Lung Parenchyma Segmentation Using Mask R-CNN in COVID-19 Chest CT Scans

Wilmer Alberto Pacheco Llacho, Eveling Castro-Gutierrez, Luis David Huallpa Tapia
Universidad Nacional de San Agustín de Arequipa, Peru

***Abstract*—During the COVID-19 pandemic, the precise evaluation of lung impairments using computed tomography (CT) scans became critical for understanding and managing the disease; however, specialists faced a high workload and the urgent need to deliver fast and accurate results. To address this, deep learning models offered a promising solution by automating lung identification and lesion localization associated with COVID-19. This study employs the semantic segmentation technique Mask R-CNN, integrated with a ResNet-50 backbone, to analyze CT scans of COVID-19 patients. The model was trained using an annotated dataset, enhancing its ability to accurately segment and delineate the lung parenchyma in CT images. The results showed that Mask R-CNN achieved a Dice Similarity Coefficient (DSC) of 93.4%, demonstrating high concordance between the segmented areas and clinically relevant regions. These findings highlight the effectiveness of the proposed approach for precise lung tissue segmentation in CT scans, enabling quantitative assessments of lung impairments and providing valuable insights for diagnosis and patient monitoring.**

***Keywords*—Mask R-CNN; ResNet-50; computed tomography; lung parenchyma; COVID-19**

I. Introduction

The COVID-19 pandemic posed significant challenges in the diagnosis and management of respiratory diseases, particularly in the accurate assessment of lung conditions. In this context, medical imaging has become a fundamental tool, as its primary objective is to generate meaningful data to analyze the physiology and anatomy of various organs or areas of the human body. Among the most prominent modalities addressing these needs is simple computed tomography (CT). Based on X-ray emission, this technique, considered the oldest within medical imaging, enables non-invasive analysis of the internal structures of the human body with high precision and accuracy [1], [2], [3].

Medical imaging emerged as an indispensable tool to address these challenges. Unlike previous viral pandemics, where its use was more limited, imaging modalities have taken on a central role by enabling the rapid and accurate identification of lung patterns characteristic of the disease [4]. Furthermore, medical imaging is essential for detailed follow-up in cases where the disease becomes complicated, allowing for monitoring of progression and evaluation of the response to clinical interventions [5].

X-rays were one of the tools that emerged during the COVID-19 pandemic due to the limitations in the sensitivity of PCR tests and their lack of immediate availability. Their use enabled early diagnoses. Meanwhile, computed tomography (CT), widely available in developed countries, proved to be highly effective in precisely detailing the condition of the lung parenchyma [6]. Particularly, CT has been fundamental in evaluating lung changes such as ground-glass opacities (GGO), consolidations, and pleural thickening. These features significantly enhance diagnostic accuracy when combined with RT-PCR tests [7] and also allow for monitoring the progression of the disease [7], [8].

However, the manual inspection of these images requires a high level of expertise, specialized human resources, and considerable time, which poses a significant challenge during high-demand situations. In this context, deep learning techniques, particularly convolutional neural networks (CNNs), have emerged as highly effective tools for automating disease detection through the analysis of medical images [9].

The main objective of this study is to segment the pulmonary parenchyma in COVID-19 patients using the Mask R-CNN technique, evaluating the performance of the ResNet50 backbone under different learning rate hyperparameter configurations. Additionally, the impact of data augmentation on the model's accuracy is analyzed, implementing various augmentation techniques to optimize results. This work aims to provide an accurate and efficient tool to assist specialist doctors in identifying and assessing affected areas in lung tissue, thus contributing to a more precise diagnosis and better clinical care.

This article is organized as follows: Section II - Literature Review, Section III - Methodology, Section IV – Results and Discussion, and the final Section V presents conclusions and future work.

II. Literature Review

The authors in study [10] used Mask R-CNN for the diagnosis of COVID-19 through chest X-ray images. This deep learning approach achieved an accuracy of 96.98%, standing out for its efficiency and superiority compared to other artificial intelligence techniques, proving to be accurate and robust in identifying the disease from these images.

D. Suganya and R. Kalpana (2023) [11] employed Mask R-CNN to classify chest CT images and differentiate between COVID-19 positive and negative patients. This model, designed to identify infection regions in the lungs, achieved a mean average precision (mAP) of 91.52% and a classification accuracy of 98.60%, making it an effective solution for the accurate diagnosis of COVID-19.

