resultado da segmentação da infecção com a MultiResUNet na faixa de 30% a 39,9% de IoU.



Fonte: Elaborada pelo autor.

Figura 5.30 – Resultado da segmentação da infecção com a MultiResUNet na faixa de 20% a 29,9% de IoU.

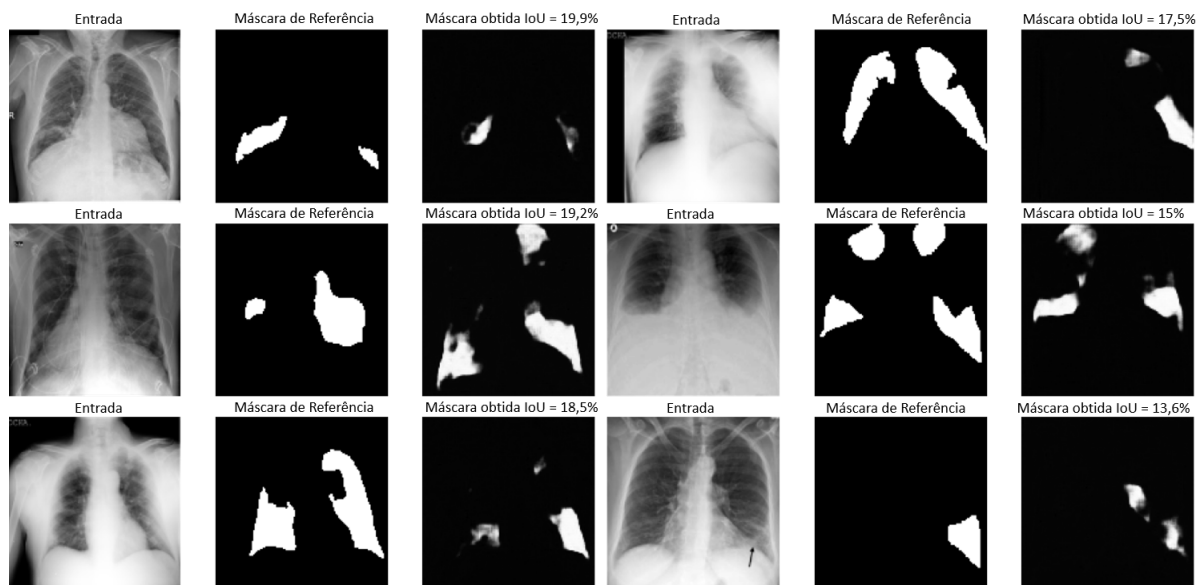


Fonte: Elaborada pelo autor.

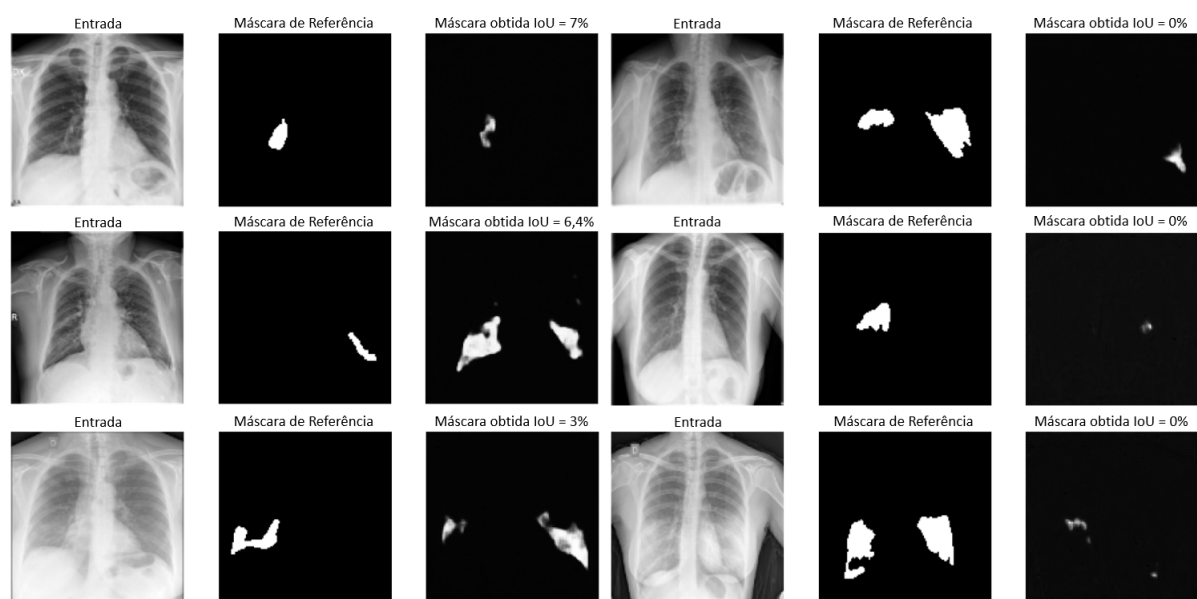
Quando o limiar é abaixo de 10%, a qualidade da segmentação é extremamente baixa. É possível observar que praticamente nenhuma área está sendo segmentada corretamente. A imagem é praticamente inútil como pode ser observado na Figura 5.32.

5.4.2 – Segmentação da Infecção COVID-19 com o LIME

A segmentação da infecção foi realizada utilizando o modelo VGG16 obtido na classificação binária treinado com imagens de COVID-19 e não COVID-19. Utilizou-se o

Figura 5.31 – Resultado da segmentação da infecção com a MultiResUNet na faixa de 10% a 19,9% de IoU.

Fonte: Elaborada pelo autor.

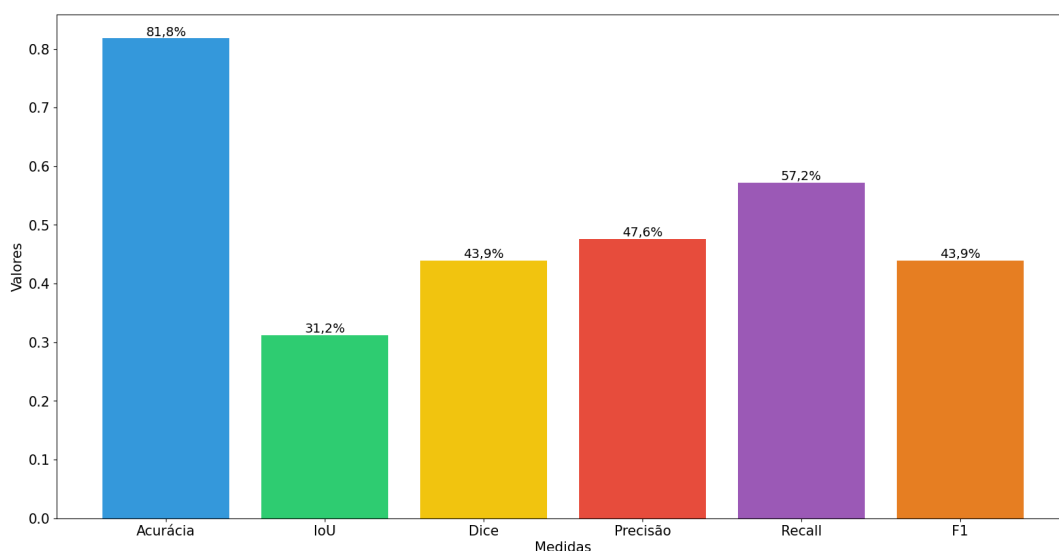
Figura 5.32 – Resultado da segmentação da infecção com a MultiResUNet na faixa de 0% a 9,9% de IoU.

Fonte: Elaborada pelo autor.

LIME a partir do modelo VGG16 para realizar a segmentação. Foi realizada a segmentação utilizando o mesmo conjunto de teste com imagens de COVID-19 que também foi utilizado pela MultiResUNet. O que diferencia no teste com LIME, é que o modelo obtido da VGG16 foi treinada com imagens de pulmão segmentado. Isso melhora a precisão do LIME na segmentação de características específicas da COVID-19. Vale ressaltar que a MultiResUNet foi treinada com imagens que não passaram por segmentação pulmonar.

A Avaliação da segmentação da infecção COVID-19 com LIME utilizou as mesmas medidas utilizadas na avaliação da segmentação obtida pela MultiResUNet, são elas: IoU, Dice, acurácia, recall, precisão e F1. Os resultados são ilustrados na Figura 5.33.

Figura 5.33 – Médias das medidas de avaliação na segmentação da infecção com o LIME.



Fonte: Elaborada pelo autor.

Os valores de cada medida foram: acurácia foi de 81,75%, IoU foi de 31,21%, Dice foi de 43,94%, precisão foi de 47,64%, recall foi de 57,20% e F1 foi de 43,94%. Para entender melhor os resultados, foi realizada uma análise da segmentação obtida pelo LIME, foi calculado as faixas de valores de IoU para verificar em quais faixas possui maior numero de imagens segmentadas com base na medida IoU. Esses resultados por faixa é exibido na Tabela 5.13.

Tabela 5.13 – Resultados segmentação da infecção por faixas usando a medida IoU.

Faixa de IoU	Número de segmentações
0% — 9,9%	104
10% — 19,9%	99
20% — 29,9%	98
30% — 39,9%	92
40% — 49,9%	88
50% — 59,9%	55
60% — 69,9%	25
70% — 79,9%	20
80% — 89,9%	8
90% — 99,9%	1
100%	0

Esses resultados sugerem que a qualidade geral das segmentações obtidas no teste é relativamente baixa. Apenas 29 segmentações (cerca de 5% do total) foram consideradas boas, ou seja, valor do IoU igual ou acima de 70%, enquanto a grande maioria (561 segmentações

ou cerca de 95% do total) foram consideradas ruins. Isso indica haver uma necessidade significativa de melhorias na qualidade das segmentações obtidas pelo LIME.

