

Segmentação de Infecções Pulmonares de COVID-19 com a Rede Mask R-CNN

Segmentation COVID-19 Lung Infections with the R-CNN Mask Network

Segmentación de infecciones pulmonares por COVID-19 con la red Mask R-CNN

Hugo Silveira Sousa¹, Abdenago Alves Pereira Neto¹, Iális Cavalcante de Paula Júnior¹,
Clara Ricardo de Melo²

1 Programa de Pós-Graduação em Engenharia Elétrica e Computação, Universidade Federal do Ceará, Sobral (CE), Brasil.

2 Curso de Engenharia de Computação, Universidade Federal do Ceará, Sobral (CE), Brasil.

Autor correspondente: Hugo Silveira Sousa
E-mail: hugosousa1111@alu.ufc.br

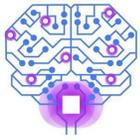
Resumo

A COVID-19 se espalhou pelo mundo causando esgotamento de recursos médicos em vários países. Métodos computacionais que analisam imagens de infecções pulmonares podem ser utilizados para diagnóstico e estimativa da evolução dessa doença. O artigo apresenta os resultados de um modelo de aprendizagem profunda (*Mask R-CNN*), para segmentação automática de infecções pulmonares em tomografias computadorizadas, utilizando a base *COVID-19 CT Lung and Infection Segmentation Dataset*. Os melhores resultados deste trabalho, com a rede que faz a segmentação de pulmões, foram 69,92% para o índice *Dice* e 55,72% para o índice *Jaccard*.

Descritores: Processamento de Imagem Assistida por Computador; COVID-19; Tomografia Computadorizada

Abstract

COVID-19 has spread around the world causing depletion of medical resources in several countries. Computational methods that analyze images of pulmonary infections can be used for diagnosis and estimation of the evolution of this disease. The paper presents the results of a deep learning model (*Mask R-CNN*), for automatic segmentation of lung infections in CT scans, using the *COVID-19 CT Lung and Infection Segmentation Dataset*.



The best results of this paper, with the network that performs the segmentation of lungs, were 69.92% for the Dice index and 55.72% for the Jaccard index.

Keywords: Image Processing, Computer-Assisted; COVID-19; CT scans

Resumen

El COVID-19 se ha extendido por todo el mundo provocando el agotamiento de los recursos médicos en varios países. Los métodos computacionales que analizan imágenes de infecciones pulmonares pueden utilizarse para el diagnóstico y estimación de la evolución de esta enfermedad. El artículo presenta los resultados de un modelo de aprendizaje profundo (Mask R-CNN), para la segmentación automática de infecciones pulmonares en tomografías computarizadas, utilizando el COVID-19 CT Lung and Infection Segmentation Dataset. Los mejores resultados de este trabajo, con la red que realiza la segmentación de pulmones, fueron 69,92% para el índice Dice y 55,72% para el índice Jaccard.

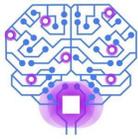
Descriptores: Procesamiento de Imagen Asistido por Computador; COVID-19; Tomografía Computarizada

Introdução

O crescimento exponencial da pandemia de COVID-19 gera preocupação e esgota os recursos médicos de inúmeros países. A criação de métodos para diagnóstico rápido e preciso são essenciais para diminuição da demanda hospitalar e ajudam a tomar medidas clínicas com pacientes de grupos de risco.

Exames de imagem como a Radiografia (RX) e as Tomografias Computadorizadas (TC) geralmente são amplamente disponíveis em hospitais e podem ser utilizados para diagnóstico de COVID-19, pois o vírus, na maioria dos casos graves, gera uma infecção nos pulmões dos pacientes, essas lesões podem ser identificadas por médicos especialistas. Em especial, as TC geram resultados mais completos sobre as áreas estudadas, pois este tipo de exame obtém imagens em camadas da região, que para este cenário é o pulmão.

As metodologias envolvendo Visão Computacional com algoritmos de *Deep Learning* podem auxiliar os médicos especialistas para acelerar o processo de detecção



e classificação de doenças como a COVID-19 com alto nível de acurácia, o que é de grande importância para o cenário atual dos hospitais de todo o mundo.

A segmentação de infecções em imagens médicas é uma fase crucial para detecção de doenças utilizando Visão Computacional. As infecções segmentadas podem ser utilizadas em algoritmos de classificação, obtendo níveis mais elevados de acurácia quando comparados com os algoritmos que não utilizam segmentação de suas imagens. Essa fase também auxilia os especialistas para entenderem o nível de dissipação de uma infecção no órgão estudado.