S. Aparna, et al. (2021) [12], introduced a Mask R-CNN-based model with the ResNet50 architecture to analyze dental

X-rays and estimate the level of filling performed on teeth. Using a dataset of different types of fillings, they trained the model, which performs pixel-level classification, improving the accuracy in diagnosing dental treatments. This approach enables machines to perform automated dental procedures by understanding in more detail the exact region and position of the treatment.

The study by Al Masarweh and colleagues [13] proposed a Mask R-CNN-based method to autonomously generate reference data in MRI images through instance segmentation. Their model achieved a mean average precision (mAP) of 98% for locating and identifying discs, along with a 70% accuracy in classifying regions of interest. This approach allows radiologists to automate the detection of relevant areas in MRI images, improving efficiency and reliability in medical diagnoses, as well as advancing automatic medical image segmentation with cutting-edge neural network technologies.

In study [14] used the ResNet-50 deep learning architecture to classify and detect human sperm heads, achieving an accuracy of 96.66%. This proposed model proved to be efficient in identifying healthy sperm, which are used in the process of intracytoplasmic sperm injection (ICSI). The automation of the process allows for faster and more accurate results, minimizing human errors and improving success rates in infertility treatment.

S. Suriyavarman and A. X. A. R. Annie proposed an algorithm based on the combination of U-Net and Efficient-Net neural networks for the segmentation and classification of lung nodules in CT images (2023) [15]. Using semi-supervised learning with a feature pyramid network (FPN) and the ResNet-50 model for feature extraction, they were able to predict unlabeled nodules. The U-Net technique, with its skip connections, allows for precise nodule localization, while Efficient-Net optimizes the scaling of depth, width, and resolution. Evaluated on the LIDC-IDRI dataset, the model achieved an accuracy of 91.67%, outperforming most existing methods and addressing issues such as high false positive rates and variability in longitudinal data.

In study [16], an improved Mask R-CNN model is proposed, specifically tailored for multi-organ segmentation in the medical field. This model introduces two major enhancements to the original framework: a multi-scale region of interest (ROI) generation method within the region proposal network (RPN) and a pre-background classification subnetwork to enhance segmentation accuracy. Experimental results on an esophageal cancer dataset demonstrated the model's effectiveness, achieving accurate segmentation of organs such as the heart, lungs, and clinically relevant volumes.

The study by E. Dandıl and M. S. Yıldırım (2021) [17] highlights the significance of computer-aided tools for automatic lung segmentation in diagnosing lung diseases. Manual segmentation by experts can introduce errors and inefficiencies. Their research proposed a Mask R-CNN-based approach, leveraging publicly available datasets like HUG-ILD and VESSEL12. The method demonstrated high performance, achieving a Dice similarity coefficient of 95.95% and a volumetric overlap error of 7.65% for the HUG-ILD dataset, and 96.80% and 6.12% for the VESSEL12 dataset. These results validate the effectiveness of the proposed method for precise lung segmentation (Dandıl & Yıldırım, 2021).

Mask R-CNN is a deep learning model that extends Faster R-CNN to perform instance segmentation, classifying objects and generating pixel-level masks for each detected instance. It introduces the RoIAlign method to improve accuracy when processing regions of interest. Its design, which combines detection and segmentation, makes it ideal for applications such as medical computer vision and autonomous driving [18]. Fig. 1 shows the Mask R-CNN framework for instance segmentation

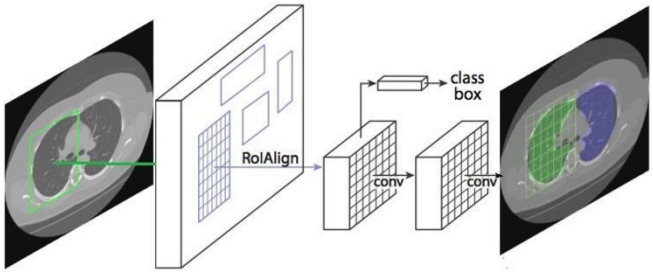

Fig. 1. The Mask R-CNN framework for instance segmentation.

III. METHODOLOGY

A. Database

A total of 20 computed tomography (CT) scans from patients, comprising 3,520 slices, were used for the training and testing phases. Of these, 10 patients were diagnosed with Covid-19 infection pathologies, and 10 patients were diagnosed as healthy or showed no evidence of pulmonary disease.