A distribuição dos resultados com base no valor do IoU mostra que a maioria das segmentações alcançou um valor baixo de IoU, com 104 segmentações (ou cerca de 18% do total) apresentando um valor de IoU entre 0% e 9,9%. Isso significa que essas segmentações têm pouca sobreposição com as máscaras de referência e são provavelmente de baixa qualidade. Além disso, outras 99 segmentações (ou cerca de 17% do total) tiveram um IoU entre 10% e 19,9%, e 98 segmentações (ou cerca de 16% do total) tiveram um IoU entre 20% e 29,9%, o que sugere que a qualidade das segmentações não melhora significativamente nesses intervalos.

Embora haja algumas segmentações com valor do IoU mais altos, como 1 segmentação (ou cerca de 0,2% do total) com IoU entre 90% e 99,9% e nenhum com IoU de 100%, a distribuição dos resultados indica haver uma queda acentuada no número de segmentações à medida que o IoU aumenta. Esses resultados sugerem haver uma necessidade significativa de melhorar a qualidade das segmentações obtidas com o LIME.

A Figura 5.34 apresenta alguns exemplos de segmentações, mas apenas um alcançou a faixa de 90% de IoU. As demais, são na faixa de 80% a 89,9% de IoU. Há poucas áreas na qual a segmentação não é tão precisa, mas são geralmente pequenas, afetando muito pouco a qualidade geral da segmentação. Essas segmentações podem ser confiáveis para uso.

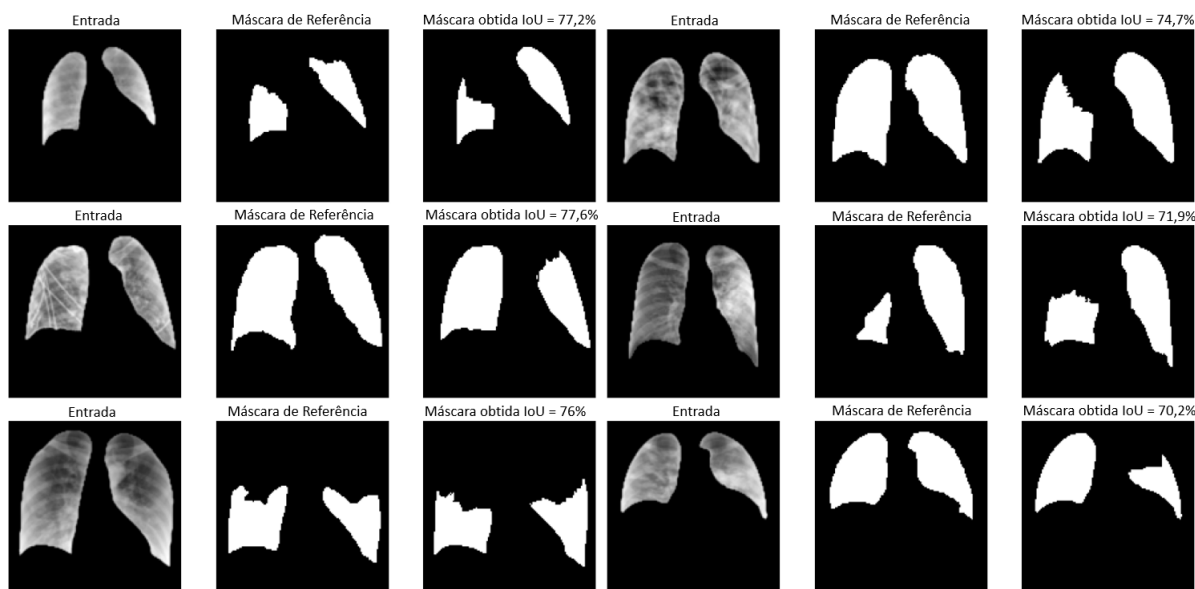
Figura 5.34 – Resultado da segmentação da infecção na faixa de 80% a 99.9% de IoU com LIME.



Fonte: Elaborada pelo autor.

A Figura 5.35 apresenta outros exemplos de segmentações, sendo na faixa de 70% a 79,9% de IoU. É possível observar que há regiões que não foram incluídas na segmentação, mas parte das regiões foram corretamente segmentados.

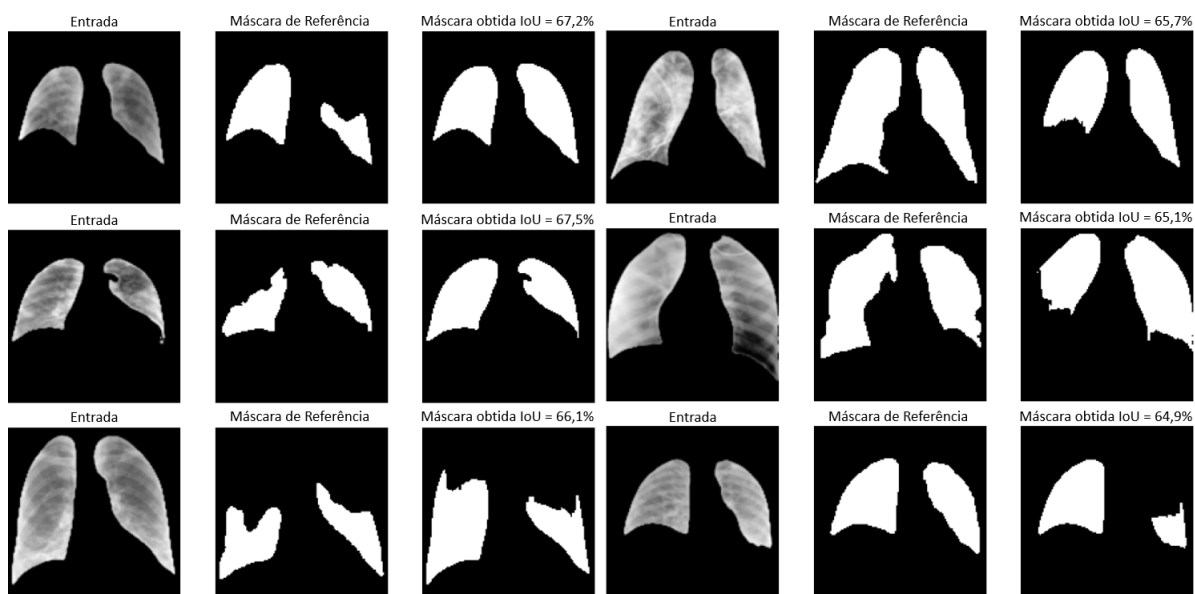
Figura 5.35 – Resultado da segmentação da infecção na faixa de 70% a 79,9% de IoU com LIME.



Fonte: Elaborada pelo autor.

A Figura 5.36 ilustra que, quando o limiar é foi entre 60% e 69,9%, a qualidade da segmentação diminui. Assim como na faixa de 70%, há áreas que não foram segmentadas, impactando na qualidade geral do resultado.

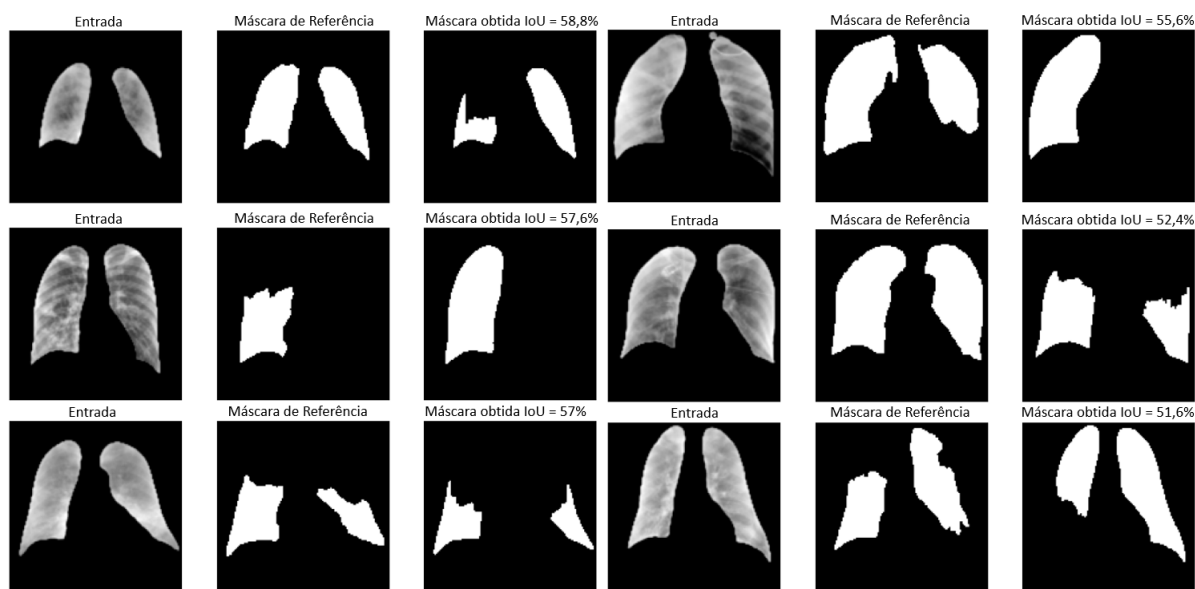
Figura 5.36 – Resultado da segmentação da infecção na faixa de 60% a 69,9% de IoU com LIME.



Fonte: Elaborada pelo autor.

Quando o valor de IoU está entre 50% e 59,9%, a quantidade de áreas não segmentadas corretamente é maior que nas faixas anteriores, e isso pode afetar significativamente a qualidade geral da segmentação. Isso pode ser visto na Figura 5.37.

Figura 5.37 – Resultado da segmentação da infecção na faixa de 50% a 59,9% de IoU com LIME.



