

MiniCovid-Unet: CT-Scan Lung Images Segmentation for COVID-19 Identification

Alvaro Salazar-Urbina¹, Elías Ventura-Molina², Cornelio Yáñez-Márquez^{1,1},
Mario Aldape-Pérez², Itzamá López-Yáñez²

¹ Instituto Politécnico Nacional,
Centro de Investigación en Computación,
Mexico

² Instituto Politécnico Nacional,
Centro de Innovación y Desarrollo,
Tecnológico en Cómputo,
Mexico

{asalazaru2020, cyanez}@cic.ipn.mx,
{eventuram, maldape, ilopez}@ipn.mx

Abstract. Detection and segmentation of Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV2 or COVID-19) is a difficult task due the different kinds of shapes, sizes and positions of the injury. Medical institutions have vast challenges because there is an urgent need for efficient tools to improve the diagnosis of COVID-19 patients. Computer tomography images (CT) are necessary for medical specialists to diagnose the patient's condition. Nevertheless, there is a lack of both in Medical Centers, mainly in rural areas. The manual analysis of CT images is time-consuming; in addition, most images have low contrast, and it is possible to find blood vessels in the background, so the difficulty of a suitable diagnosis increases. Nowadays, deep learning methods are an alternative method to perform the detection and segmentation task. In this work, we propose a novel light model to detect and identify COVID-19 using CT images: MiniCovid-Unet. It is an improved version of U-net; main differences reside on the decoder and encoder architecture, MiniCovid-Unet needs fewer convolution layers and filters because it focuses only on COVID-19 images. Also, as a result of employing fewer parameters, it can be trained in less time, and the resulting model is light enough to be downloaded to a mobile device. In this way, it is possible to have a quick and confident diagnosis in remote areas, where there exists an absence of internet connection and medical specialists.

Keywords. Deep learning, image segmentation, COVID-19, computer tomography, Mask R-CNN, Unet, MiniCovid-Unet.

1 Introduction

SARS-CoV2, better known as COVID-19 or Coronavirus, is an acute fatal disease identified in December 2019 in Wuhan province, China. This virus spread worldwide with great speed [1], declaring itself a pandemic on March 11, 2020 [2]. As of October 31, 2020, 45,428,731 cases have been confirmed in the world, causing 1,185,721 deaths [3].

COVID-19 is spread through droplets of secretion released from the mouth and nose of an infected individual [4] and is transmitted by direct or indirect contact (through contaminated objects and surfaces) to mucosal areas of the skin such as the mouth, nose, or tear ducts. Symptoms may include dry cough, fever, headache, fatigue, shortness of breath, loss of taste or smell, and shortness of breath. Symptoms usually appear 2 to 14 days after infection [5]. An early diagnosis is important because it is one of the most effective methods to stop the disease progression [6].

There are studies that have shown that COVID19 virus mainly attacks human lungs, after that there is a possibility of an infection and a lung disease [7]. Therefore, the diagnosis using a patient's chest computed tomography (CT) is so relevant.

The main aspect in a CT image of COVID is the presence of ground glass opacity (GGO) [8,9]. Some experts have identified three main types of anomalies in CT lung images related to COVID-19: ground glass opacification, consolidation and pleural effusion [10,11].

The manual observation is the main technique to decide whether the patients are infected or not. However, the job is exhausted and there aren't enough medical doctor's staff to do the job. So, an automatic segmentation system is necessary in order to identify and delimit the boundary of the region of interest in the lung [12].

Deep Learning (DL), a subfield of Machine Learning, is a tool commonly used in re-search areas for speech recognition, computer vision, natural language processing, and image processing [13]. The main advantage of DL methods is that they do not require experts to perform feature extraction; it is done automatically and implicitly by multiple flexible linear and non-linear processing units in a deep architecture.

In recent years, Deep Learning has been a useful tool for classifying medical images [14], among its techniques the convolutional neural network (CNN) model [15] stands out; a neural network inspired by the connectivity of the animal visual cortex. CNN is a multi-layer neural network that uses minimal processing of convolution operations on the pixels of the images. This technique extracts the relevant features from image sets to detect features regardless of their position.

Nowadays, the computer's power has made it possible to apply deep learning in a wide range of applications in the medical field, such as deciding whether a tumor is in a radiograph [16] or detect a cardiovascular risk. For the task about semantic segmentation, there is a constant improvement in the accuracy of segmentation with models such as Fully Convolutional Network (FCN) [17], U-net [18], Fast RCNN [19] and Mask RCNN [18] among others.

There are a lot of models that detect Covid19 cases from chest x ray images [20–22], yielding a prediction value of 90% [23]. However, this kind of model cannot provide a quantitative analysis of infection severity because they just classify between Covid19 and regular pneumonia.

