ISSN 1870-4069

SmallCovid-Net: COVID-19 Segmentation Using CT Scan Lung Images

Alvaro Salazar Urbina[1], Elías Ventura Molina[2], Sergio Flores Cortes[3], Antonio de Jesús Méndez Agüero[4], Daniel Jiménez Alcantar [5]

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

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

[3] Instituto Politécnico Nacional,
Escuela Superior de Ingeniería Mecánica y Eléctrica Unidad Zacatenco,
Mexico

[4] Instituto Politécnico Nacional,
Dirección de Sistemas Informáticos,
Mexico

[5] Comisión Federal de Electricidad,
Departamento de Mantenimiento,
Mexico

{asalazaru, eventuram, amendeza}@ipn.mx,
sfloresc2400@alumno.ipn.mx, daniel.jimenez@cfe.mx

Abstract. Since the beginning of the COVID-19 pandemic declared in March 2020, the medical field has faced various challenges. If an injured region can be detected and segmented automatically, it would be a huge help for doctors to diagnoses the patient’s infection. However, this is a difficult task because the virus can have different shapes and sizes, more- over it can also be located in any region of the lung. In order to diagnose a patient, lung Computer Tomography Scan (CT) images are required. Nevertheless, the manual review of CT images is a hard task because it requires medical specialist and is a slow process. To be able to segment automatically we proposed a new network named SmallCovid-Net, which is an improved version of the Segnet and Unet model. Because it only focuses on CT COVID-19 scans, the network requires few convolution and filter layers. As a consequence, it requires less training time and has obtained competitive results.

Keywords: Image segmentation, deep learning, COVID-19, computer tomography, Unet, mask R-CNN.

1 Introduction

The SARS-Cov2 or COVID-19 is an infectious and acute fatal disease that has had a devastating effect on the lives of people around the world. It was identified in the Chinese province of Wuhan in December 2019 and after a few months it spread throughout the world [4].

The COVID-19 infection begins in the mucous membranes of the throat and spreads to the lungs through the respiratory tract. Symptoms may include fever, dry cough, difficulty breathing, fatigue, loss of smell and taste, and dizziness. Which can appear from 2 to 14 days after being infected.

For the diagnosis of COVID-19, medical images from computed tomography (CT) and scanned x-ray (XR) have been used with good results [5]. However, in order to adequately interpret the images, it is necessary to have expert radiologist because the disease is very similar to other lung diseases. Furthermore, due to the complex nature of COVID-19 and the high degree of mortality, accurate and time-consuming diagnosis is necessary [13].

The COVID-19 virus primarily attacks the lungs, subsequently there is the possibility of developing infection and lung disease [11]. In order to make an adequate diagnosis, CT images are used. A characteristic that stands out in CT images of the COVID-19 virus is the appearance of ground glass opacity (GGO). However, there are more anomalies in the images that have a relationship with COVID-19, which are: consolidation and pleural effusion [17-19]. To decide whether a patient is healthy or sick, physical observation is used. Nevertheless, this is strenuous work and doctors with sufficient experience are rare. Therefore, a system is necessary that allows automatic classification and indicates the region of interest in the image [21].

Deep Learning (DL) is a tool regularly used in different areas of research such as: computer vision, speech recognition, image processing and natural language, among others. In models that use DL, feature extraction is automatic, so medical experts are not required to perform it. With a deep architecture and multiple processing units, this task can be achieved [9].

DL has been used to perform the classification of medical images with good results, using techniques that use convolutional neural network (CNN) models [10-18]. CNNs use a minimal process of convolution operations on each pixel of the images to extract the set of relevant features regardless of their position [6].

Also, DL has been applied in a wide range of the medical field, CNN has been used to determine if an x-ray contains a malignant tumor, to indicate risks of heart disease, among others [20]. For the segmentation task that indicates the region of interest, architectures such as Full Convolutional Networks (FCN) and U-Net have been used [15].

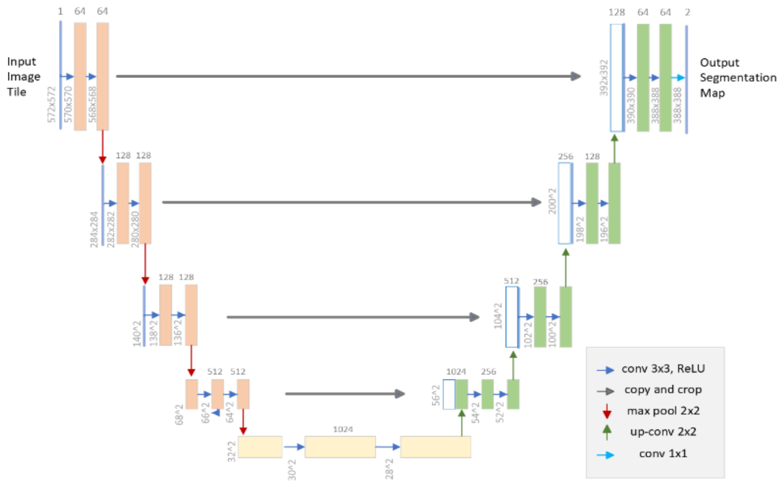

Fig. 1. The U-net model [12].