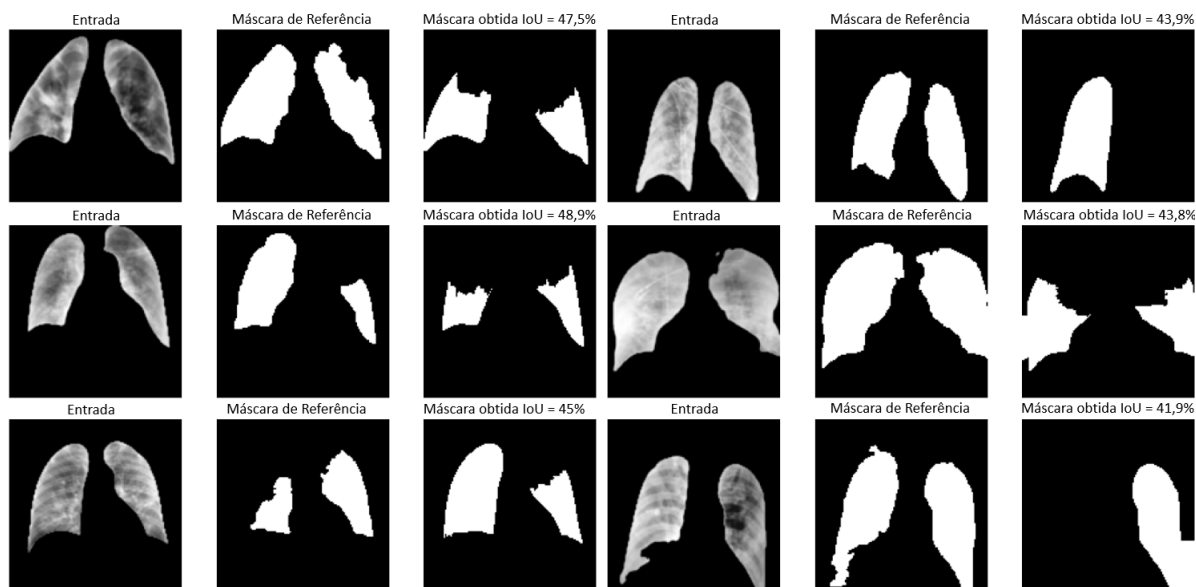
Fonte: Elaborada pelo autor.

Quando a porcentagem o valor de IoU está entre 30% e 49,9%, a qualidade da segmentação é baixa. É perceptível que muitas áreas não são identificadas ou segmentadas corretamente, resultando em uma precisão geral baixa. Além disso, as áreas de interesse podem ser difíceis de identificar devido à imprecisão. As Figuras 5.38 e 5.39 evidenciam que a imagem é de baixa qualidade e inadequada para análise e interpretação.

Quando o valor de IoU é entre 20% e 29,9%, a qualidade da segmentação é muito baixa. Nesse caso, muitas áreas não são identificadas ou segmentadas corretamente, resultando em uma segmentação insuficiente das áreas de interesse. Essa baixa qualidade torna a segmentação inadequada para uso, deixando muito a desejar. As imagens que ilustram essa situação apresentadas na Figura 5.40.

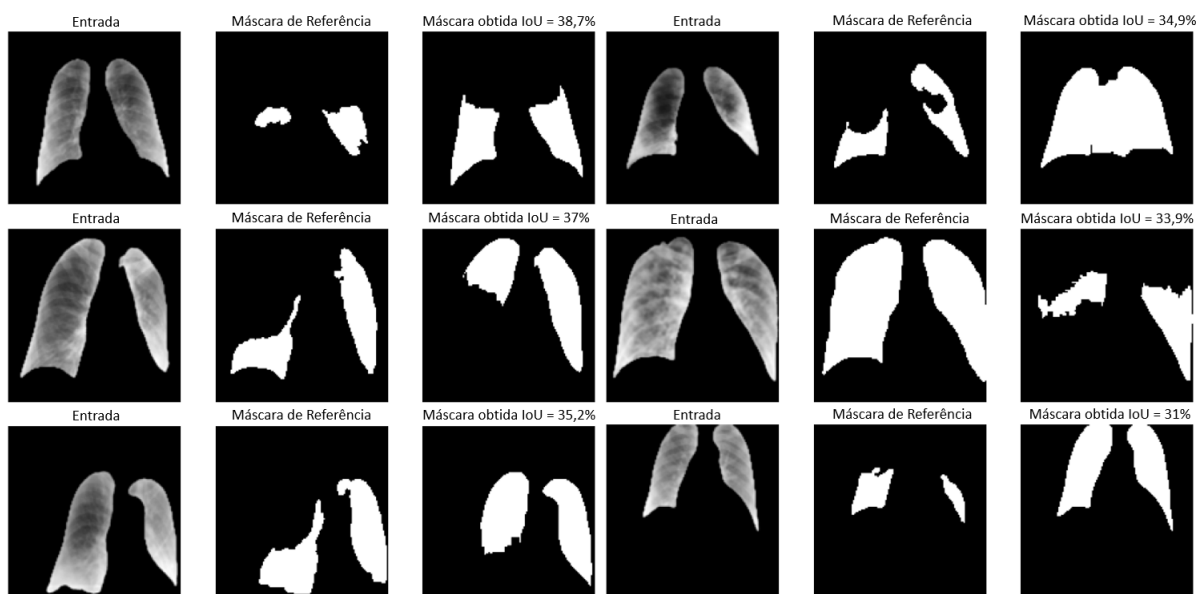
A qualidade da segmentação é pior quando o limiar está entre 10% e 19,9%, como mostrado na Figura 5.41. A maioria das áreas não é identificada e segmentada corretamente. Se o valor de IoU é menor que 10%, a qualidade da segmentação se torna extremamente baixa, com praticamente nenhuma área segmentada corretamente. A imagem se torna inútil, como visto na Figura 5.42.

Figura 5.38 – Resultado da segmentação da infecção na faixa de 40% a 49,9% de IoU com LIME.



Fonte: Elaborada pelo autor.

Figura 5.39 – Resultado da segmentação da infecção na faixa de 30% a 39,9% de IoU com LIME.

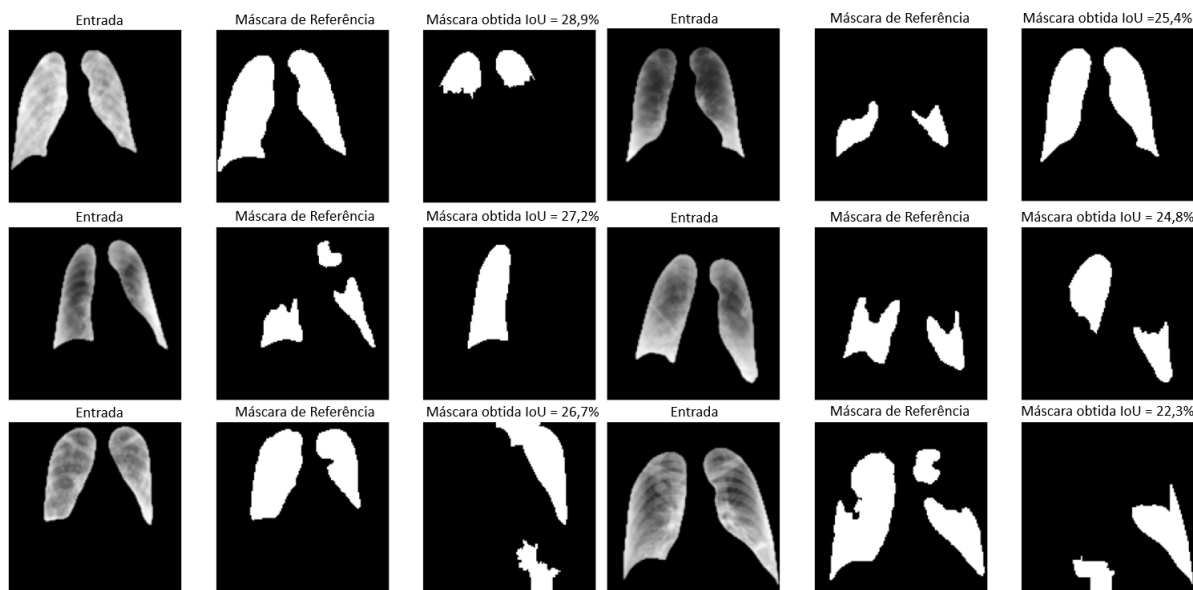


Fonte: Elaborada pelo autor.

5.4.3 – Segmentação da infecção COVID-19 com MultiResUNet vs LIME

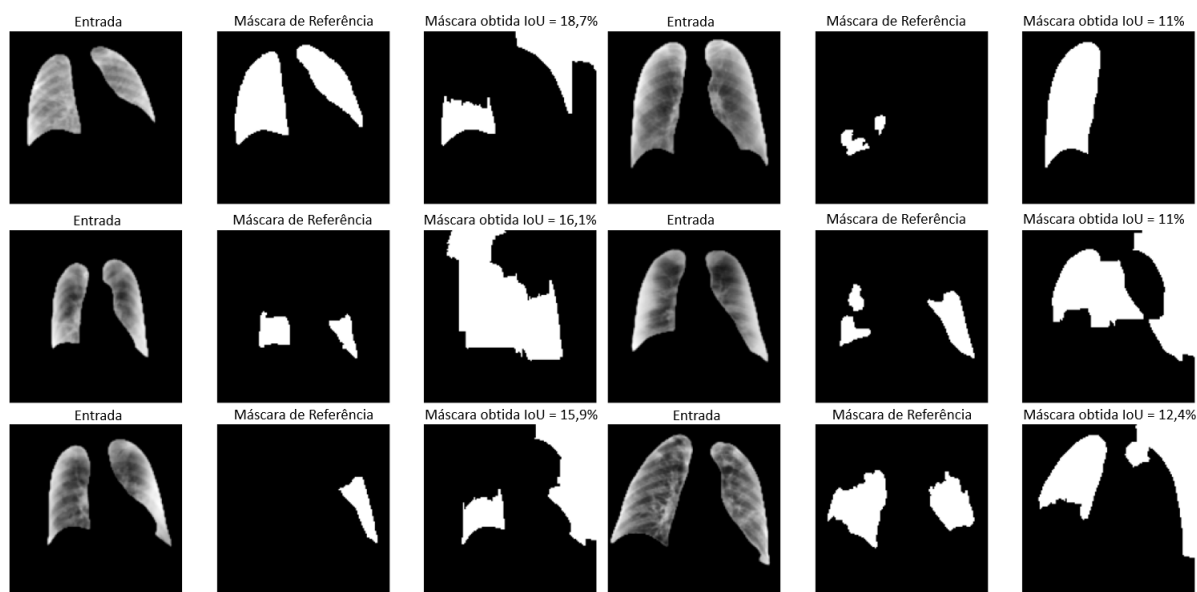
A segmentação da infecção COVID-19 ocorreram de duas maneiras distintas, essas duas formas de segmentação foram apresentadas nas subseções 5.4.1 e 5.4.2, foi realizado uma comparação entre a MultiResUNet e LIME para segmentar a infecção da COVID-19. A Tabela 5.14 apresenta uma comparação do desempenho da segmentação em relação à medida de IoU. A Tabela mostra o número de segmentações em cada faixa de IoU para cada técnica.