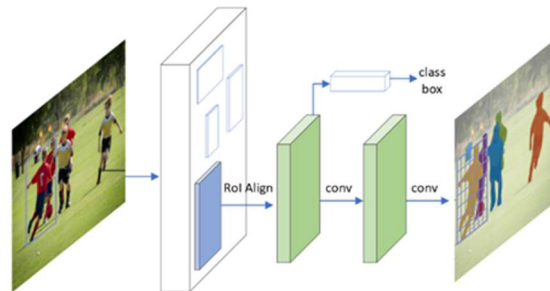


Fig. 1. Framework of the Mask R-CNN method used for detection and segmentation COVID-19 in CT images

2 Related Work

2.1 Mask R-CNN

Mask R-CNN is a framework focused on instance segmentation. This task combines elements of object detection (classify individual objects and localize every instance with a bounding box) and semantic segmentation (classify every pixel in a set of categories).

The Figure 1 shows a representation of the Mask R-CNN framework that contains two main phases; the first one consists of a Faster R-CNN architecture [19]. It has three elements: the backbone, the region proposal network (RPN) and the object detection [18]. The backbone takes advantage of a CNN architecture for image feature extraction and generating feature maps.

The RPN uses these maps and creates proposed bounding boxes (anchors) to do the object detection task, dispersed over each feature map. These bounding boxes or anchors are classified in two classes: positive anchors or foreground, which refers to the anchors located in regions that represent features on the objects to be detected, and the negative ones or background which are located outside these objects.

The positive anchors are used to perform a task called region of interest (ROI) alignment; they are centered to the located object and mark the ROIs for the next part. The object detection is the last part and classifies every class inside each ROI. The second phase consists of a new branch in order to do the instance segmentation task over

every detected object inside the image. This new branch is made by a fully convolutional mask [18].

2.2 Unet

Unet is one of most popular models for the task of image segmentation in the medical field. It was developed to understand in a visual way different types of images. And it is based on an encoder decoder neural network architecture. There are two main parts: con-tractive and expansive. The contracting one is built with several layers of convolution, filters of size 3 x 3 and strides in both directions, with ReLU layers at the end.

This part is important because it extracts the essential features of the input and the result is a feature vector of a particular dimension. The second part recover information from the contractive part by coping and cropping. However, the feature vector is built by convolutions and generate an output segmentation map. In this architecture the main component is the linking operation between the first and second part.

In this way the network gets correct information from the first part, so it could generate an accurate segmentation mask [18].

2.3 SegNet

SegNet is a deep fully convolutional neural network architecture for semantic segmentation [24]. It was originally designed for road and interior scene segmentation tasks. This requires the network to converge using an unbalanced dataset because the pixels of the road, sky, and buildings dominate. The main elements consist of an encoder network, a corresponding decoder followed by a pixel classification layer.

The encoder network is almost the same as the 13 convolutional layers of the VGG16 network [25]. The task of the decoder network is to map low resolution encoder feature maps to full input resolution feature maps for pixel classification. The main feature of SegNet is the way the decoder upsamples its lower resolution input feature maps; in this part, the decoder network uses clustering indices computed in the maximum clustering step of the corresponding encoder to perform non-linear upsampling.

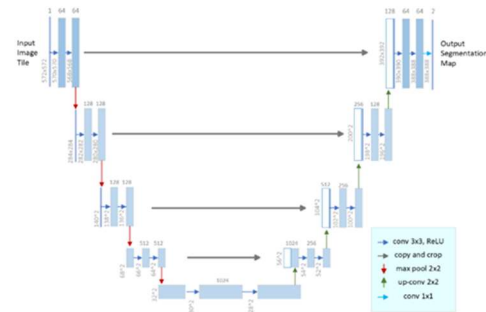


Fig. 2. U-net architecture [18]

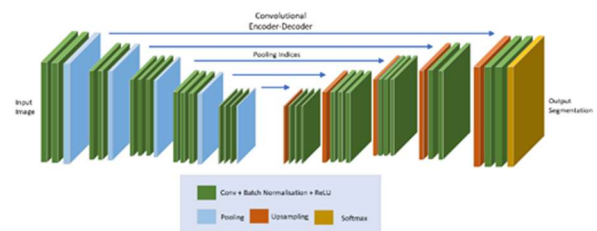


Fig. 3. SegNet architecture [24]

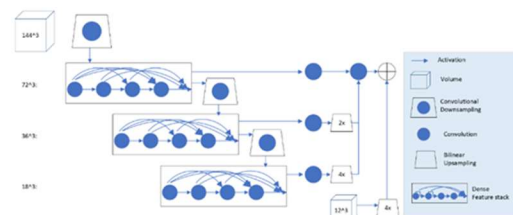


Fig. 4. Dense V-Net architecture [26]

2.4 Dense V-Net

Dense V-Net is a fully connected convolutional neural network that has performed well on the organ segmentation task. You can establish a voxel-voxel connection between the input and output images [26].