The number of CT scans used for the training phase included 18 patients, representing 90% of the patients, with a total of 3,126 slices. The number of CT scans for the testing phase included two patients, with a total of 394 slices per image, as shown in Table I.

TABLE I. DATASET FOR TRAINING AND TESTING

Data Set	*%*	*Nro Patient*	*Slides*	*Health*	*Covid*
Training	90	18	3126	9	9
Test	10	2	394	1	1

For the Testing phase, a total of eight computed tomography (CT) scans were used, comprising 259 chest images, as shown in Table II.

TABLE II. DATASET OF HEALTHY AND COVID-19 PATIENTS

Data Set	*%*	*Nro Patient*	*Slides*
Covid-19	75	6	167
Healthy	25	2	92

All CTs for the training and testing phases were obtained from the Zenodo repository in NIFTI format [19]. The CTs for the evaluation phase were obtained from the Research, Technology Transfer, and Software Development Center (CiTeSoft) at the National University of San Agustín.

B. Preprocessing

The images for training, validation, and testing were obtained from medical files in NIFTI format. These slices were converted to PNG format with three channels to preserve the information, scaled to a size of 512 x 512 pixels, and their values were normalized to a range of 0 to 256.

C. Data Augmentation

To avoid overfitting and lack of information, four data augmentation techniques were applied: rotation, horizontal flipping, grid distortion, and elastic transformation. These techniques increased the dataset by 50%, reaching a total of 1760 images distributed between training and testing data, as shown in Table III.

TABLE III. DATASET WITH DATA AUGMENTATION

Data Set	*+%*	*Data Augmentation*
Training	50	1563
Testing	50	197

D. Mask RCNN Architecture

1) Pre-training: In this study, a transfer learning strategy has been adopted by using pre-trained weights from the ImageNet database in the training process of Mask R-CNN. The ImageNet database contains a wide variety of images from different categories, allowing the pre-trained weights to capture general and meaningful visual features. Leveraging these pre-trained weights from ImageNet provides our model with weight initializations that already have a deep understanding of visual patterns, textures, and details in images. This significantly accelerates and improves the training process of Mask R-CNN for our specific task of instance segmentation in medical images of chest CT scans from COVID-19 patients

2) Learning rate: Within the Mask R-CNN architecture, the learning rate is an important hyperparameter that controls the size of the steps the SGD (Stochastic Gradient Descent) optimization algorithm takes to adjust the weights in the model during the training process. This hyperparameter determines how quickly the model converges to a local minimum in the loss function and is responsible for finding a balance between fast and stable model fitting.

3) Implementation: For the implementation of the Mask R-CNN architecture, this study relied on the implementation available in the 'Mask_RCNN' repository by Matterport [9]. This repository provides a Python implementation using TensorFlow, which allowed us to develop and train our instance segmentation model on medical images of chest CT scans from COVID-19 patients.

To carry out the experiments and training of the proposed model, the Mask R-CNN architecture was used with an NVIDIA TITAN RTX graphics card with 24 GB of GDDR6 memory.

The Mask R-CNN parameter configuration was adapted to the specific characteristics of our medical CT image dataset from chest scans of COVID-19 patients. A total of three classes were defined: a) Background and the class corresponding to the areas of interest in the images of the lung regions, b) left, and c) right. For the neural network, 'resnet50' was chosen as the backbone, which leverages the ResNet architecture for feature extraction. The input image dimensions were defined, setting both the minimum and maximum size at 512x512 pixels to ensure consistency in processing. Anchor selection for the Region Proposal Network (RPN) was done using anchor scales of (32, 64, 128, 256, 512) to address different object sizes in our images. During training, a maximum of 200 regions of interest (ROIs) per image and a maximum of 5 true instances for instance detection were established.

For the inference phase, the maximum number of ROIs after the non-maximum suppression (NMS) process was set to 1000, and for training, it was set to 2000. Additionally, a minimum detection confidence threshold of 0.7 was established to ensure the accuracy of the predictions. To mitigate overlap, an NMS threshold of 0.3 was applied to filter redundant detections and improve the consistency of the resulting instance segmentation.

E. Evaluation

For each of the different pre-trained models with a Resnet50 backbone, using different learning rates (0.001, 0.0001, 0.00001) and different epochs, they were evaluated with the following metrics: a) Jaccard Index (1) and b) Dice Coefficient (2).