Figura 5.40 – Resultado da segmentação da infecção na faixa de 20% a 29,9% de IoU com LIME.



Fonte: Elaborada pelo autor.

Figura 5.41 – Resultado da segmentação da infecção na de 10% a 19,9% de IoU com LIME.

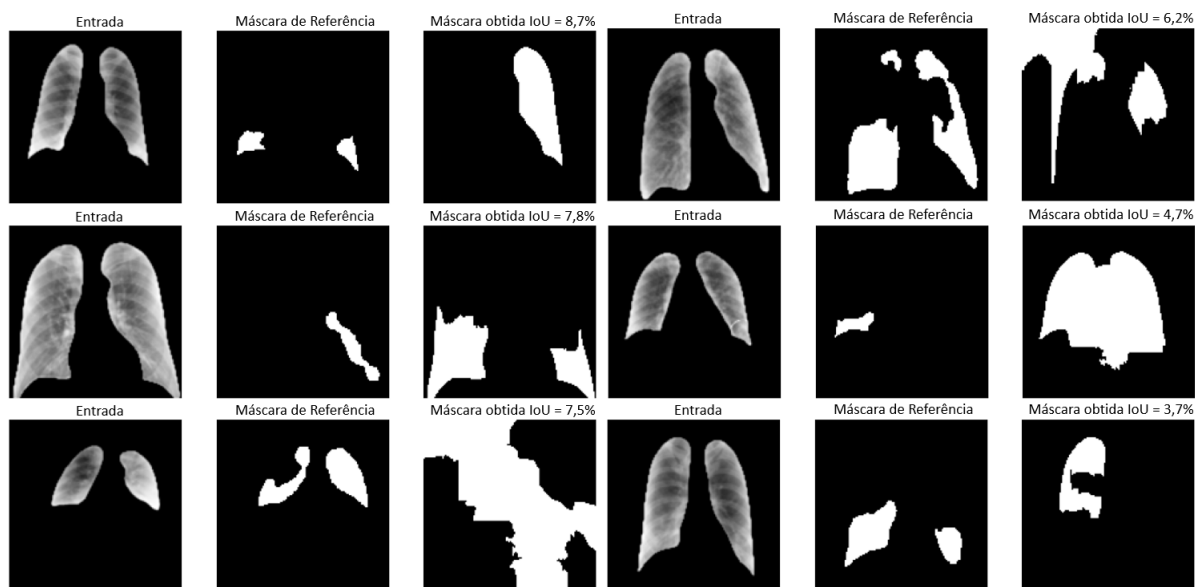


Fonte: Elaborada pelo autor.

Analisando a Tabela, é possível observar que a MultiResUNet obteve mais segmentações com valor de IoU acima de 50%, enquanto o LIME obteve mais segmentações com valor de IoU abaixo de 50%. Isso sugere que a MultiResUNet é mais precisa na tarefa de segmentação de imagens.

Ainda observando a Tabela, é possível notar que a MultiResUNet gerou 105 segmentações com valor de IoU entre 80% e 89,9%, enquanto o LIME gerou apenas 8 segmentações na mesma faixa. Isso reforça a superioridade da MultiResUNet em termos de

Figura 5.42 – Resultado da segmentação da infecção na faixa de 0% a 9,9% de IoU com LIME.



Fonte: Elaborada pelo autor.

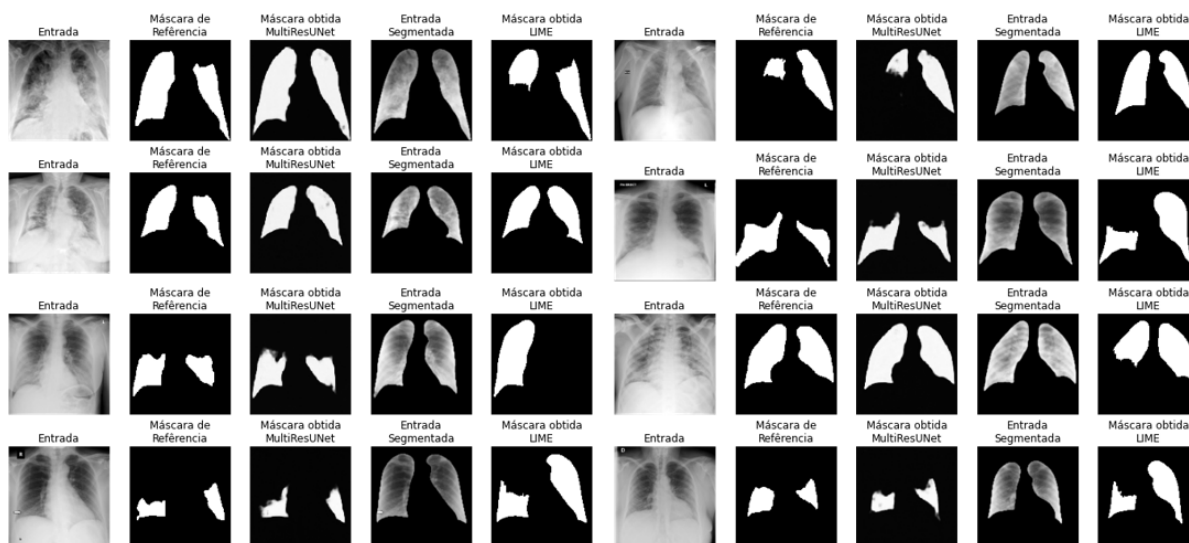
Tabela 5.14 – Resultados segmentação da infecção por faixas usando a medida IoU.

Faixa de IoU	Segmentações MultiResUNet	Segmentações LIME
0% — 9,9%	32	104
10% — 19,9%	24	99
20% — 29,9%	43	98
30% — 39,9%	51	92
40% — 49,9%	56	88
50% — 59,9%	66	55
60% — 69,9%	83	25
70% — 79,9%	94	20
80% — 89,9%	105	8
90% — 99,9%	36	1
100%	0	0

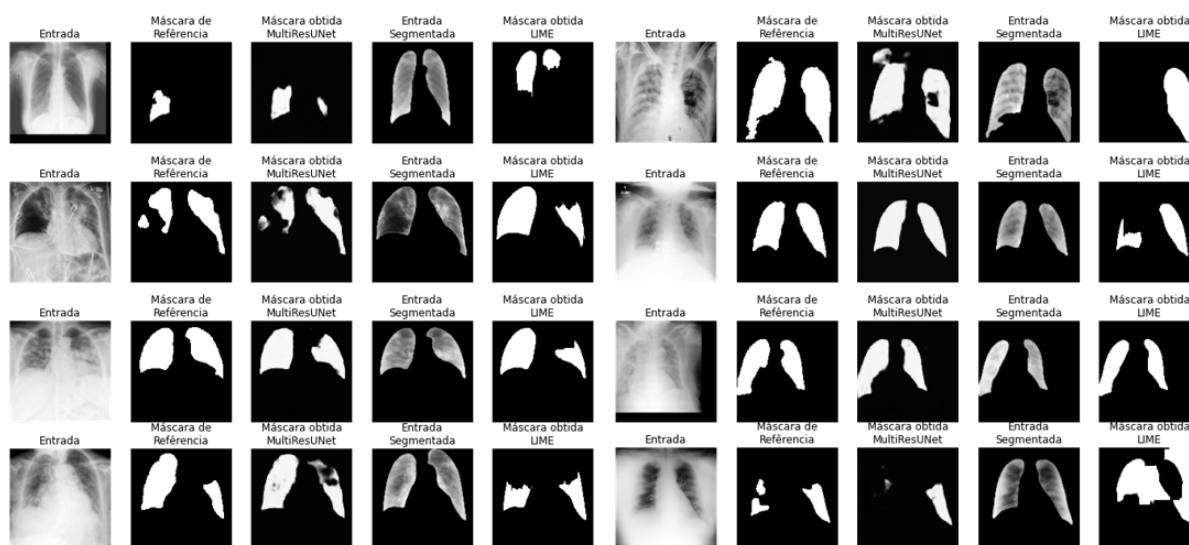
precisão. Além disso, a MultiResUNet gerou 36 segmentações com valor de IoU acima de 90%, enquanto o LIME gerou apenas uma. Isso sugere que a MultiResUNet é mais confiável para gerar segmentações da infecção COVID-19.

Alguns exemplos da segmentação da infecção COVID-19 obtidas por ambos os modelos são ilustrados nas Figuras 5.43 e 5.44. São apresentadas três máscaras para cada exemplo: a máscara de referência, a máscara obtida com a MultiResUNet e a máscara obtida pelo método LIME.

Os resultados obtidos mostraram que, em geral, a máscara obtida pela MultiResUNet foi superior à maioria das máscaras obtidas pelo LIME. No entanto, foi observado que algumas das máscaras obtidas pelo LIME conseguiram produzir resultados bastante similares à máscara de referência. Isso sugere que o LIME possui uma margem de melhoria.

Figura 5.43 – Comparativo da segmentação obtida pelo LIME em imagens com segmentação.

Fonte: Elaborada pelo autor.

Figura 5.44 – Alguns outros exemplos do comparativo da segmentação obtida pelo LIME em imagens com segmentação.

Fonte: Elaborada pelo autor.