It consists of three layers of dense feature stacks whose outputs are concatenated after a convolution on the missing connection and bilinear oversampling [27]. There are 723 feature maps that are computed using a convolution step.

It then continues with a cascade of convolutions and dense feature stacks to generate activation maps with resolutions of three outputs. A convolution unit is applied on each output resolution to reduce the number of features. At the end it generates the segmentation logit.

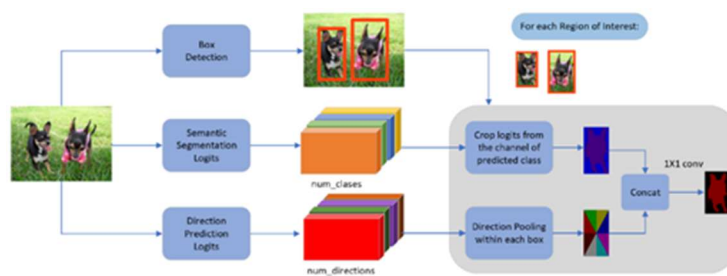


Fig. 5. MaskLab architecture [29]

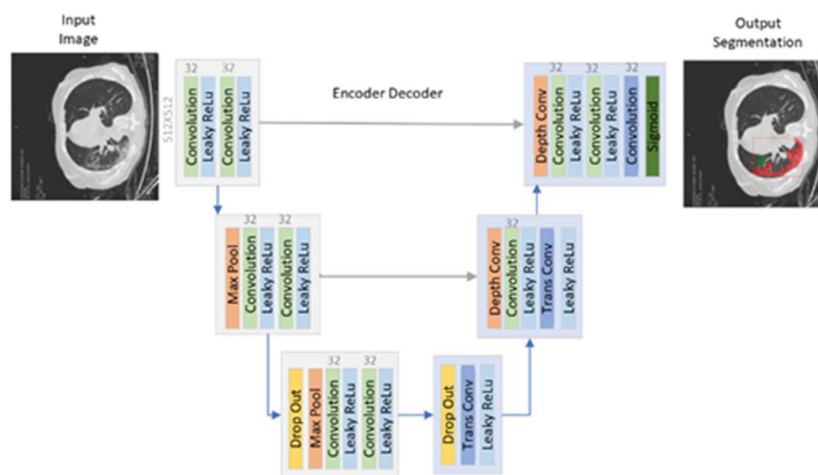


Fig. 6. MiniCovid-Unet architecture

Dense V-Net differs from V-Net [28] in several respects: the downsampling sub-network is a sequence of three dense feature stacks connected by downsampling strided convolutions; each skip connection is a unique convolution to the output of the corresponding dense feature stack. The upsampling network comprises bilinear upsampling to the final segmentation resolution.

2.5 MaskLab

MaskLab is an instance segmentation model [29], refines object detection with address and semantic features based on Faster R-CNN [19]. This model produces three out-puts: box detection, semantic segmentation logits for pixel classification, and direction prediction logits to predict the direction of each pixel around its instance center.

Therefore, MaskLab is based on the Faster R-CNN object detector, the predicted frames provide precise location of object instances. Within each region of interest, MaskLab performs fore-ground and background segmentation by combining semantic and direction prediction.

2.6 MiniCovid-Unet

The ground glass opacities are important features of COVID-19 infection regions in CT images scans. However, these image characteristics cannot be extracted efficiently by conventional CNNs, where the original images are taken as input and the learning processes begin from pixel level features. Hence, to reflect more regional features of infections we use different filters to highlight the region of interest.

As shown in Figure 6, the proposed COVID-19 segmentation model applies the Unet like structure

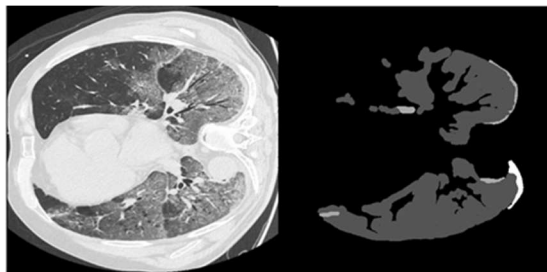


Fig. 7. Image and mask sample. CT scan (left) and labeled classes (right), where dark gray is ground glass opacities, gray is pleural effusion and white is consolidation

as backbone. There are two basic sections: contractive and expansive. We have used the activation function Leaky ReLU in all blocks of layers because it is faster and it reduces the complexity of the network. Every convolution layer has 32 filters for images of 512 X 512 pixels.