2 Related Work

2.1 U-Net

In the medical field, one of the most popular models for the segmentation task, which is based on an encoder-decoder architecture. It was proposed to have a tool that could use different types of images that would allow the doctor to have a better overview of the injures. It consists of two parts: a contractive path to capture context and an asymmetric expansion that enables precise locations [12].

The down sampling or contracting part has an architecture that resembles a fully convolutional network (FCN) that extracts features with 3X convolutions. The importance of this part is that it can extract the main features from the input image and the result is a feature vector.

On the other hand, the expansion part recovers the information from the first part by copying and trimming. For this phase the feature vector is constructed by convolutions and generates a segmentation output map.

In this architecture, an important part is the link operation between the two parts previously explained. Therefore, the model can generate an accurate segmentation mask (see Figure 1).

There are works based on this model that allow segmentation of medical images, but they require good quality images and the training time as well as its parameters are greater [16].

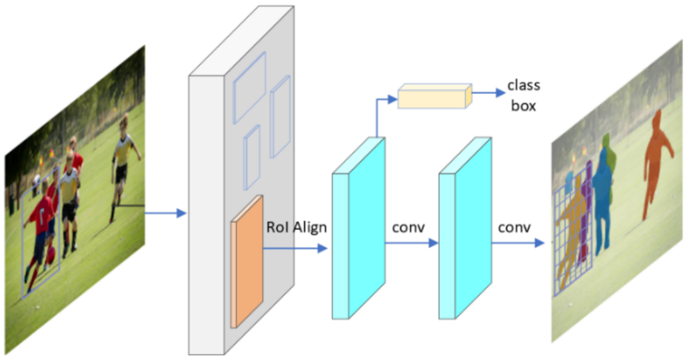

Fig. 2. Mask R-CNN used for detection and segmentation.

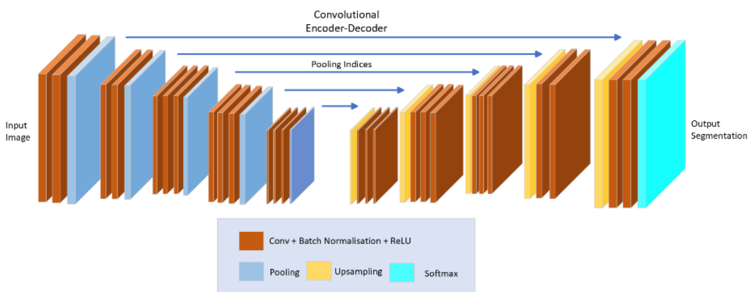

Fig. 3. The SegNet model [3].

2.2 Mask R-CNN

Instance segmentation is challenging due to the fact it requires the correct identification of all objects in the image as well as the precise segmentation of each instance. Therefore, this combines elements of computer vision task such as object detection (the objective is to individually classify objects and locate using bounding box) and semantic segmentation (the objective is to classify each pixel within a fixed set of categories) [12].

Badri Narayanan et al. proposed an encoder decoder architecture for image segmentation [3]. The model does not have fully connected layers, so it is completely convolutional. Its original objective was to do road segmentation. The network uses an unbalanced data set because the pixels of roads and buildings predominate. Segnet is composed of an encoder network, a corresponding decoder and a pixel-wise classification layer. The topology of the encoder network is identical to the 13 convolutional layers of the VGG16 network [8].

Mask R-CNN is a general framework for object instance segmentation. This approach identifies objects in an image while generating a segmentation mask for each instance. This architecture is simple to train and contains two main phases (Figure 2).