É importante destacar que, embora a MultiResUNet seja mais precisa do que o LIME na segmentação de imagens, existem casos que ocorram erros na segmentação. Além disso, em alguns casos, o LIME segmentou partes da imagem que não têm relação com o objeto de interesse, o que pode levar a resultados imprecisos.

Neste trabalho, a rede MultiResUNet foi empregada para segmentação pulmonar em radiografias de tórax. Observou-se que os melhores resultados ocorreram quando não foi utilizada nenhuma técnica de aumento de dados ou filtro CLAHE. A acurácia do modelo foi de 98,4% com variação de $\pm 0,001$, enquanto o valor de IoU foi de 92,39% com variação de $\pm 0,0$. O valor de Dice foi de 96,1% com variação de $\pm 0,001$. O valor da precisão do modelo foi de 97,2% com variação de $\pm 0,004$, o valor do recall foi de 94,9% com variação de $\pm 0,003$. O valor F1 foi de 96,1% com variação de $\pm 0,001$.

As máscaras de segmentação obtidas pela MultiResUNet foram consideradas confiáveis e com qualidade próxima das máscaras de referência. A segmentação pulmonar alcançou faixas de IoU, com 6 máscaras entre as faixas de 60% e 69,9% de IoU, 28 máscaras entre as faixas de 70% e 79,9% de IoU, 242 máscaras entre as faixas de 80% e 89,9% de IoU e 1044 máscaras entre as faixas de 90% e 99,9% de IoU, sem nenhuma máscara nas demais faixas. Esses resultados são promissores e reforçam a eficácia da rede MultiResUNet para a segmentação pulmonar em radiografias de tórax.

Posteriormente, as redes neurais convolucionais pré-treinadas foram utilizadas para classificar imagens de radiografias de tórax. Para isso, foram utilizadas as CNNs: VGG19, VGG16, InceptionV3, DenseNet201, Xception, ResNet50, InceptionResNetV2 e NASNetLarge. O objetivo foi avaliar o desempenho desses modelos e determinar qual deles é o mais eficaz na classificação de radiografias de tórax. As medidas utilizadas para avaliar os modelos foram: acurácia, sensibilidade, especificidade, precisão e F1.

Durante os experimentos alguns modelos se destacaram. A VGG16 se destacou na classificação binária entre COVID-19 e Não COVID-19, obtendo os melhores resultados em nas medidas de desempenho: a acurácia foi de 86,3% com variação $\pm 0,021$, a sensibilidade foi de 81,6% com variação $\pm 0,069$, a especificidade foi de 91,3% $\pm 0,031$, a precisão foi de 91,00% com variação $\pm 0,022$ e F1 de 85,9% com variação $\pm 0,028$ com exceção da sensibilidade, que a

VGG16 obteve 81,6% com variação $\pm 0,069$ contra 89,3% com variação $\pm 0,004$ da ResNet50. Os resultados indicam que os modelos têm dificuldade em identificar casos de COVID-19 e Não COVID-19, nenhum modelo atingiu valores superiores a 92%.

Na classificação multiclases, a VGG16 alcançou os melhores resultados em três das cinco medidas, acurácia foi de 80,9% com variação $\pm 0,010$, sensibilidade foi de 86,57% com variação $\pm 0,032$ e F1 de 90,4% com variação $\pm 0,014$. A maior especificidade foi alcançada pela InceptionResNetV2 com 95,2% com variação $\pm 0,012$. A precisão maior foi pela VGG19 com 95% com variação $\pm 0,015$.

Esses resultados indicam a dificuldade dos modelos em classificar corretamente as imagens. Essa dificuldade pode estar relacionada ao conjunto de dados utilizado, que foi criado a partir da junção de outros conjuntos, conforme informações descritas no banco. Além disso, a alta semelhança entre as características apresentadas por pacientes com COVID-19 e pneumonia pode ter sido um fator adicional.

No caso da COVID-19, a OMS ainda não estabeleceu uma recomendação específica para os modelos de aprendizado de máquina na detecção da doença em radiografias e tomografias de tórax. No entanto, em relação à detecção de tuberculose em radiografias de tórax, a OMS recomenda que os modelos de aprendizado de máquina apresentem uma sensibilidade mínima de 90% e uma especificidade mínima de 70%(WHO, 2014).

Após a classificação de imagens, ocorreu a realização da segmentação da infecção COVID-19 utilizando a MultiResUNet como também usando o LIME. Embora os valores das medidas de desempenho tenham sido inferiores na segmentação da infecção em comparação com os resultados da segmentação pulmonar, o teste 2 da segmentação da infecção apresentou os melhores resultados, sem a utilização de aumento de dados ou aplicação do filtro CLAHE. Esse teste obteve uma acurácia de 94,2%, o valor do IoU foi de 67,3%, o valor de Dice foi de 80,4%, o valor da precisão foi de 86,6%, o valor do recall foi de 75,1% e o valor do F1 foi de 80,4%.

No entanto, as máscaras de segmentação da infecção obtidas pela MultiResUNet apresentaram qualidade inferior quando comparadas às máscaras obtidas na segmentação pulmonar devido à dificuldade em segmentar a infecção COVID-19. Houve uma quantidade significativa de segmentações com IoU abaixo de 70%, indicando a necessidade de melhoria na segmentação da infecção para minimizar essas ocorrências.

Apesar das limitações da MultiResUNet na segmentação da infecção COVID-19, seus resultados ainda são promissores e abrem caminho para futuras melhorias. Em resumo, embora tenham sido encontradas dificuldades, a MultiResUNet apresentou bom desempenho na segmentação em parte das imagens de infecção, indicando a possibilidade de melhoria no futuro.

A segmentação da infecção utilizando a técnica de LIME apresentou desempenho inferior quando comparada à segmentação obtida pela MultiResUNet. Os resultados mostraram uma acurácia de 81,75%, o valor do IoU foi de 31,21%, o valor de Dice foi de 43,94%, o valor da precisão foi de 47,64%, o valor do recall foi de 57,20% e o valor do F1 foi de 43,94%.

A análise dos valores de IoU evidenciou que a grande maioria das segmentações obteve valores abaixo de 70%. Comparando com a segmentação da infecção obtida pela MultiResUNet, a MultiResUNet apresentou menor quantidade de segmentações com IoU abaixo de 70%, o que indica que a técnica foi mais eficaz em segmentar a infecção COVID-19. O LIME apresentou maior dificuldade em segmentar a infecção.

6.1 – Sugestões de Trabalhos Futuros

1. Comparar o desempenho da MultiResUNet com outras redes de segmentação em termos de segmentação pulmonar e infecção COVID-19 em imagens de radiografia de tórax.
2. Investigar a eficácia da segmentação pulmonar e infecção COVID-19 usando a MultiResUNet em diferentes conjuntos de dados de radiografia de tórax.
3. Investigar como diferentes técnicas de pré-processamentos podem melhorar a segmentação pulmonar e a classificação de imagens de radiografia de tórax.
4. Desenvolver um modelo de classificação de radiografia de tórax com um desempenho mais consistente na detecção de COVID-19 em imagens de radiografia de tórax.
5. Explorar a eficácia do uso de redes neurais convolucionais para a detecção de outras doenças pulmonares em imagens de radiografia de tórax.
6. Melhoria do desempenho da rede MultiResUNet: Mesmo com resultados promissores na segmentação pulmonar, ainda é possível buscar maneiras de melhorar o desempenho da rede MultiResUNet.

7. Desenvolvimento de um sistema de apoio à decisão para segmentação pulmonar e a capacidade de alguns modelos pré-treinados para classificar imagens de radiografias em diferentes classes, é possível desenvolver um sistema de apoio à decisão para auxiliar os médicos na interpretação de imagens de radiografia de tórax, especialmente em situações de diagnóstico difícil ou incerto. Esse sistema poderia fornecer sugestões ou opções de tratamento com base na análise automatizada das imagens.

6.2 – Artigos Publicados

1. FARIAS, R. S; OLIVEIRA, K. R. S. ; REGIS, C. D. M; COSTA, C. M ; SEIXAS, J. M. Redes Neurais Convolucionais para Classificação da COVID-19 em Imagens de Raio-x de Tórax. In: Congresso Brasileiro de Inteligência Computacional, 2021, Joinville, SC. Anais do 15. Congresso Brasileiro de Inteligência Computacional, 2021. p. 1.
2. COSTA, C. M ; SEIXAS, J. M ; GUTFILEN, B; SOUZA, S. A. L; OLIVEIRA, S. M; REGIS, C. D. M; FARIAS, R. S; OLIVEIRA, K. R. S. Apoio para Triagem da Infecção por COVID-19 Através de Aprendizagem Profunda Sobre Imagens Segmentadas de Raio-X do Tórax. In: Congresso Brasileiro de Inteligência Computacional, 2021, Joinville, SC. Anais do 15. Congresso Brasileiro de Inteligência Computacional. p. 1.

Referências Bibliográficas

- Abbas, Asmaa, Abdelsamea, Mohammed M, & Gaber, Mohamed Medhat. 2021. Classification of COVID-19 in chest X-ray images using DeTraC deep convolutional neural network. *Applied Intelligence*, **51**, 854–864.
- Abirami, S., & Chitra, P. 2020. Chapter Fourteen - Energy-efficient edge based real-time healthcare support system. *Pages 339–368 of: Raj, Pethuru, & Evangeline, Preetha (eds), The Digital Twin Paradigm for Smarter Systems and Environments: The Industry Use Cases. Advances in Computers*, vol. 117, no. 1. Elsevier.
- Academy, Data Science. 2019. *Deep Learning Book*. <https://www.deeplearningbook.com.br/funcao-de-ativacao/>. Accessed: 06-02-2021.
- Afridi, Muhammad Jamal, Ross, Arun, & Shapiro, Erik M. 2018. On automated source selection for transfer learning in convolutional neural networks. *Pattern recognition*, **73**, 65–75.
- Aggarwal, Priya, Mishra, Narendra Kumar, Fatimah, Binish, Singh, Pushpendra, Gupta, Anubha, & Joshi, Shiv Dutt. 2022. COVID-19 image classification using deep learning: Advances, challenges and opportunities. *Computers in Biology and Medicine*, 105350.
- Ai, Tao, Yang, Zhenlu, Hou, Hongyan, Zhan, Chenao, Chen, Chong, Lv, Wenzhi, Tao, Qian, Sun, Ziyong, & Xia, Liming. 2020. Correlation of chest CT and RT-PCR testing in coronavirus disease 2019 (COVID-19) in China: a report of 1014 cases. *Radiology*.
- Al-Antari, Mugahed A, Han, Seung-Moo, & Kim, Tae-Seong. 2020. Evaluation of deep learning detection and classification towards computer-aided diagnosis of breast lesions in digital X-ray mammograms. *Computer methods and programs in biomedicine*, **196**, 105584.