There are less layers of convolution because the improvement was slightly better, however it increased the time of training and the computer resources needed. The model we proposed has good performance for computers with limited resources and is small enough to use in a mobile device.

3 Materials and Methods

Images of the dataset are Computed Tomography (CT) scans that belong to the Italian Society of Medical and Interventional Radiology [30]. The dataset contains one hundred one-slice CT scans in png format, whose dimensions are 512 x 512. There are also masks that show the region labeled by experts of the medical field [31].

In the original dataset there are three kinds of injuries related with Covid19: ground-glass opacities, consolidation and pleural effusion (Figure 7). However, we just try to identify whether an image has an injury in the lung and where it is located. The images are of people who had been infected with COVID-19.

The training of Mask R-CNN used a total of 72 of lung CT images and lung segmentation masks labels, the original image's size remained and no data enhancement was used for training. The validation set used 18 images and its masks. The

training set iteration was 16 with 500 steps per iteration. The learning rate was 0.001. We set aside 10 images to visualize the performance of the trained and validated model with the training and validation data sets.

For this experiment the backbone CNN architecture used was ResNet50 because of the small graphic card [32]. The experiment used pre-trained COCO weights [18,33]. The total number of parameters for Mask R-CNN is 44,662,942.

There is a problem with imbalance classes, because the task is to segment only the COVID-19 infected region. But with this configuration we have two classes: COVID-19 region and non-COVID-19. In this case, we have more pixels from healthy regions ($2.4482e + 7$) than from infected regions ($2.119975e + 6$). So, the unbalance ratio is 11 and the data set is unbalanced, that's the reason we have chosen metrics for the segmentation task.

3.1 Implementation Details

The Jupyter notebook interactive development environment was used to build and visualize the model and results. Python 3.6 was used as a programming language and the hardware configured to execute the experiments was a personal computer with a processor Intel(R) Core (TM) i7-6700 CPU @ 3.40GHz with 8 cores. NVIDIA GeForce GTX 1050 Ti (GPU 0), CUDA Toolkit 10.0 and CUDNN 7.4.1 were used to drop the time training.

Be-cause of the small GPU the training configuration was set to use one image in every step and it was needed to use a small backbone (resnet50). On average the full execution of this model took 57 minutes.

3.2 Evaluation Metrics

In order to evaluate the performance of the models, we used the following classification and segmentation measures: precision, recall, Dice coefficient and mean Intersection over Union (mIoU). These metrics are also used in the medical field, and are defined be-low.

Precision is the ratio of pixels correctly predicted as COVID-19 divided by the total predicted as COVI-19:

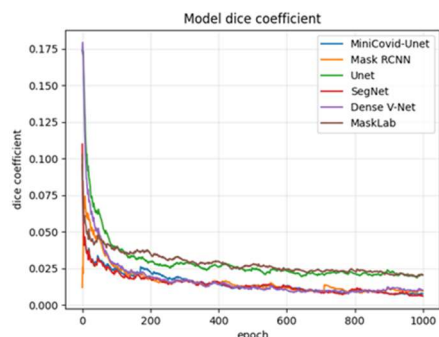


Fig. 8. Loss during the training phase

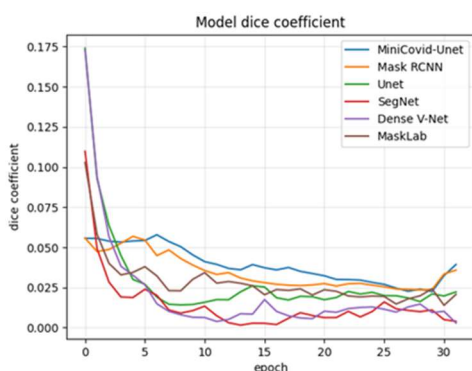


Fig. 9. Loss during the validation phase

$$Precision = \frac{TP}{TP+FP}, \quad (1)$$

where TP is the true positive (i.e., the number of pixels labeled as COVID-19 correctly) and FP refers to the false positive (i.e., the number of pixels labeled as COVID-19 wrong).

Recall is the ratio of pixels correctly predicted as COVID-19 divided by total number of actual COVID-19:

$$Recall = \frac{TP}{TP+FN}, \quad (2)$$

where FN refers to the false negatives (i.e., the pixels that are labeled wrong as non-COVID-19).