Para este trabalho, inicialmente foi realizado o levantamento bibliográfico das principais técnicas utilizadas nos estudos de identificação de COVID-19 em imagens médicas, depois foram realizados experimentos com a rede *Mask Region-Based Convolutional Neural Network (Mask R-CNN)* neste cenário.

O objetivo deste artigo é apresentar os resultados de um modelo de aprendizagem profunda (*Mask R-CNN*), para segmentação automática de infecções pulmonares em tomografias computadorizadas, utilizando a base *COVID-19 CT Lung and Infection Segmentation Dataset*.

Métodos

Os métodos deste trabalho envolvem treinamento e teste de modelos *Mask R-CNN*, a figura 1 apresenta o fluxograma de treinamento do modelo de segmentação de infecções, inicialmente a base de dados de treino é pré-processada, com o ajuste de contraste e o corte da região dos pulmões utilizando as segmentações de especialistas, por fim, o modelo é treinado.

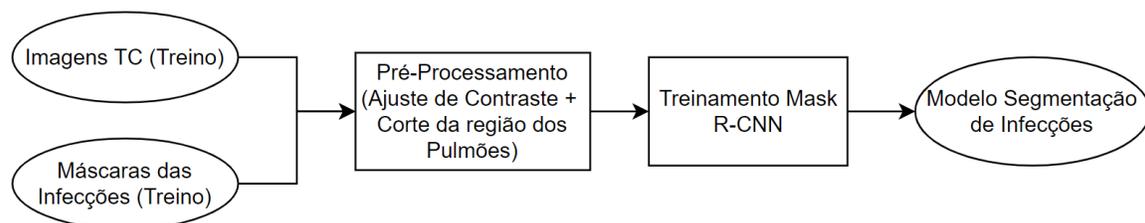
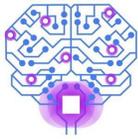


Figura 1 – Fluxo de treinamento do modelo de segmentação de infecções



A figura 2 apresenta o fluxograma do modelo de segmentação de pulmões, a base de dados de treino é pré-processada com o ajuste de contraste, então o modelo de *Mask R-CNN* é treinado.

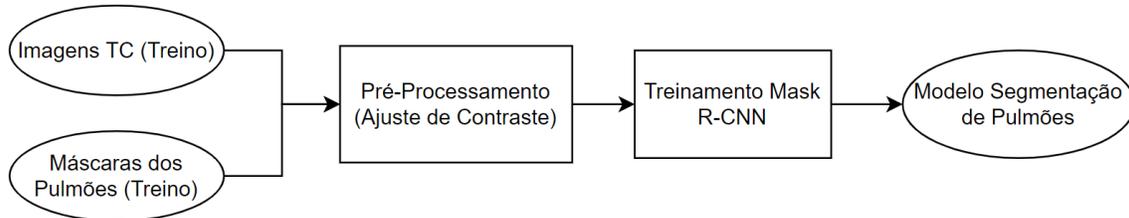


Figura 2 – Fluxo de treinamento do modelo de segmentação de pulmões

A figura 3 apresenta a fase de teste dos modelos. As imagens de TC da base de dados de teste são pré-processadas, realizando o ajuste de contraste e o corte da região dos pulmões. A segmentação nessa fase é realizada pelo modelo de segmentação de pulmões, depois essas imagens são enviadas para o modelo de segmentação de infecções, resultando nas máscaras preditas. Por fim, as máscaras preditas são comparadas com as máscaras reais das imagens de teste, através do cálculo dos coeficientes de similaridade.

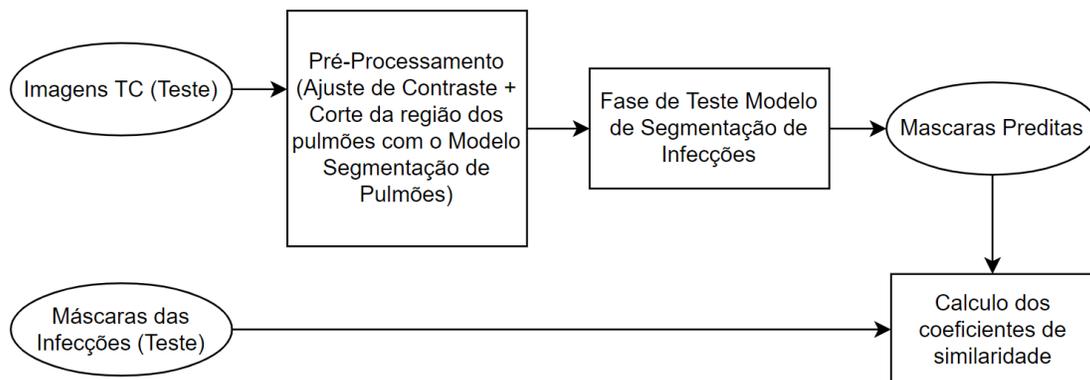
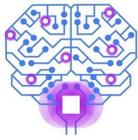