The equations corresponding to these metrics are presented below.

$$J(A,B) = \frac{D(A,B)}{2+D(A,B)} \quad (1)$$

$$D(A,B) = \frac{2.J(A,B)}{1+J(A,B)} \quad (2)$$

IV. RESULTS AND DISCUSSION

Results were evaluated under two scenarios: with and without Data Augmentation. The best performance was achieved with a learning rate of 0.001. Metrics such as Jaccard Index and Dice Coefficient highlighted the model's segmentation capabilities, which are shown in Table IV.

TABLE IV. SUMMARY OF THE BEST RESULTS

Metric	*With Augmentation*	*Without Augmentation*
Jaccard Index	0.8794	0.8901
Dice Coefficient	0.9266	0.9340

Pixel-level analysis showed high sensitivity (93.39%), specificity (99.56%), and accuracy (98.84%) with augmentation, underscoring the model's reliability in identifying lung regions and lesions.

A. Training with Data Augmentation

In relation to the results obtained for the Jaccard Index metric, better results were achieved with a learning rate of 0.001, with maximum and minimum values of 0.8600 and 0.8794, respectively. Regarding the Dice Coefficient metric, minimum and maximum values of 0.9105 and 0.9266 were

recorded, as shown in Fig. 2. Table V details the results corresponding to each data subset used for training (Fold).

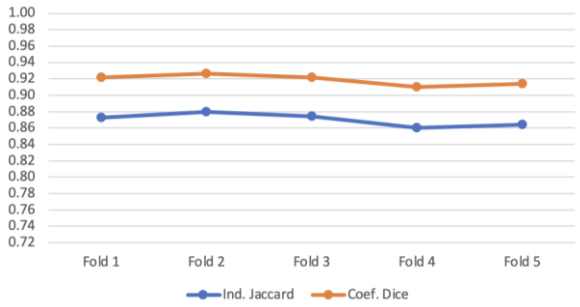

Fig. 2. Learning rate with Jaccard metrics and dice coefficient with data augmentation technique and learning rate of 0.001.

TABLE V. JACCARD METRICS AND DICE COEFFICIENT WITH DATA AUGMENTATION AND LEARNING RATE OF 0.001

Folds	*Jaccard*	*Coef. Dice*
Fold 1	0.8730	0.9215
Fold 2	0.8794	0.9266
Fold 3	0.8742	0.9218
Fold 4	0.8600	0.9105
Fold 5	0.8639	0.9142

In the context of pixel analysis, the metrics of Accuracy, Sensitivity, Specificity, and Precision were evaluated. These metrics provide a detailed understanding of how the model performs in the classification and precise segmentation of pixels in the images. The results corresponding to this evaluation are presented in Fig. 3, Table VI, offering a comprehensive view of the model's effectiveness at the pixel level.

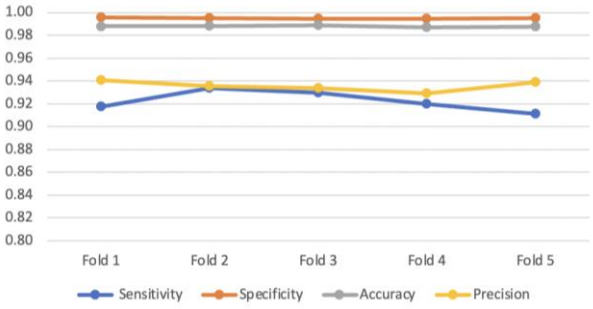

Fig. 3. Learning rate with pixel-level metrics with data augmentation and learning rate of 0.001.

TABLE VI. PIXEL-LEVEL METRICS WITH DATA AUGMENTATION AND LEARNING RATE OF 0.001

Folds	*Sensitivity*	*Specificity*	*Accuracy*	*Precision*
Fold 1	0.9175	0.9956	0.9879	0.9406
Fold 2	0.9339	0.9950	0.9884	0.9357
Fold 3	0.9298	0.9947	0.9885	0.9337
Fold 4	0.9196	0.9946	0.9872	0.9289
Fold 5	0.9113	0.9949	0.9877	0.9390

B. Training without Data Augmentation

In relation to the results obtained for the Jaccard Index metric, better results were achieved with a learning rate of 0.001, with minimum and maximum values of 0.8787 and 0.8901, respectively. Regarding the Dice Coefficient metric, minimum and maximum values of 0.9226 and 0.9339 were recorded, as shown in Fig. 4. Table VII details the results corresponding to each data subset used for training (Fold).