However, these two measures are not frequently used as evaluation metrics because of their sensibility to segment size, in other words, they penalize errors in small segments more than in large ones [28, 34, 35].

Dice coefficient or *Dice score* (DSC) is a metric for image segmentation:

$$Dice = \frac{2|A \cap B|}{|A| + |B|}, \quad (3)$$

where A and B refers to the predicted and ground truth masks.

4 Results and Discussion

All models that we have used in this work predict a probability for every pixel and we have to set a threshold in order to identify if a pixel is in the segment of COVID-19 or is in the healthy part. So, we have decided that the threshold value of 0.9 is the best to do the Task.

We used the validation method five-fold cross validation to evaluate the segmentation performance of the models on the COVID-19 dataset. First of all, we set aside 10 images to test the model after we have trained it. With the remaining 90 images, the new data set is used to apply five-cross validation.

We divided the new dataset into 5 parts, one of which was selected as the validation set and the other four parts were used for the training set in order to train the model. When the training had finished the loss, metrics were calculated and we repeated all the experiments until all the parts were used as a validation set, then the average of metrics was calculated to get the performance evaluation value of the model.

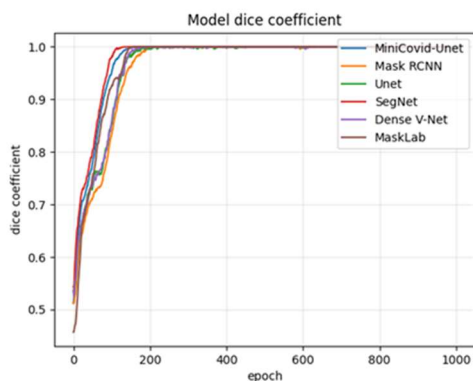
Figures 8 and 9 show the loss during the training and validation phase. At the beginning of the training phase, the difference between all the models is noticeable, but over time, all the ones converge. The models only detect where the lesion is, so we don't ask about the class of lesion.

Table 1 shows the metrics to evaluate the performance of the model. The Dice metric can be used to compare predicted segmentation pixels and their corresponding ground truth. Dense V-Net is the model that has the best performance in terms of metric accuracy. On the other hand, the proposed model achieves a better performance with respect to the Dice and Recall metric.

All models were able to detect the foreground from the back-ground; however, they were unable to detect the lesion class of the background. The best scores were obtained during the training phase compared to the validation phase, as can be seen in Figures 10 and 11.

Table 1. Performance metrics associated with different algorithms for the images in the testing dataset

Method	Dice	Precision	Recall
Mask R-CNN			
Unet	0.7801	0.7857	0.7333
SegNet	0.8202	0.6190	0.8667
Dense V-Net	0.8001	0.7667	0.8333
MaskLab	0.7905	0.8667	0.8467
	0.7885	0.8001	0.8402
Proposed	0.8301	0.8254	0.8684

**Fig. 10.** Dice coefficient during the training phase

4.1 Inference

Figure 12 illustrates the segmentation results of lung infections from an example of lung CT slices taken from the test set using different segmentation networks.

The original image is on the left side (a), the expertly labeled mask is on the right side (b). All models have located the correct position on the image of the COVID-19 related injury, but do not retrieve the exact shape of the injury.

Unet misses true infected areas with small sizes. Mask RCNN works better than Unet to determine the infected region, however, some tissues close to infections are segmented incorrectly. Segnet and Dense V-Net provide good performance in segmenting medium-sized infection regions, but several overestimates of normal tissues as infections.

MaskLab cannot provide full segmentation of some regions. On the contrary, the proposed MiniCovid-Unet provides superior performance to

previous methods, regarding the recognition and segmentation of small and medium infections.

The shape of the infected area was complex and could be located anywhere within the image, the contrast between the infected and healthy parts was low.

In addition, the original Mask R-CNN model has been trained with millions of images of people and different types of objects, which could explain the low score against MiniCovid-Unet.

Furthermore, the other models were unable to retrieve the exact shape of the COVID-19 lesion, as can be seen in Table 1.

5 Conclusion and Future Work

In this paper, we propose the MiniCovid-Unet network with novel structure for COVID-19 infection region segmentation in lung CT slices. We also presented other models applied to detect and segment injuries related to COVID-19.

The models were selected because it is simple to implement for a custom dataset of images. However, a GPU is necessary in order to train the model in a reasonable time.

All models were able to identify the regions where lesions were found, but had difficulties in correctly segmenting the shape of the lesion. Figure 12 shows that a healthy lung could be differentiated from a diseased one, and even completely healthy lungs could be detected. However, the results for the segmentation task were poor.