Figura 3 – Fluxo de teste

Base de Imagens

A base de imagens utilizada neste trabalho foi a *COVID-19 CT Lung and Infection Segmentation Dataset* ⁽¹⁾. Essa base contém 20 TC de tórax de pacientes com COVID-19, de duas fontes diferentes, 10 TC da organização *Radiopedia* e 10 TC da organização *Coronacases*.



Além das imagens dos exames, a base acompanha as máscaras de segmentação das infecções e dos pulmões, criadas por dois radiologistas e verificadas por um terceiro radiologista mais experiente, todos usaram as mesmas técnicas de segmentação manual. Na figura 4 pode ser observada uma camada de um exame dessa base, acompanhada da máscara da infecção do pulmão.



Figura 4 – Amostra da base. Da esquerda para a direita: TC, Infecção, Pulmões

Como as TC foram realizadas em equipamentos diferentes, com diversas espessuras de camadas, os exames variam em quantidade de imagens, no total, os 20 exames têm 2112 imagens. Foram eliminadas todas as imagens que não apresentavam infecções, sobrando 1615 imagens. Por fim, foram eliminadas as camadas problemáticas, estas não apresentavam conteúdo, com a imagem inteira com *pixels* brancos ou *pixels* pretos, restando para o estudo 1609 imagens.

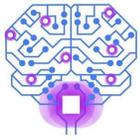
Essas imagens foram divididas aleatoriamente em dois grupos, o de Treino com 80% (1287 imagens) e o de Teste com 20% (322 imagens).

Ambiente de Desenvolvimento

Os experimentos deste trabalho foram implementados com a linguagem Python 3.7, no ambiente *Google Colaboratory Pro* com uma máquina remota de 12GB de memória RAM, com acesso a processamento dedicado em GPU.

As principais bibliotecas utilizadas foram:

- NiBabel: para manipulação das TC que estão no formato ".nii";
- OpenCV: para trabalhar com as imagens e para realizar o pré-processamento;
- NumPy: facilita a manipulação de matrizes dentro do Python;
- Tensorflow: biblioteca de *machine learning* para aplicações de visão computacional, sendo utilizada para criação dos modelos de *Mask R-CNN*.



Pré-Processamento

Todas as imagens passaram por uma fase de pré-processamento. Durante os testes, foi observado que essa fase é essencial para elevar os valores das métricas utilizadas como coeficientes de semelhança das máscaras.

São utilizadas duas fases no pré-processamento das imagens, na primeira é realizado um ajuste de contraste utilizando a Equalização de Histograma Adaptativo Limitada de Contraste (*CLAHE*), essa função melhora a qualidade das imagens estudadas, o que eleva os índices de similaridade durante o treinamento dos modelos.

Na segunda fase do pré-processamento, as imagens são recortadas para remoção das regiões que não são os pulmões para o treinamento do modelo de segmentação de infecções. A remoção dessas regiões eliminou o problema da detecção de infecção fora dos pulmões (falsos positivos), além de ajudar o modelo a convergir mais rápido durante o treinamento, visto que as imagens eram menores e apresentavam menos detalhes. Nessa fase foram utilizadas as máscaras dos pulmões criadas por especialistas. No experimento final da metodologia, os cortes são realizados pelo modelo treinado de segmentação de pulmões. A figura 5 apresenta uma amostra depois da fase de pré-processamento e sua máscara de infecção.