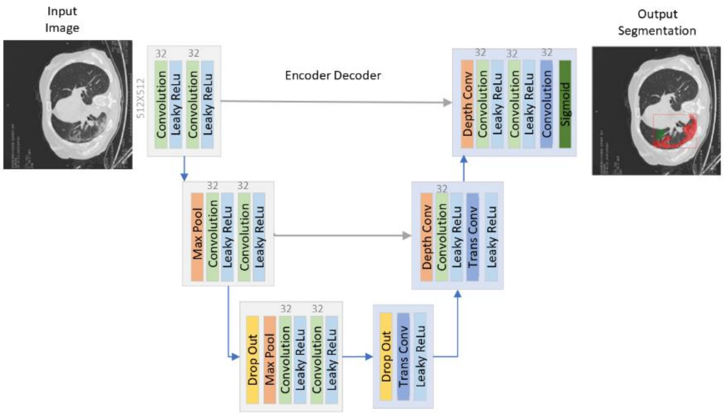

Fig. 4. SmallCovid-Net model.

The first one is the R-CNN architecture [14]; it has three elements: the backbone, Region Proposal Network (RPN) and object detection. The backbone is used for image feature extraction and map generation. Maps are used by RPN and integrate bounding boxes to achieve the detection task. Bounding boxes are classified as positive or foreground. The positives are used to create a Region of Interest (ROI) alignment. The second phase uses a new branch to perform the instance segmentation task [12, 2].

2.3 Segnet

Its main novelty is the way in which the decoder performs upsampling of a low-resolution feature map. The up-sampled maps are convolved with trained filters to produce dense feature maps.

Segnet is designed to be efficient in terms of memory and computing time during the inference phase because its objective lies in the segmentation of panorama images (Figure 3). It uses the stochastic gradient and a smaller number of train- able parameters compared to other architecture.

2.4 SmallCovid-Net

One of the most important features of COVID-19 is the detection of ground-glass opacities on a CT scan. GGO refers to an area of interest fading in the lung on a CT scan. Nevertheless, it is not possible to extract GGO using conventional CNN. In this architecture, the original image is the input and the training and learning process starts from a pixel-level feature. So, to highlight more areas of infections, we have used different filters.

The main objective of the proposed model is to perform the segmentation task of COVID-19 CT images to locate and mark the region containing the lesion. It uses the SegNet architecture as a base because it contains two phases: convolution and deconvolution. We have used the ReLU activation function in the convolution block

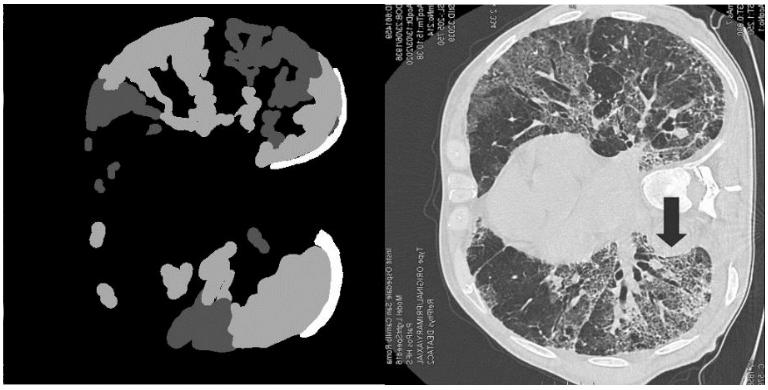

Fig. 5. CT scan image (right) and mask sample (left). In the right white is consolidation, dark gray is ground glass opacities and gray is pleural effusion.

because we only handle positive values, and it reduces the complexity of the model. Convolution blocks of 24 filter have been used for images of 256 X 256 pixels.

The fact that there are fewer layers means that there is an improvement in the training phase. On the other hand, the time for this phase has increased due to the needs of computing resources. Finally, the proposed model has obtained good results compared to other state of the art models.

3 Materials and Methods

The main source of materials for this work are the images from the Italian Society of Medicine which consists of scanned computed tomography images. The images have been segmented by radiological experts and for each image there is its segmentation mask counterpart. The format of the images is gray scale and their dimensions are 512 x 512 pixels [1]. The main source of materials for this work are the images from the Italian Society of Medicine which consists of scanned computed tomography images.

The images have been segmented by radiological experts and for each image there is its segmentation mask counterpart. The format of the images is gray scale and their dimensions are 512 x 512 pixels [1].

In this work the objective is to do a semantic segmentation although in the selected data set three types of lesions related to COVID-19 can be detected: consolidation, pleural effusion and ground glass opacities. The images belong to people who were infected in the early phases of the COVID-19 pandemic in European countries [7].

Of the 100 images that exist in the database, 72 were used for the training phase, 18 for the validation and 10 were reserved for the testing one. A process was carried out on the images to reduce them to 256 x 256 pixels.

The training phase consisted of a set of 20 iterations with 600 steps per iteration. The learning ratio was 0.0001 for the Gradient Descent optimizer.