- Albahli, Saleh, & Albattah, Waleed. 2020. Deep Transfer Learning for COVID-19 Prediction: Case Study for Limited Data Problems. *Current medical imaging*.
- Alom, Md Zahangir, Rahman, MM, Nasrin, Mst Shamima, Taha, Tarek M, & Asari, Vijayan K. 2020. Covid_mtnet: Covid-19 detection with multi-task deep learning approaches. *arXiv preprint arXiv:2004.03747*.
- Araújo, Flávio HD, Carneiro, AC, Silva, Romuere RV, Medeiros, Fátima NS, & Ushizima, Daniela M. 2017. Redes neurais convolucionais com tensorflow: Teoria e prática. *SOCIEDADE BRASILEIRA DE COMPUTAÇÃO. III Escola Regional de Informática do Piauí. Livro Anais-Artigos e Minicursos*, **1**, 382–406.
- Arias-Garzón, Daniel, Alzate-Grisales, Jesús Alejandro, Orozco-Arias, Simon, Arteaga-Arteaga, Harold Brayan, Bravo-Ortiz, Mario Alejandro, Mora-Rubio, Alejandro, Saborit-Torres, Jose Manuel, Serrano, Joaquim Ángel Montell, de la Iglesia Vayá, Maria, Cardona-Morales, Oscar, *et al.* 2021. COVID-19 detection in X-ray images using convolutional neural networks. *Machine Learning with Applications*, **6**, 100138.
- Baratella, Elisa, Crivelli, Paola, Marrocchio, Cristina, Bozzato, Alessandro Marco, Vito, Andrea De, Madeddu, Giordano, Saderi, Laura, Confalonieri, Marco, Tenaglia, Luigi, & Cova, Maria Assunta. 2020. Severity of lung involvement on chest X-rays in SARS-coronavirus-2 infected patients as a possible tool to predict clinical progression: an observational retrospective analysis of the relationship between radiological, clinical, and laboratory data. *Jornal Brasileiro de Pneumologia*, **46**(5).
- Bhadoria, Riddhi Kumari, Saha, Jayita, Biswas, Suparna, & Chowdhury, Chandreyee. 2021. IoT-based location-aware smart healthcare framework with user mobility support in normal and emergency scenario: a comprehensive survey. *Healthcare Paradigms in the Internet of Things Ecosystem*, 137–161.
- Bhadra, Rajarshi, & Kar, Subhajit. 2020. Covid Detection from CXR Scans using Deep Multi-layered CNN. *Pages 214–218 of: 2020 IEEE Bombay Section Signature Conference (IBSSC)*. IEEE.
- Brahme, Anders. 2014. *Comprehensive biomedical physics*. Newnes.

- Ceccon, Denny. 2020. *Funções de ativação: definição, características e quando usar cada uma.* <https://iaexpert.academy/2020/05/25/funcoes-de-ativacao-definicao-caracteristicas-e-quando-usar-cada-uma/>. Accessed: 30-04-2021.
- Center, Math Works Helper. 2019. *NASNetLarge.* <https://la.mathworks.com/>. Accessed: 30-04-2021.
- Chamorro, E Martínez, Tascón, A Díez, Sanz, L Ibáñez, Vélez, S Ossaba, & Nacenta, S Borrue. 2021. Radiologic diagnosis of patients with COVID-19. *Radiología (English Edition)*, **63**(1), 56–73.
- Chamseddine, Ekram, Mansouri, Nesrine, Soui, Makram, & Abed, Mourad. 2022. Handling class imbalance in COVID-19 chest X-ray images classification: Using SMOTE and weighted loss. *Applied Soft Computing*, **129**, 109588.
- Chauhan, Tavishee, Palivela, Hemant, & Tiwari, Sarveshmani. 2021. Optimization and fine-tuning of DenseNet model for classification of COVID-19 cases in medical imaging. *International Journal of Information Management Data Insights*, **1**(2), 100020.
- Chen, Jui-Chi, Lan, Tzu-Peng, Lian, Zhen-You, & Chuang, Cheng-Hung. 2021. A Study of Intervertebral Disc Segmentation Based on Deep Learning. *Pages 85–87 of: 2021 IEEE 4th International Conference on Knowledge Innovation and Invention (ICKII)*. IEEE.
- Cheng, Shao-Chung, Chang, Yuan-Chia, Chiang, Yu-Long Fan, Chien, Yu-Chan, Cheng, Mingte, Yang, Chin-Hua, Huang, Chia-Husn, & Hsu, Yuan-Nian. 2020. First case of Coronavirus Disease 2019 (COVID-19) pneumonia in Taiwan. *Journal of the Formosan Medical Association*, **119**(3), 747–751.
- Chollet, François. 2017a. Xception: Deep learning with depthwise separable convolutions. *Pages 1251–1258 of: Proceedings of the IEEE conference on computer vision and pattern recognition.*
- Chollet, François. 2017b. *Deep Learning with Python*. Manning Publications.
- Choudhury, Manabendra D., Atteya, Walid A., Dahal, Keshav, Chetia, Pankaj, Choudhury, Karabi D., & Paradkar, Anant. 2018. Chapter 11 - Virtual Screening of Phytochemicals.

- Pages 301–334 of:* Sarker, Satyajit D., & Nahar, Lutfun (eds), *Computational Phytochemistry*. Elsevier.
- Chowdhury, Muhammad EH, Rahman, Tawsifur, Khandakar, Amith, Mazhar, Rashid, Kadir, Muhammad Abdul, Mahbub, Zaid Bin, Islam, Khandakar Reajul, Khan, Muhammad Salman, Iqbal, Atif, Al Emadi, Nasser, *et al.* 2020. Can AI help in screening viral and COVID-19 pneumonia? *IEEE Access*, **8**, 132665–132676.
- Costa, Cecília, Seixas, José, Gutflen, Bianca, Souza, Sérgio, Medeiros, Susie, Régis, Carlos, Farias, Ricardo, & Oliveira, Kaique. 2021. Apoio para Triagem da Infecção por COVID-19 Através de Aprendizagem Profunda Sobre Imagens Segmentadas de Raio-X do Tórax. *Pages 1–7 of:* Bastos Filho, Carmelo José Albanez, Siqueira, Hugo Valadares, Ferreira, Danton Diego, Bertol, Douglas Wildgrube, & de Oliveira, Roberto Célio Limão (eds), *Anais do 15º Congresso Brasileiro de Inteligência Computacional*. Joinville, SC: SBIC.
- DeepAI. 2020. *What is the Jaccard Index?* <https://deepai.org/machine-learning-glossary-and-terms/jaccard-index/>. Accessed: 30-04-2021.
- Degerli, Aysen. 2021. *QaTa-COV19 Dataset*. <https://www.kaggle.com/datasets/aysendegerli/qatacov19-dataset>. Acesso em: 10 de abril de 2022.
- Degerli, Aysen, Ahishali, Mete, Yamac, Mehmet, Kiranyaz, Serkan, Chowdhury, Muhammad EH, Hameed, Khalid, Hamid, Tahir, Mazhar, Rashid, & Gabbouj, Moncef. 2021. COVID-19 infection map generation and detection from chest X-ray images. *Health information science and systems*, **9**(1), 15.
- DiGregorio, Justin, Arezza, Giordano, Gibicar, Adam, Moody, Alan R, Tyrrell, Pascal N, & Khademi, April. 2021. Intracranial volume segmentation for neurodegenerative populations using multicentre flair MRI. *Neuroimage: Reports*, **1**(1), 100006.
- Dilshad, Sara, Singh, Nikhil, Atif, M, Hanif, Atif, Yaqub, Nafeesah, Farooq, WA, Ahmad, Hijaz, Chu, Yu-ming, & Masood, Muhammad Tamoor. 2021. Automated image classification of chest X-rays of COVID-19 using deep transfer learning. *Results in Physics*, **28**, 104529.
- Elfiky, Abdo A. 2020. Ribavirin, Remdesivir, Sofosbuvir, Galidesivir, and Tenofovir against SARS-CoV-2 RNA dependent RNA polymerase (RdRp): A molecular docking study. *Life sciences*, 117592.