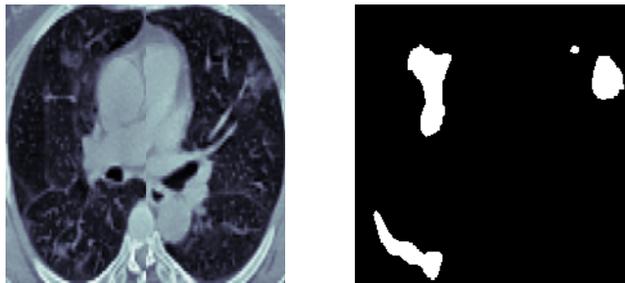
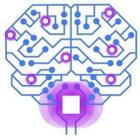


Figura 4 – Amostra da base. Da esquerda para a direita: TC, Infecção

Treinamento

A rede *Mask R-CNN* (Redes Neurais Convolucionais Baseadas em Região) é uma arquitetura de rede neural profunda, que estende a rede *Faster R-CNN* ⁽²⁾. A rede *Faster* faz a detecção de objetos em imagens, já a *Mask* faz a segmentação semântica, que é importante para a identificação de onde estão objetos com diferentes formas em uma imagem, como os diferentes níveis de infecções causadas pela COVID-19.



Esta rede foi escolhida pois ela é bem difundida no cenário de detecção de doenças com imagens. É simples de realizar a fase de treinamento da rede *Mask R-CNN*, além de já existirem implementações em bibliotecas populares como o *Tensorflow*. É uma rede eficiente, conseguindo elevados valores de coeficientes de similaridade em diversos cenários.

A *Mask R-CNN* utilizada neste trabalho tem como *backbone* uma rede ResNet50, que é uma rede residual amplamente aplicada em trabalhos de Visão Computacional, a rede desse trabalho é constituída de 50 blocos residuais. Essa rede foi pré-treinada com a base *COCO dataset* ⁽³⁾, e os seus pesos foram reajustados na fase de treinamento com as imagens dos exames de TC e as máscaras de segmentação das infecções e pulmões, feitas por especialistas.

Coeficientes de Similaridade

Para avaliar a qualidade das segmentações são necessários coeficientes que avaliem a similaridade entre a imagem segmentada pelo algoritmo e segmentada por um profissional especialista, para isso foram utilizadas as métricas *Jaccard* (equação 1), *Dice* (equação 2), precisão (equação 3), sensibilidade (equação 4) e especificidade (equação 5), apresentados nas equações abaixo.

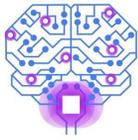
As equações são definidas com os termos Verdadeiro Positivo (VP), Verdadeiro Negativo (VN), Falso Positivo (FP) e Falso Negativo (FN). VP são os *pixels* selecionados corretamente na segmentação que fazem parte da região de interesse da imagem real, VN são os *pixels* que não foram selecionados e não fazem parte da região de interesse, FP são *pixels* selecionados erroneamente, pois não fazem parte da região de interesse e FN são os *pixels* que não foram selecionados durante a segmentação, mas fazem parte da região de interesse da imagem real.

$$Jaccard = \frac{VP}{VP + FP + FN} \quad (1)$$

$$Dice = \frac{2VP}{2VP + FP + FN} \quad (2)$$

$$Precisão = \frac{VP}{VP + FP} \quad (3)$$

$$Sensibilidade = \frac{VP}{VP + FN} \quad (4)$$



$$\text{Especificidade} = \frac{VN}{VN+FP} \quad (5)$$

Resultados e Discussões

Os resultados desse trabalho são baseados na média dos coeficientes de similaridade entre as imagens preditas pelas redes e as imagens segmentadas por especialistas. Os valores das tabelas de resultados foram truncados, com precisão de duas casas decimais.

Foram realizados experimentos para encontrar os melhores parâmetros para as duas redes, variando a taxa de aprendizagem (TA), *momentum* M e taxa de decaimento dos pesos (TDP), e número de épocas, os parâmetros que retornaram as redes com os menores valores de perda durante o treinamento foram escolhidos.

A rede de segmentação de pulmões teve como parâmetros TA = 0,005; M = 0,9; TDP = 0.0001 e 50 épocas. Na tabela 1 são apresentados os resultados para a segmentação das imagens de treino e de teste, os resultados são as médias (μ) e os desvios-padrão (σ), para cada coeficiente de similaridade.

Tabela 1 – Resultados da segmentação de pulmões

Base	Dice		Jaccard		Precisão		Sensibilidade		Especificidade	
	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ
Treino	72,82	10,32	58,21	9,11	63,89	12,43	83,14	8,74	93,27	2,86
Teste	73,77	10,71	59,36	10,91	66,72	11,55	83,09	7,47	93,91	1,44

Percebe-se que a rede de segmentação de pulmões apresenta bons resultados com as imagens de teste, e que esses resultados não estão muito abaixo das imagens de treino, indicando que não houve *overfitting* na fase de treinamento, e que conseguiu generalizar as características das imagens estudadas.