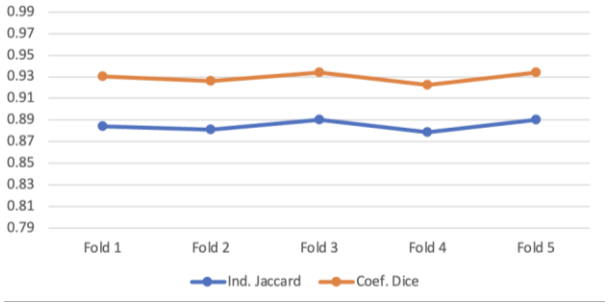

Fig. 4. Learning rate without Jaccard metrics and dice coefficient with technique without data augmentation and learning rate of 0.001.

TABLE VII. JACCARD METRICS AND DICE COEFFICIENT WITHOUT DATA AUGMENTATION AND LEARNING RATE OF 0.001

Folds	*Ind. Jaccard*	*Coef. Dice*
Fold 1	0.8839	0.9302
Fold 2	0.8808	0.9260
Fold 3	0.8901	0.9340
Fold 4	0.8787	0.9226
Fold 5	0.8901	0.9339

In the context of pixel analysis, the metrics of Accuracy, Sensitivity, Specificity, and Precision were evaluated. These metrics provide a detailed understanding of how the model performs in the classification and precise segmentation of pixels in the images. The results corresponding to this evaluation are presented in Fig. 5, Table VIII, offering a comprehensive view of the model's effectiveness at the pixel level.

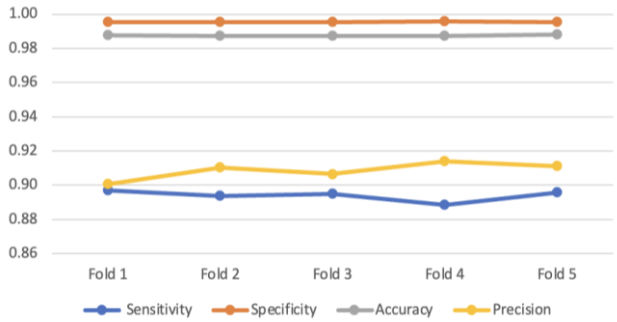

Fig. 5. Learning rate with pixel-level metrics without data augmentation and learning rate of 0.001.

TABLE VIII. PIXEL-LEVEL METRICS WITHOUT DATA AUGMENTATION AND LEARNING RATE OF 0.001

Folds	*Sensitivity*	*Specificity*	*Accuracy*	*Precision*
Fold 1	0.8969	0.9952	0.9877	0.9005
Fold 2	0.8936	0.9954	0.9874	0.9103
Fold 3	0.8947	0.9953	0.9874	0.9064
Fold 4	0.8885	0.9957	0.9872	0.9140
Fold 5	0.8956	0.9954	0.9879	0.9110

In Table XI and Fig. 6, the most notable results regarding the Jaccard Index and Dice Coefficient metrics are presented.

TABLE IX. PIXEL-LEVEL METRICS WITHOUT DATA AUGMENTATION AND LEARNING RATE OF 0.001

Learning Rate	*Ind. Jaccard*	*Coef. Dice*
0.001	0.8901(3)	0.9340(3)
0.0001	0.8777(5)	0.9234(5)
0.00001	0.8264(5)	0.8788(4)

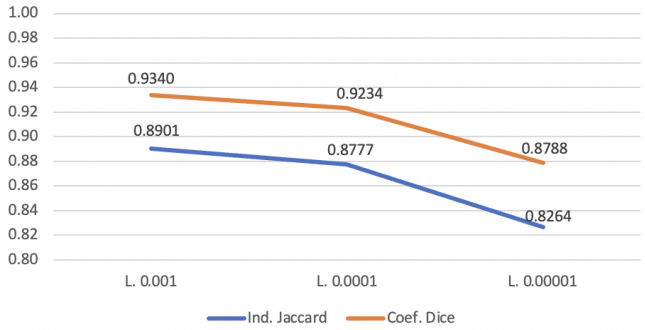

Fig. 6. Learning rate with pixel-level metrics without data augmentation and learning rate of 0.001.