Although the models can identify the injury, it does not indicate the type of injury. We used a small dataset available for the segmentation task, however the MiniCovid-Unet's results obtained so far in this work represent an alternative to use deep learning to help in the objective diagnosis of COVID-19 using CT images of the Lung.

As future work, we want to get more images to train the framework. We also hope to be able to perform the segmentation taking into account the three existing classes in the dataset.

It is also proposed to make a comparison against other models such as U-Net++ [36], which are frameworks focused on COVID-19 medical images.

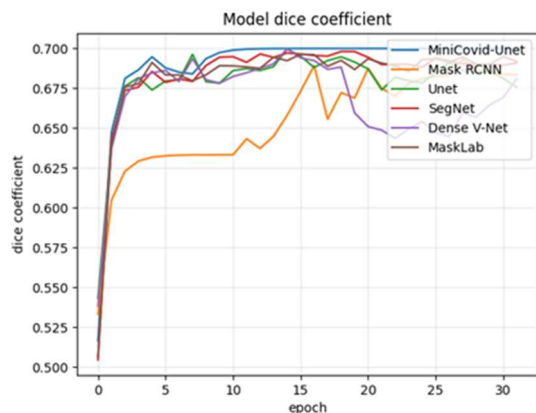


Fig. 11. Dice coefficient during the validation phase

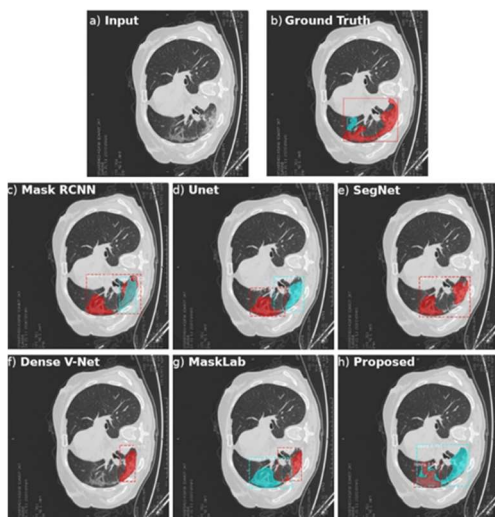


Fig. 12. Visual comparison of COVID-19 segmentation results. From left to right, (a) the first image is input CT scans, (b) the second is mask label visualization image or ground truth. The models (c) to (h) are the ones listed in Table 1. The color labeled part is the infected area

Acknowledgments

First author thanks to CONAHCYT for the scholarship granted. This research received no external funding.

References

1. Platto, S., Wang, Y., Zhou, J., Carafoli, E. (2021). History of the COVID-19 pandemic: Origin,

explosion, worldwide spreading. *Biochemical and Biophysical Research Communications*, Vol. 538, pp. 14–23. DOI: 10.1016/j.bbrc.2020.10.087.

2. WHO (2020). Director-General's opening remarks at the media briefing on COVID-19. <https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020>.
3. WHO (2020). Coronavirus (COVID-19) Dashboard, 2020. <https://covid19.who.int/>.
4. Li, S., Li, S., Disoma, C., Zheng, R., Zhou, M., Razzaq, A., Liu, P., Zhou, Y., Dong, Z., Du, A., Peng, J., Hu, L., Huang, J., Feng, P., Jiang, T., Xia Z. (2021). SARS-CoV-2: Mechanism of infection and emerging technologies for future prospects. *Reviews in Medical Virology*, Vol. 31, No. 2, pp. e2168.31. DOI: 10.1002/rmv.2168.
5. Chen, V. C. C., Wong, S. C., Chen, J. H. K., Yip, C. C. Y., Chuang, V. W. M., Tsang, O. T. Y., Sridhar, S., Chan, J. F. W., Ho, P. L., Yuen, K. Y. (2020). Escalating infection control response to the rapidly evolving epidemiology of the coronavirus disease 2019 (COVID-19) due to SARS-CoV-2 in Hong Kong. *Infection Control & Hospital Epidemiology*, 41, 493–498. DOI: 10.1017/ice.2020.58.
6. Chen, M., Tu, C., Tan, C., Zheng, X., Wang, X., Wu, J., Huang, Y., Wang, Z., Yan, Y., Li, Z., Shan, H., Liu, J., Huang, J. (2020). Key to successful treatment of COVID-19: accurate identification of severe risks and early intervention of disease progression. *MedRxiv*. DOI: 10.1101/2020.04.06.20054890.
7. Liu, Z., Jin, C., Wu, C. C., Liang, T., Zhao, H., Wang, Y., Wang, Z., Li, F., Zhou, J., Cai, S., Zeng, L., Yang, J. (2020). Association between Initial chest CT or clinical features and clinical course in patients with coronavirus disease 2019 pneumonia. *Korean J Radiology*, Vol. 21, No. 6, pp. 736–745. DOI: 10.3348/kjr.2020.0171.
8. Zhou, X., Pu, Y., Zhang, D., Xia, Y., Guan, Y., Liu, S., Fan, L. (2022). CT findings and dynamic imaging changes of COVID-19 in 2908 patients: A systematic review and meta-analysis. *Acta Radiologica*, Vol. 63, No. 3, pp. 291–310. DOI: 10.1177/0284185121992655.
9. Suri, J. S., Agarwal, S., Chabert, G. L., Carriero, A., Paschè, A., Danna, P. S. C., Saba, L., Mehmedović, A., Faa, G., Singh, I. M., Turk, M., Chadha, P. S., Johri, A. M., Khanna, N. N., Mavrogeni, S., Laird, J. R., Pareek, G., Miner, M., Sobel, D. W., Balestrieri, A. (2022). COVLIAS 1.0 lesion vs. medseg: An artificial intelligence framework for automated lesion segmentation in