A rede de segmentação de infecções teve como parâmetros TA = 0,005; M = 0,9; TDP = 0.0001 e 100 épocas. Foram realizados dois processos de teste, na primeira as imagens são cortadas na fase de pré-processamento com as imagens segmentadas por especialistas (ESP), e na segunda as imagens são cortadas pela rede de segmentação de pulmões (RSP). Na tabela 2 são apresentados os resultados para a segmentação das imagens de treino e de teste, para os dois processos de pré-processamento utilizados,



os resultados são as médias (μ) e os desvios-padrão (σ), para cada coeficiente de similaridade.

Tabela 2 – Resultados da segmentação de infecções

Pré.	Base	Dice		Jaccard		Precisão		Sensib.		Espec.	
		μ	σ	μ	σ	μ	σ	μ	σ	μ	σ
ESP	Treino	80,43	18,49	70,12	18,90	82,86	18,86	80,30	20,00	98,75	2,08
	Teste	81,73	16,36	71,38	16,82	84,01	16,80	80,92	18,24	98,78	1,81
RSP	Treino	69,01	15,43	54,87	17,47	62,95	16,18	77,14	15,08	95,12	1,80
	Teste	69,92	14,98	55,72	16,53	63,81	15,74	78,87	14,48	95,20	1,67

Nota-se que os resultados com as imagens de teste não estão muito abaixo da fase de teste, para os dois pré-processamentos. O resultado com pré-processamento realizado pela rede de segmentação de pulmões é satisfatório, mas ainda está abaixo do realizado com as imagens segmentadas por especialistas.

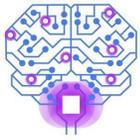
Na tabela 3 os resultados desse trabalho são comparados com os resultados encontrados no estado da arte dessa problemática, estes utilizam diferentes bases de imagens e técnicas.

Tabela 3 – Comparação dos resultados encontrados com a literatura

Trabalho	Técnica	Base	Dic.	Jac.	Pre.	Sen.	Esp.
(4)	E. Kapur	275 TC	71,4	57,3	73,9	73,3	99,4
(5)	CNN	449 TC	88,0	-	-	90,2	99,7
(6)	COPL-Net	558 TC	80,72	-	-	-	-
(7)	U-Net	473 TC	83,1	-	-	86,7	99,3
(8)	CNN	100 img *	86,04	75,5	-	-	-
(9)	CoSinGAN	100 img *	71,3	-	-	-	-
(10)	Inf-Net	100 img *	59,7	-	51,5	86,5	97,7
(11)	CNN	100 img *	64,0	-	58,1	63,0	95,3
(12)	CNN	20 TC **	69,8	-	85,7	-	66,2
(13)	3D U-Net	20 TC **	76,1	-	-	73,0	99,9
Rede 1	Mask R-CNN	20 TC **	81,73	71,38	84,01	80,92	98,78
Rede 2	Mask R-CNN	20 TC **	69,92	55,72	63,81	78,87	95,20

*Base: COVID-19 CT Segmentation Dataset

**Base: COVID-19 CT Lung and Infection Segmentation Dataset



Rede 1: Pré-Proc. com segmentações dos pulmões realizadas por especialistas

Rede 2: Pré-Proc. com modelo de segmentações dos pulmões

Os melhores resultados deste trabalho, de modo geral, ficaram abaixo dos valores dos melhores resultados dos trabalhos comparados, mas ao serem analisados aqueles que utilizam a base *COVID-19 CT Lung and Infection Segmentation Dataset*, os índices *Dice* e sensibilidade são maiores.

Entre os 12 resultados de *Dice*, que é a principal métrica utilizada por todos os trabalhos, a rede 1 ficou em 4º lugar. A rede 2, que utiliza o modelo de segmentação de pulmões, está com resultado próximo dos trabalhos que utilizaram a mesma base.

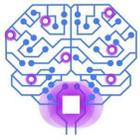
Conclusão

Este trabalho implementa a rede *Mask R-CNN* para realizar a segmentação de infecções pulmonares em imagens de Tomografias Computadorizadas (TC) de tórax de pacientes com COVID-19. As imagens estudadas constituem de 20 TC de pacientes com COVID-19 da base pública *COVID-19 CT Lung and Infection Segmentation Dataset*.