- Elhamraoui, Zahra. 2020. *InceptionResNetV2 Simple Introduction*. <https://medium.com/@zahraelhamraoui1997/inceptionresnetv2-simple-introduction-9a2000edcdb6/>. Accessed: 27-05-2021.
- Fang, Yicheng, Zhang, Huangqi, Xie, Jicheng, Lin, Minjie, Ying, Lingjun, Pang, Peipei, & Ji, Wenbin. 2020. Sensitivity of chest CT for COVID-19: comparison to RT-PCR. *Radiology*.
- Fogante, Marco, Cavagna, Enrico, & Rinaldi, Giovanni. 2021. COVID-19 follow-up: chest X-ray findings with clinical and radiological relationship three months after recovery. *Radiography*.
- Freitas, André Ricardo Ribas, Medeiros, Nicole Montenegro de, Frutuoso, Livia Carla Vinhal, Beckedorff, Otto Albuquerque, Martin, Lucas Mariscal Alves de, Coelho, Marcela Montenegro de Medeiros, Freitas, Giovanna Gimenez Souza de, Lemos, Daniele Rocha Queiróz, & Cavalcanti, Luciano Pamplona de Góes. 2020. Tracking excess deaths associated with the COVID-19 epidemic as an epidemiological surveillance strategy—preliminary results of the evaluation of six Brazilian capitals. *Revista da Sociedade Brasileira de Medicina Tropical*, **53**.
- Ghofrani, Fatemeh, Behnam, Hamid, & Motlagh, Hamid Didari Khamseh. 2020. Liver Segmentation in CT Images Using Deep Neural Networks. *Pages 1–6 of: 2020 28th Iranian Conference on Electrical Engineering (ICEE)*. IEEE.
- Goldberg, Mateus Arnaud Santos de Sousa. 2021. Análise de técnicas de compressão em redes neurais profundas por poda em dataset de imagens.
- Guidotti, Riccardo, Monreale, Anna, Ruggieri, Salvatore, Turini, Franco, Giannotti, Fosca, & Pedreschi, Dino. 2018. A survey of methods for explaining black box models. *ACM Computing Surveys (CSUR)*, **51**(5), 1–42.
- Guo, Yanhui, & Ashour, Amira S. 2019. Neutrosophic sets in dermoscopic medical image segmentation. *Pages 229–243 of: Neutrosophic Set in Medical Image Analysis*. Elsevier.
- Gupta, Anunay, Gupta, Shreyansh, Katarya, Rahul, *et al.* 2021. Instacovnet-19: A deep learning classification model for the detection of covid-19 patients using chest x-ray. *Applied Soft Computing*, **99**, 106859.

- Hajabdollahi, Mohsen, Esfandiarpour, Reza, Sabeti, Elyas, Karimi, Nader, Soroushmehr, SM Reza, & Samavi, Shadrokh. 2020. Multiple abnormality detection for automatic medical image diagnosis using bifurcated convolutional neural network. *Biomedical Signal Processing and Control*, **57**, 101792.
- Han, Zhimeng, Jian, Muwei, & Wang, Gai-Ge. 2022. ConvUNeXt: An efficient convolution neural network for medical image segmentation. *Knowledge-Based Systems*, **253**, 109512.
- Haykin, Simon. 2001. *Redes neurais: princípios e prática*. 900 p.
- Hemdan, Ezz El-Din, Shouman, Marwa A, & Karar, Mohamed Esmail. 2020. Covidx-net: A framework of deep learning classifiers to diagnose covid-19 in x-ray images. *arXiv preprint arXiv:2003.11055*.
- Hertel, Robert, & Benlamri, Rachid. 2022. A deep learning segmentation-classification pipeline for x-ray-based covid-19 diagnosis. *Biomedical Engineering Advances*, 100041.
- Huang, Gao, Liu, Zhuang, Van Der Maaten, Laurens, & Weinberger, Kilian Q. 2017. Densely connected convolutional networks. *Pages 4700–4708 of: Proceedings of the IEEE conference on computer vision and pattern recognition*.
- Huang, Mei-Ling, & Wu, Yi-Zhen. 2022. Semantic segmentation of pancreatic medical images by using convolutional neural network. *Biomedical Signal Processing and Control*, **73**, 103458.
- Ibtehaz, Nabil, & Rahman, M Sohel. 2020. MultiResUNet: Rethinking the U-Net architecture for multimodal biomedical image segmentation. *Neural Networks*, **121**, 74–87.
- Iranpoor, Rasool, Mahboob, Amir Soltany, Shahbandegan, Shakiba, & Baniasadi, Nasrin. 2020. Skin lesion segmentation using convolutional neural networks with improved U-Net architecture. *Pages 1–5 of: 2020 6th Iranian Conference on Signal Processing and Intelligent Systems (ICSPIS)*. IEEE.
- Jiao, Zhicheng, Choi, Ji Whae, Halsey, Kasey, Tran, Thi My Linh, Hsieh, Ben, Wang, Dongcui, Eweje, Feyisope, Wang, Robin, Chang, Ken, Wu, Jing, *et al.* 2021. Prognostication of patients with COVID-19 using artificial intelligence based on chest x-rays and clinical data: a retrospective study. *The Lancet Digital Health*, **3**(5), e286–e294.

- Jordan, Jeremy. 2018. *Evaluating image segmentation models*. <https://www.jeremyjordan.me/evaluating-image-segmentation-models/>. Accessed: 23-03-2021.
- Kathamuthu, Nirmala Devi, Subramaniam, Shanthi, Le, Quynh Hoang, Muthusamy, Suresh, Panchal, Hitesh, Sundararajan, Suma Christal Mary, Alrubaie, Ali Jawad, & Zahra, Musaddak Maher Abdul. 2023. A deep transfer learning-based convolution neural network model for COVID-19 detection using computed tomography scan images for medical applications. *Advances in Engineering Software*, **175**, 103317.
- Khan, Asif Iqbal, Shah, Junaid Latief, & Bhat, Mohammad Mudasar. 2020. CoroNet: A deep neural network for detection and diagnosis of COVID-19 from chest x-ray images. *Computer Methods and Programs in Biomedicine*, **196**, 105581.
- Khan, Irfan Ullah, & Aslam, Nida. 2020. A deep-learning-based framework for automated diagnosis of COVID-19 using X-ray images. *Information*, **11**(9), 419.
- Kingma, Diederik P, & Ba, Jimmy. 2014. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.
- Kucirka, Lauren M, Lauer, Stephen A, Laeyendecker, Oliver, Boon, Denali, & Lessler, Justin. 2020. Variation in false-negative rate of reverse transcriptase polymerase chain reaction-based SARS-CoV-2 tests by time since exposure. *Annals of internal medicine*, **173**(4), 262–267.
- Kusakunniran, Worapan, Karnjanapreechakorn, Sarattha, Siriapisith, Thanongchai, Borwarnginn, Punyanuch, Sutassananon, Krittanat, Tongdee, Trongtum, & Saiviroonporn, Pairash. 2021. COVID-19 detection and heatmap generation in chest x-ray images. *Journal of Medical Imaging*, **8**(S1), 014001.
- Lakhani, Paras, & Sundaram, Baskaran. 2017. Deep learning at chest radiography: automated classification of pulmonary tuberculosis by using convolutional neural networks. *Radiology*, **284**(2), 574–582.
- Leão, Pedro Paulo de Souza, de Sousa Freire, Natalia, Pinto, Rafael Albuquerque, Pinto, Kely Maciel Braule, Giusti, Rafael, & dos Santos, Eulanda Miranda. 2021. Detecção de Covid-19 em Imagens de Raio-x Utilizando Redes Convolucionais. *Journal of Health Informatics*, **12**.

- LeCun, Yann, Bengio, Yoshua, & Hinton, Geoffrey. 2015. Deep learning. *nature*, **521**(7553), 436–444.
- Liang, Yuanhao, He, Ran, Li, Yongshuai, & Wang, Zhili. 2019. Simultaneous segmentation and classification of breast lesions from ultrasound images using mask R-CNN. *Pages 1470–1472 of: 2019 IEEE International Ultrasonics Symposium (IUS)*. IEEE.
- Lin, Hui, Zhang, Benjian, Guo, Xin, Guo, Dong, Jing, Jia, Huiyuan, Luo, Yumeng, Zhao, Hu, Min, & Ren, Fuji. 2021. An Automatic Method for Brain Tumors Segmentation Based on Deep Convolutional Neural Network. *Pages 1–8 of: 2021 IEEE International Conference on Medical Imaging Physics and Engineering (ICMIPE)*. IEEE.
- Liu, Qiong, & Wu, Ying. 2012. Supervised Learning. 01.
- Liu, Yu Han. 2018. Feature extraction and image recognition with convolutional neural networks. *Page 062032 of: Journal of Physics: Conference Series*, vol. 1087. IOP Publishing.
- Lou, Ange, Guan, Shuyue, Kamona, Nada, & Loew, Murray. 2019. Segmentation of Infrared Breast Images Using MultiResUnet Neural Networks. *Pages 1–6 of: 2019 IEEE Applied Imagery Pattern Recognition Workshop (AIPR)*. IEEE.
- Luo, Lei, Liu, Dan, Liao, Xin-long, Wu, Xian-bo, Jing, Qin-long, Zheng, Jia-zhen, Liu, Fang-hua, Yang, Shi-gui, Bi, Bi, Li, Zhi-hao, *et al.* 2020. Modes of contact and risk of transmission in covid-19: a prospective cohort study 4950 close contact persons in guangzhou of China.
- Maharjan, Jenish, Calvert, Jacob, Pellegrini, Emily, Green-Saxena, Abigail, Hoffman, Jana, McCoy, Andrea, Mao, Qingqing, & Das, Ritankar. 2021. Application of deep learning to identify COVID-19 infection in posteroanterior chest X-rays. *Clinical imaging*, **80**, 268–273.
- Maior, Caio BS, Santana, João MM, Lins, Isis D, & Moura, Márcio JC. 2021. Convolutional neural network model based on radiological images to support COVID-19 diagnosis: Evaluating database biases. *Plos one*, **16**(3), e0247839.
- Nour, Majid, Cömert, Zafer, & Polat, Kemal. 2020. A novel medical diagnosis model for COVID-19 infection detection based on deep features and Bayesian optimization. *Applied Soft Computing*, **97**, 106580.