Thorough tests were conducted to optimize the configuration of the semantic segmentation model. In order to explore different approaches, experiments were carried out with the approach a) With data augmentation and b) Without data augmentation, and learning rates were adjusted to 0.001, 0.0001, and 0.00001. After a thorough analysis, the most promising results were obtained by combining the absence of Data Augmentation with a learning rate of 0.001. These findings highlight the importance of precisely and custom tailoring the hyperparameters to maximize the performance of our model. Table IX and Fig. 6, presents the performance metrics under different learning rate configurations, demonstrating that a rate of 0.001 yielded the highest Dice Coefficient and Jaccard Index values.

One of the main limitations of this research was the need to work with all the computed tomography images due to the knowledge provided by the specialists. According to them, COVID-19 could manifest both in the upper and lower images, which complicated the analysis and reduced our detection rate, as the affected areas were very small and difficult to identify. Additionally, we faced the challenge of accurately segmenting the pulmonary parenchyma due to the pathologies caused by COVID-19 in the lungs, which added significant complexity to the analysis process.

The study by Shu, Nian, Yu, and Li (2020) [16] employed Mask R-CNN for lung segmentation, reporting high Dice similarity coefficients (98.1% for the right lung and 97.6% for the left lung). However, the images used do not specify whether the lungs had any pathologies, and the tests were conducted on a range of CT scan slices that included complete organs as well as the lungs, which could influence the model's accuracy when segmenting the lungs exclusively.

In contrast, our research addresses a more complex scenario by working exclusively with CT scan images from patients with COVID-19. This approach not only ensures that the images include clear cases of the disease but also utilizes all slices of the CT scans, including those with smaller or limited affected lung regions. Despite these additional challenges, our model achieved a Dice coefficient of 93.4%, demonstrating its effectiveness in real clinical scenarios related to COVID-19.

Additionally, the work by Dandıl and Yıldırım (2021) [17] also used Mask R-CNN but focused exclusively on lung images with interstitial lung diseases. This study reported Dice coefficients of 95.95% and 96.80% on the HUG-ILD and VESSEL12 datasets, respectively. However, these datasets do not include specific cases of COVID-19, which limits their direct applicability in diagnosing and managing this disease in clinical settings.

In summary, our research stands out by directly addressing lung segmentation in images from patients with COVID-19 and utilizing all the available slices in the CT scans. This ensures a more detailed and relevant approach to the specific challenges posed by this disease.

V. CONCLUSION

This study demonstrates the effectiveness of Mask R-CNN with a ResNet-50 backbone for segmenting lung parenchyma in COVID-19 chest CT scans. Optimal performance was achieved with a learning rate of 0.001 and without Data Augmentation, achieving a Dice Similarity Coefficient (DSC) of 93.4%. Future work will focus on expanding the dataset, exploring alternative backbone architectures, and enhancing segmentation in heterogeneous clinical settings.

Additionally, the use of pre-trained ImageNet weights significantly enhanced the model's performance. By capturing general visual features, these weights accelerated the training process and improved segmentation accuracy. This highlights the importance of transfer learning in specialized tasks, such as the segmentation of medical images from COVID-19 patients.

When evaluating the impact of data augmentation, the study found that while this technique achieved a maximum DSC of 92.7%, training without data augmentation outperformed it, yielding a higher DSC of 93.4%. These results suggest that, for this specific task, excluding data augmentation contributes to better segmentation performance, challenging common assumptions about the universal benefits of augmentation.

Extensive experiments were conducted using a public dataset for training and a custom dataset from Arequipa, Peru,

provided by the Research, Technology Transfer, and Software Development Center I+D+i - CiTeSoft. The findings validated the adaptability of the Mask R-CNN method, demonstrating its effectiveness even when applied to a regional population with potentially different characteristics from the training data.

This study opens new opportunities for improving and expanding the application of the proposed method. Future work will focus on experimenting with different backbone architectures, comparing models trained from scratch versus those using pre-trained weights, and determining the optimal number of epochs to achieve well-trained models without overfitting. Furthermore, testing will be extended to medical images of other anatomical regions to evaluate the method's adaptability and robustness across diverse clinical scenarios.