The number of classes was two: the part of the image that contains the lesion and the part that is not a lesion. In this aspect, there is a considerable class imbalance because there are more pixels in healthy regions than in the infected ones. In this work the objective is to do a semantic segmentation although in the selected data set three types of lesions related to COVID-19 can be detected: consolidation, pleural effusion and ground glass opacities. The images belong to people who were infected in the early phases of the COVID-19 pandemic in European countries [7].

Of the 100 images that exist in the database, 72 were used for the training phase, 18 for the validation and 10 were reserved for the testing one. A process was carried out on the images to reduce them to 256 x 256 pixels.

The training phase consisted of a set of 20 iterations with 600 steps per iteration. The learning ratio was 0.0001 for the Gradient Descent optimizer.

The number of classes was two: the part of the image that contains the lesion and the part that is not a lesion. In this aspect, there is a considerable class imbalance because there are more pixels in healthy regions than in the infected ones.

3.1 Network Training

The training process was been done using open source tools such as Python 3.9 as a programming language and the hardware configured to execute the experiments was a personal computer with a processor Intel(R) Core(TM)©i9- 6700 CPU @ 4.20 Ghz with 16 cores and NVIDIA©GeForce GTX 1050 Ti, CUDA Toolkit 10.0 and CUDNN 7.4.1 were used to drop the time training.

Using a GPU reduce the training process and can complete simple task faster because it is possible to decompose complex task in simple ones and process them in parallel.

3.2 Performance Measures

To fully quantify the performance of our model, we have used performance mea- sures for the classification and segmentation task: precision, dice coefficient and recall. These measures are used in the medical field:

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

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

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

In the Equation 1 and 3 TP or true positive is the number of pixels labeled by the model as COVID-19 is correct. The FP or false positive is number of pixels labeled by the model as COVID-19 is wrong. The FN or false negative is the number of pixels labeled by the model as non-COVID-19 is wrong. The Equations 2 A refers to the predicted mask and B is the ground truth one.

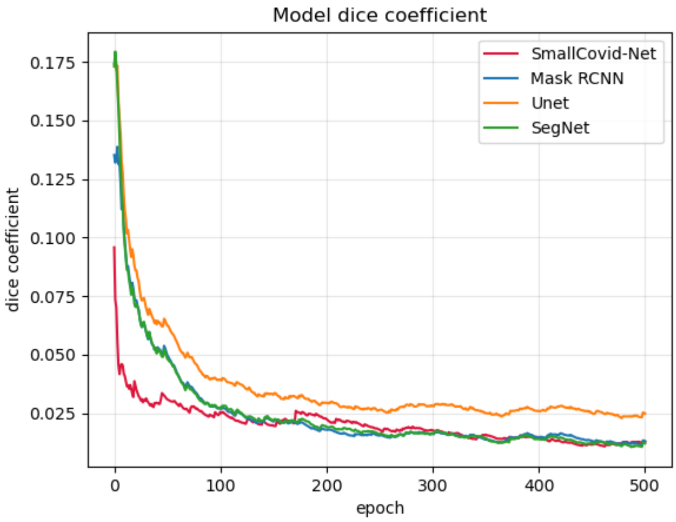

Fig. 6. Training loss.

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

Method	Dice	Precision	Recall
Mask R-CNN	0.6801	0.6857	0.6333
Unet	**0.7202**	0.5190	**0.7667**
SegNet	0.7001	0.6667	0.7333
Proposed	0.7011	**0.7121**	0.7321

4 Results and Discussion

The proposed model uses a probability to determine whether each pixel is in a segment classified as COVID-19 or is a healthy region. In order to achieve the above, it is necessary to select a threshold that allows reliable results, which is the reason it has been decided to place the threshold at 0.8.

Due to the small size of the data set used for the segmentation task, it was decided to use the k-fold cross validation method. Of the 100 images, 10 have been selected randomly to carry out the testing phase. The remaining 90 images are used for training and validation phase.

Initially 90 images are divided into 5 equal sets (folds), 4 of these are used for the training phase and the remaining one is used for the validation one. During 5 times the training is carried out, in each one the validation set will be different. At the end of each

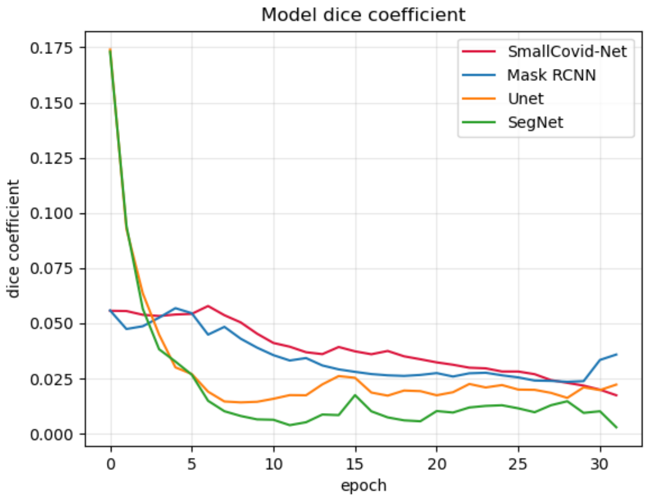

Fig.7. Validation loss.