- Oliveira, Bruno, Ziviani, Hugo, Oliveira, Juscelino, Viegas, Alan, & Calvo, Daniel. 2021. Suporte para Diagnóstico de COVID-19 por Meio de Classificação Automática de Imagens de Raio-x e Modelos Explicáveis. *Pages 1–8 of: Anais do 15° Congresso Brasileiro de Inteligência Computacional*. Sociedade Brasileira de Inteligência Computacional (SBIC).
- OMS. 2023. *WHO Coronavirus (COVID-19) Dashboard*. <https://covid19.who.int/>. Accessed: 09-04-2023.
- Rahman, Tawsifur, Khandakar, Amith, Qiblawey, Yazan, Tahir, Anas, Kiranyaz, Serkan, Kashem, Saad Bin Abul, Islam, Mohammad Tariqul, Al Maadeed, Somaya, Zughailer, Susu M, Khan, Muhammad Salman, *et al.* 2021. Exploring the effect of image enhancement techniques on COVID-19 detection using chest X-ray images. *Computers in biology and medicine*, **132**, 104319.
- Ranjbar, Sajad, Nejad, Fereidoon Moghadas, Zakeri, Hamzeh, & Gandomi, Amir H. 2020. Computational intelligence for modeling of asphalt pavement surface distress. *New Materials in Civil Engineering*.
- Ranjbarzadeh, Ramin, Kasgari, Abbas Bagherian, Ghouschi, Saeid Jafarzadeh, Anari, Shokofeh, Naseri, Maryam, & Bendeche, Malika. 2021. Brain tumor segmentation based on deep learning and an attention mechanism using MRI multi-modalities brain images. *Scientific Reports*, **11**(1), 1–17.
- Refaeilzadeh, Payam, Tang, Lei, Liu, Huan, *et al.* 2009. Cross-validation. *Encyclopedia of database systems*, **5**, 532–538.
- Ren, Jing, Green, Mark, & Huang, Xishi. 2021. From traditional to deep learning: Fault diagnosis for autonomous vehicles. *Pages 205–219 of: Learning Control*. Elsevier.
- Renard, Félix, Guedria, Soulaïmane, De Palma, Noel, & Vuillerme, Nicolas. 2020. Variability and reproducibility in deep learning for medical image segmentation. *Scientific Reports*, **10**(1), 1–16.
- Rezende, Vanessa Castro, Costa, Michel, Santos, Adam, & ao de Oliveira, Roberto Célio Lim 2019. Aplicação de Redes Neurais Convolucionais na Classificação de Doenças de Plantas. *Pages 1–7 of: Fernandes, Bruno José Torres, & {Pereira Júnior}, Antônio (eds), Anais do 14 Congresso Brasileiro de Inteligência Computacional*. Curitiba, PR: ABRICOM.

- Ribeiro, Marco Tulio, Singh, Sameer, & Guestrin, Carlos. 2016. "Why should i trust you?" Explaining the predictions of any classifier. *Pages 1135–1144 of: Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining.*
- Riquelme, Diego, & Akhloufi, Moulay A. 2020. Deep learning for lung cancer nodules detection and classification in CT scans. *AI*, **1**(1), 28–67.
- Rodrigues, Diego Alves. 2018. Deep Learning e redes neurais convolucionais: reconhecimento automático de caracteres em placas de licenciamento automotivo.
- Rodrigues, Iago, Santos, Guto Leoni, Sadok, Djamel FH, & Endo, Patricia Takako. 2021. Classifying COVID-19 positive X-ray using deep learning models. *IEEE Latin America Transactions*, **19**(6), 884–892.
- Rousan, Liqa A, Elobeid, Eyhab, Karrar, Musaab, & Khader, Yousef. 2020. Chest x-ray findings and temporal lung changes in patients with COVID-19 pneumonia. *BMC Pulmonary Medicine*, **20**(1), 1–9.
- Santos, Carlos Alexandre Silva dos. 2017. Reconhecimento de imagens de marcas de gado utilizando redes neurais convolucionais e máquinas de vetores de suporte.
- Schmidhuber, Jürgen. 2015. Deep learning in neural networks: An overview. *Neural networks*, **61**, 85–117.
- Sharma, Neeraj, & Aggarwal, Lalit M. 2010. Automated medical image segmentation techniques. *Journal of medical physics/Association of Medical Physicists of India*, **35**(1), 3.
- Sharma, Neeraj, Saba, Luca, Khanna, Narendra N, Kalra, Mannudeep K, Fouda, Mostafa M, & Suri, Jasjit S. 2022. Segmentation-Based Classification Deep Learning Model Embedded with Explainable AI for COVID-19 Detection in Chest X-ray Scans. *Diagnostics*, **12**(9), 2132.
- Silva, Iago RR, Silva, Gabriela SL, de Souza, Rodrigo G, dos Santos, Wellington P, & Roberta, A de A. 2019. Model Based on Deep Feature Extraction for Diagnosis of Alzheimer's Disease. *Pages 1–7 of: 2019 International Joint Conference on Neural Networks (IJCNN).* IEEE.

- Silva, Rodrigo Emerson Valentim da. 2018. Um estudo comparativo entre redes neurais convolucionais para a classificação de imagens.
- Simonyan, Karen, & Zisserman, Andrew. 2014. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.
- Singh, Pratibha, Chaudhury, Santanu, & Panigrahi, Bijaya Ketan. 2021. Hybrid MPSO-CNN: Multi-level Particle Swarm optimized hyperparameters of Convolutional Neural Network. *Swarm and Evolutionary Computation*, **63**, 100863.
- Srivastava, Nitish, Hinton, Geoffrey, Krizhevsky, Alex, Sutskever, Ilya, & Salakhutdinov, Ruslan. 2014. Dropout: a simple way to prevent neural networks from overfitting. *The journal of machine learning research*, **15**(1), 1929–1958.
- Su, Run, Zhang, Deyun, Liu, Jinhui, & Cheng, Chuandong. 2021. MSU-Net: Multi-scale U-Net for 2D medical image segmentation. *Frontiers in Genetics*, **12**, 140.
- Szegedy, Christian, Liu, Wei, Jia, Yangqing, Sermanet, Pierre, Reed, Scott, Anguelov, Dragomir, Erhan, Dumitru, Vanhoucke, Vincent, & Rabinovich, Andrew. 2014. Going deeper with convolutions. *Pages 1–9 of: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*.
- Szegedy, Christian, Liu, Wei, Jia, Yangqing, Sermanet, Pierre, Reed, Scott, Anguelov, Dragomir, Erhan, Dumitru, Vanhoucke, Vincent, & Rabinovich, Andrew. 2015a. Going deeper with convolutions. *Pages 1–9 of: Proceedings of the IEEE conference on computer vision and pattern recognition*.
- Szegedy, Christian, Vanhoucke, Vincent, Ioffe, Sergey, Shlens, Jon, & Wojna, Zbigniew. 2015b. Rethinking the inception architecture for computer vision. *Pages 2818–2826 of: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*.
- Szegedy, Christian, Ioffe, Sergey, Vanhoucke, Vincent, & Alemi, Alexander. 2017. Inception-v4, inception-resnet and the impact of residual connections on learning. *In: Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 31.
- Tahir, Anas M, Chowdhury, Muhammad EH, Khandakar, Amith, Rahman, Tawsifur, Qiblawey, Yazan, Khurshid, Uzair, Kiranyaz, Serkan, Ibtehaz, Nabil, Rahman, M Shohel, Al-Madeed,

- Somaya, *et al.* 2021a. COVID-19 Infection Localization and Severity Grading from Chest X-ray Images. *arXiv preprint arXiv:2103.07985*.
- Tahir, Anas M., Chowdhury, Muhammad E. H., Qiblawey, Yazan, Khandakar, Amith, Rahman, Tawsifur, Kiranyaz, Serkan, Khurshid, Uzair, Ibte haz, Nabil, Mahmud, Sakib, & Ezeddin, Maymouna. 2021b. *COVID-QU-Ex*. <https://doi.org/10.34740/kaggle/dsv/3122958>.
- Tahir, Anas Mohammed. 2021. *COVID-QU-Ex Dataset*. <https://www.kaggle.com/datasets/anasmohammedtahir/covidqu>. Acesso em: 10 de abril de 2022.
- Toennies, Klaus D. 2012. *Advances in Computer Vision and Pattern Recognition*.
- Wang, Cheng, Chen, Delei, Hao, Lin, Liu, Xuebo, Zeng, Yu, Chen, Jianwei, & Zhang, Guokai. 2019. Pulmonary image classification based on inception-v3 transfer learning model. *IEEE Access*, **7**, 146533–146541.
- Wang, Zekun, Zou, Yanni, & Liu, Peter X. 2021. Hybrid dilation and attention residual U-Net for medical image segmentation. *Computers in Biology and Medicine*, **134**, 104449.
- Wang, Zhiwei, Lv, Jinxin, Yang, Yunqiao, Lin, Yi, Li, Qiang, Li, Xin, & Yang, Xin. 2022. Accurate scoliosis vertebral landmark localization on X-ray images via shape-constrained multi-stage cascaded CNNs. *Fundamental Research*.
- WHO. 2014. *High-priority target product profiles for new tuberculosis diagnostics: report of a consensus meeting*.
- Wörner, Alexandra, & Spitzer, Marcel. 2019. *Machine Learning Interpretability: Explaining Blackbox Models with LIME (Part II)*. <https://www.inovex.de/blog/lime-machine-learning-interpretability/>. Accessed: 06-02-2021.
- Yu, Hang, Yang, Laurence T, Zhang, Qingchen, Armstrong, David, & Deen, M Jamal. 2021. Convolutional neural networks for medical image analysis: state-of-the-art, comparisons, improvement and perspectives. *Neurocomputing*.
- Zhang, Moqi. 2021. *A Comparison of Image Classification with Different Activation Functions in Balanced and Unbalanced Datasets*. Ph.D. thesis, Virginia Tech.

- Zhang, Pengyi, Zhong, Yunxin, Deng, Yulin, Tang, Xiaoying, & Li, Xiaoqiong. 2020a. Drr4covid: Learning Automated COVID-19 Infection Segmentation From Digitally Reconstructed Radiographs. *Ieee Access*, **8**, 207736–207757.
- Zhang, Ziang, Wu, Chengdong, Coleman, Sonya, & Kerr, Dermot. 2020b. DENSE-INception U-net for medical image segmentation. *Computer methods and programs in biomedicine*, **192**, 105395.
- Zhao, Songwei, Wang, Pengjun, Heidari, Ali Asghar, Chen, Huiling, Turabieh, Hamza, Mafarja, Majdi, & Li, Chengye. 2021. Multilevel threshold image segmentation with diffusion association slime mould algorithm and Renyi's entropy for chronic obstructive pulmonary disease. *Computers in Biology and Medicine*, **134**, 104427.
-