Inicialmente foi realizada uma fase de pré-processamento nas imagens que melhora os resultados das segmentações. Para essa fase, foi treinado um modelo que faz a segmentação dos pulmões das imagens, que é utilizado para remover as regiões que não são dos pulmões. A rede *Mask R-CNN* foi utilizada para treinar os modelos de segmentação de pulmões e infecções.

Os resultados encontrados são relevantes, estão próximos dos melhores resultados do estado da arte, mas ainda estão abaixo destes. Os melhores resultados deste trabalho foram 81,73%, 71,38%, 84,01%, 80,92% e 98,78% para os índices *Dice*, *Jaccard*, precisão, sensibilidade e especificidade, respectivamente, para a rede que utiliza como pré-processamento imagens dos pulmões segmentadas por especialistas. Para a rede que utiliza no pré-processamento o modelo de segmentação de pulmões, teve como resultados 69,92%, 55,72%, 63,81%, 78,87%, 95,20% para os índices *Dice*, *Jaccard*, precisão, sensibilidade e especificidade, respectivamente.

Futuramente, pretende-se estender a metodologia deste trabalho para uma fase inicial que realizará a classificação de um grupo de doenças que causam infecções no

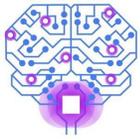


pulmão, entre elas a doença COVID-19, para posteriormente realizar os passos desse trabalho para segmentação da infecção.

Além disso, as técnicas desse trabalho devem ser aplicadas em outras bases de imagem, como a *COVID-19 CT Segmentation Dataset*, para analisar se os resultados se mantêm e facilitar a comparação com outros trabalhos.

Referências

1. Jun M, Cheng G, Yixin W, Xingle A, Jiantao G, Ziqi Y, Minqing Z, Xin L, Xueyuan D, Shucheng C, Hao W. COVID-19 CT lung and infection segmentation dataset.
2. He K, Gkioxari G, Dollár P, Girshick R. Mask r-cnn. IEEE international conference on computer vision 2017 (pp. 2961-2969).
3. Sawatzky J, Souri Y, Grund C, Gall J. What object should i use?-task driven object detection. IEEE/CVF Conference on Computer Vision and Pattern Recognition 2019 (pp. 7605-7614).
4. Oulefki A, Agaian S, Trongtirakul T, Laouar AK. Automatic COVID-19 lung infected region segmentation and measurement using CT-scans images. Pattern recognition. 2021 Jun 1;114:107747.
5. Amyar A, Modzelewski R, Li H, Ruan S. Multi-task deep learning based CT imaging analysis for COVID-19 pneumonia: Classification and segmentation. Computers in Biology and Medicine. 2020 Nov 1;126:104037.
6. Wang G, Liu X, Li C, Xu Z, Ruan J, Zhu H, Meng T, Li K, Huang N, Zhang S. A noise-robust framework for automatic segmentation of COVID-19 pneumonia lesions from CT images. IEEE Transactions on Medical Imaging. 2020 Jun 5;39(8):2653-63.
7. Zhou T, Canu S, Ruan S. An automatic covid-19 ct segmentation network using spatial and channel attention mechanism. arXiv 2020. arXiv preprint arXiv:2004.06673.
8. El-Bana S, Al-Kabbany A, Sharkas M. A multi-task pipeline with specialized streams for classification and segmentation of infection manifestations in COVID-19 scans. PeerJ Computer Science. 2020 Oct 19;6:e303.
9. Zhang P, Zhong Y, Deng Y, Tang X, Li X. CoSinGAN: learning COVID-19 infection segmentation from a single radiological image. Diagnostics. 2020 Nov 3;10(11):901.
10. Fan DP, Zhou T, Ji GP, Zhou Y, Chen G, Fu H, Shen J, Shao L. Inf-net: Automatic covid-19 lung infection segmentation from ct images. IEEE Transactions on Medical Imaging. 2020 May 22;39(8):2626-37.
11. Elharrouss O, Subramanian N, Al-Maadeed S. An encoder-decoder-based method for COVID-19 lung infection segmentation. arXiv preprint arXiv:2007.00861. 2020 Jul 2.



12. Yao Q, Xiao L, Liu P, Zhou SK. Label-free segmentation of COVID-19 lesions in lung CT. IEEE transactions on medical imaging. 2021 Mar 24;40(10):2808-19.
13. Müller D, Rey IS, Kramer F. Automated chest ct image segmentation of covid-19 lung infection based on 3d u-net. arXiv preprint arXiv:2007.04774. 2020 Jun 24.