- COVID-19 lung computed tomography scans. *Diagnostics*, Vol. 12, No. 5, pp. 1283. DOI: 10.3390/diagnostics12051283.
10. **Shi, H., Han, X., Jiang, N., Cao, Y., Alwalid, O., Gu, J., Fan, Y., Zheng, C. (2020).** Radiological findings from 81 patients with COVID-19 pneumonia in Wuhan, China: a descriptive study. *The Lancet. Infectious Diseases*, Vol. 20, pp. 425–434. DOI: 10.1016/S1473-3099(20)30086-4.
 11. **Ye, Z., Zhang, Y., Wang, Y., Huang, Z., Song, B. (2020).** Chest CT manifestations of new coronavirus disease 2019 (COVID-19): A pictorial review. *European Radiology*, Vol. 30, pp. 4381–4389. DOI: 10.1007/s00330-020-06801-0.
 12. **Al-Shehri, W., Almalki, J., Mehmood, R., Alsaif, K., Alshahrani, S. M., Jannah, N., Alangari, S. (2022).** A novel COVID-19 detection technique using deep learning-based approaches. *Sustainability*, Vol. 14, No. 19, DOI: 10.3390/su141912222.
 13. **LeCun, Y., Bengio, Y., Hinton, G. (2015).** Deep learning. *Nature*, Vol. 521, pp. 436. DOI: 10.1038/nature14539.
 14. **Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., van der Laak, J. A., van Ginneken, B., Sánchez, C. I. (2017).** A survey on deep learning in medical image analysis. *Medical Image Analysis*, Vol. 42, pp. 60–88. DOI: 10.1016/j.media.2017.07.005.
 15. **Gatys, L. A., Ecker, A. S., Bethge, M. (2017).** Texture and art with deep neural networks. *Current Opinion in Neurobiology*, Vol. 46, pp. 178–186. DOI: 10.1016/j.conb.2017.08.019.
 16. **Wang, S., Yang, D. M., Rong, R., Zhan, X., Fujimoto, J., Liu, H., Minna, J., Wistuba, I. I., Xie, Y., Xiao, G. (2019).** Artificial intelligence in lung cancer pathology image analysis. *Cancers*, Vol. 11, No. 11, pp. 1673. DOI: 10.3390/cancers11111673.
 17. **Shelhamer, E., Long, J., Darrell, T. (2017).** Fully convolutional networks for semantic segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 39, pp. 640–651. DOI: 10.1109/TPAMI.2016.2572683.
 18. **Minaee, S., Boykov, Y., Porikli, F., Plaza, A., Kehtarnavaz, N., Terzopoulos, D. (2022).** Image segmentation using deep learning: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 44, pp. 3523–3542. DOI: 10.1109/TPAMI.2021.3059968.
 19. **Ren, S., He, K., Girshick, R., Sun, J. (2017).** Faster R-CNN: towards real-time object detection with region proposal networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 39, pp. 1137–1149. DOI: 10.1109/TPAMI.2016.2577031.
 20. **Luján-García, J., Villuendas-Rey, Y., López-Yáñez, I., Camacho-Nieto, O., Yáñez-Márquez, C. (2021).** Nanochest-net: A simple convolutional network for radiological studies classification. *Diagnostics*, Vol. 11, No. 5, pp. 775. DOI: 10.3390/diagnostics11050775.
 21. **Constantinou, M., Exarchos, T., Vrahatis, A. G., Vlamos, P. (2023).** COVID-19 classification on chest X-ray images using deep learning methods. *International Journal of Environmental Research and Public Health*, Vol. 20, No. 3, pp. 2035. DOI: 10.3390/ijerph20032035.
 22. **Teixeira, L. O., Pereira, R. M., Bertolini, D., Oliveira, L. S., Nanni, L., Cavalcanti, G. D. C., Costa, Y. M. G. (2021).** Impact of lung segmentation on the diagnosis and explanation of COVID-19 in chest X-ray images. *Sensors*, Vol. 21, No. 21, pp. 7116. DOI: 10.3390/s21217116.
 23. **Wang, L., Lin, Z. Q., Wong, A. (2020).** COVID-Net: a tailored deep convolutional neural network design for detection of COVID-19 cases from chest X-ray images. *Scientific Reports*, Vol. 10, pp. 19549. DOI: 10.1038/s41598-020-76550-z.
 24. **Badrinarayanan, V., Kendall, A., Cipolla, R. (2017).** SegNet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 39, No. 12, pp. 2481–2495. DOI: 10.1109/TPAMI.2016.2644615.
 25. **Jiang, Z. P., Liu, Y. Y., Shao, Z. E., Huang, K. W. (2021).** An improved VGG16 model for pneumonia image classification. *Applied Sciences*, Vol. 11, No. 23, pp. 11185. DOI: 10.3390/app112311185.
 26. **Gibson, E., Giganti, F., Hu, Y., Bonmati, E., Bandula, S., Gurusamy, K., Davidson, B., Pereira, S. P., Clarkson, M. J., Barratt, D. C. (2018).** Automatic multiorgan segmentation on abdominal CT with dense V-Networks. *IEEE Transactions on Medical Imaging*, Vol. 37, No. 8, pp. 1822–1834. DOI: 10.1109/TMI.2018.2806309.
 27. **Wu, J., Hu, W., Wen, Y., Tu, W., Liu, X. (2020).** Skin lesion classification using densely connected convolutional networks with attention residual learning. *Sensors*, Vol. 20, No. 24, pp. 7080. DOI: 10.3390/s20247080.
 28. **Liu, X., Song, L., Liu, S., Zhang, Y. (2021).** A review of deep-learning-based medical image segmentation methods. *Sustainability*, Vol. 13, No. 3, pp. 1224. DOI: 10.3390/su13031224.
 29. **Sharma, R., Saqib, M., Lin, C. T., Blumenstein, M. (2022).** A survey on object instance segmentation.