REFERENCES

[1] Y. Huérfano et al., "Imagenología médica: fundamentos y alcance," *Archivos Venezolanos de Farmacología y Terapéutica*, vol. 35, no. 3, pp. 71–76, 2016.

[2] Y. X. Tay, S. Kothan, S. Kada, S. Cai, and C. W. K. Lai, "Challenges and optimization strategies in medical imaging service delivery during COVID-19," *World Journal of Radiology*, vol. 13, no. 5, pp. 102–121, 2021.

[3] Z. Y. Zu et al., "Coronavirus disease 2019 (COVID-19): A perspective from China," *Radiology*, vol. 296, no. 2, pp. E15–E25, 2020.

[4] V. Varadarajan, M. Shabani, B. Ambale Venkatesh, and J. A. C. Lima, "Role of Imaging in Diagnosis and Management of COVID-19: A Multiorgan Multimodality Imaging Review," *Frontiers in Medicine*, vol. 8, p. 765975, 2021.

[5] N. Brandi and M. Renzulli, "The role of imaging in detecting and monitoring COVID-19 complications in the Intensive Care Unit (ICU) setting," *APS*, vol. 2, no. 3, p. 3, Jan. 2024.

[6] D. Gandhi et al., "Current role of imaging in COVID-19 infection with recent recommendations of point of care ultrasound in the contagion: a narrative review," *Annals of Translational Medicine*, vol. 8, no. 17, p. 1094, 2020.

[7] D. Dong *et al.*, "The Role of Imaging in the Detection and Management of COVID-19: A Review," *IEEE Reviews in Biomedical Engineering*, vol. 14, pp. 16-29, 2021.

[8] M. Alhasan and M. Hasaneen, "Digital imaging, technologies and artificial intelligence applications during COVID-19 pandemic," *Computerized Medical Imaging and Graphics*, vol. 91, p. 101933, 2021.

[9] P. Kumar, D. Jayaswal, M. Khan, and B. Singh, "A relative analysis of different CNN based models for COVID-19 detection using CXR and CT images," *Procedia Computer Science*, vol. 235, pp. 3163–3173, 2024.

[10] S. Podder, S. Bhattacharjee, and A. Roy, "An efficient method of detection of COVID-19 using Mask R-CNN on chest X-ray images," *AIMS Biophysics*, vol. 8, no. 3, pp. 281–290, 2021.

[11] D. Suganya and R. Kalpana, "Prognosticating various acute COVID lung disorders from COVID-19 patient using chest CT images," *Engineering Applications of Artificial Intelligence*, vol. 119, p. 105820, 2023.

[12] S. Aparna, K. Muppavaram, C. C. V. Ramayanam, and K. S. Sai Ramani, "Mask RCNN with RESNET50 for Dental Filling Detection," *Int. J. Adv. Comput. Sci. Appl.*, vol. 12, no. 10, 2021.

[13] M. Al Masarweh, O. Oluseyi, A. Alkafri, H. Alsmadi, and T. Alwadan, "Automatic Detection of Lumbar Spine Disc Herniation," Int. J. Adv. Comput. Sci. Appl., vol. 15, no. 11, 2024.

[14] A. A. Mashaal, M. A. A. Eldosoky, L. N. Mahdy, and K. A. Ezzat, "Classification of Human Sperms using ResNet-50 Deep Neural Network," *Int. J. Adv. Comput. Sci. Appl.*, vol. 14, no. 2, 2023.

[15] S. Suriyavarman and A. X. A. R. Annie, "Lung Nodule Segmentation and Classification using U-Net and Efficient-Net," Int. J. Adv. Comput. Sci. Appl., vol. 14, no. 7, 2023.

[16] J.-H. Shu, F.-D. Nian, M.-H. Yu, and X. Li, "An Improved Mask R-CNN Model for Multiorgan Segmentation," *Mathematical Problems in Engineering*, vol. 2020, Article ID 8351725, 11 pages, 2020.

[17] E. Dandıl and M. S. Yıldırım, "A Mask R-CNN based Approach for Automatic Lung Segmentation in Computed Tomography Scans," 2021 International Conference on INnovations in Intelligent SysTems and Applications (INISTA), Kocaeli, Turkey, 2021, pp. 1-6.

[18] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask R-CNN," *arXiv*, 2018.

[19] "COVID-19 CT Lung and Infection Segmentation Dataset," Zenodo, Apr. 20, 2020.