- SN Computer Science, Vol. 3, No. 6, pp. 499. DOI: 10.1007/s42979-022-01407-3.
30. **COVID-19 CT Segmentation Dataset (2020)**. <http://medicalsegmentation.com/covid19/>.
31. **Hofmanninger, J., Prayer, F., Pan, J., Röhrich, S., Prosch, H., Langs, G. (2020)**. Automatic lung segmentation in routine imaging is primarily a data diversity problem, not a methodology problem. *European Radiology Experimental*, Vol. 4, No. 1, pp. 1–3. DOI: 10.1186/s41747-020-00173-2.
32. **Xie, J., Pang, Y., Nie, J., Cao, J., Han, J. (2022)**. Latent feature pyramid network for object detection. *IEEE Transactions on Multimedia*, Vol. 25, pp. 2153–2163. DOI: 10.1109/TMM.2022.3143707.
33. **Rostianingsih, S., Setiawan, A., Halim, C. I. (2020)**. COCO (creating common object in context) dataset for chemistry apparatus. *Procedia Computer Science*, Vol. 171, pp. 2445–2452. DOI: 10.1016/j.procs.2020.04.264.
34. **Udupa, J. K., LeBlanc, V. R., Zhuge, Y., Imielinska, C., Schmidt, H., Currie, L. M., Hirsch, B. E., Woodburn, J. (2006)**. A framework for evaluating image segmentation algorithms. *Computerized Medical Imaging and Graphics*, Vol. 30, No. 2, pp. 75–87. DOI: 10.1016/j.compmedimag.2005.12.001.
35. **Moorthy, J., Gandhi, U. D. (2022)**. A Survey on medical image segmentation based on deep learning techniques. *Big Data and Cognitive Computing*, Vol. 6, No. 4, 117. DOI: 10.3390/bdcc6040117.
36. **Zhou, Z., Siddiquee, M. M. R., Tajbakhsh, N., Liang, J. (2020)**. UNet++: redesigning skip connections to exploit multiscale features in image segmentation. *IEEE Transactions on Medical Imaging*, Vol. 39, No. 6, pp. 1856–1867. DOI: 10.1109/TMI.2019.2959609.

*Article received on 06/09/2023; accepted on 18/10/2023.
Corresponding author is Cornelio Yáñez-Márquez.