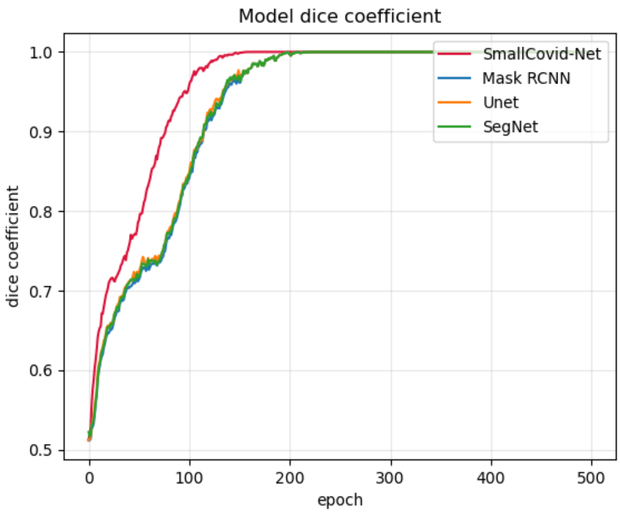

Fig. 8. Dice coefficient during the training phase.

training, the loss measure is calculated and with the average the model performance evaluation is obtained.

In these experiments we perform the segmentation task, but only the semantic one. All models are focused to identify where the injure is. In the Figures 6 and 7 we can see that all models have learned the way to identify the regions marked as COVID-19.

To show the results, the Table 1 has been created where the performance metrics of the models can be displayed. In order to evaluate the segmentation task, the Says coefficient has been used, which allows comparing the pixel segmentation generated

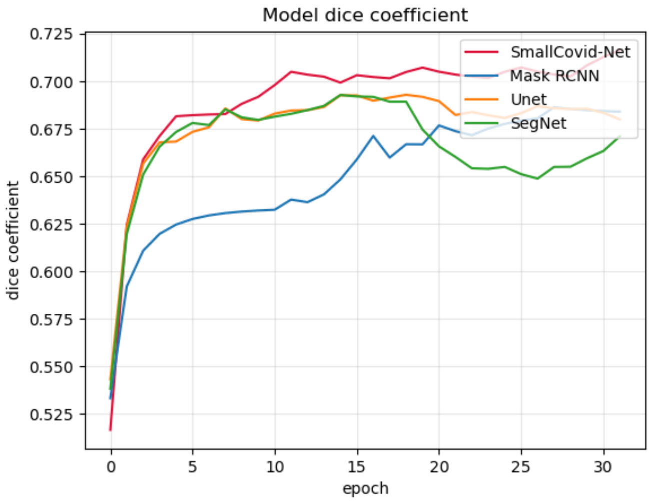

Fig. 9. Dice coefficient during the validation phase.

by the model against the ground truth. The Unet model obtained a better result for the dice and recall metrics, however our proposed model obtained the highest precision metric.

All models used were able to perform COVID-19 lesion classification from the image background, although they were unable to do instance segmentation because they could not determine the type of COVID-19 lesion they had segmented.

In Figures 8 and 9 we can identify that the best scores of the coefficient dice were obtained during the training phase.

5 Conclusions

In this article, we have proposed the SmallCovid-Net model for the segmentation of areas infected with the COVID-19 virus using computed tomography images.

We have made a comparison of our model with other models that perform image segmentation. As has been observed in this work, the segmentation of lesions caused by viruses is a very difficult task because there is no defined shape of lesions, and they can be located anywhere in the lungs. Due to the above, the results using the dice coefficient have yielded very low results, around 70%.

Considering that a perfect segmentation would yield a dice coefficient value of 100%, the model can be a support for medical personnel who want to detect lung injuries.

In the future, we want to apply this model to detect other types of lesions in other organs such as the brain. We also want to improve our model so that it can perform instance segmentation and not just semantic.

Acknowledgments. No funding was received to assist with the preparation of this manuscript.

References

1. Medical Segmentation.: COVID-19 CT segmentation dataset (2020) http://medicalsegmentation.com/covid19/
2. Abdulla, W.: Mask R-CNN for Object Detection and Instance Segmentation on Keras and Tensorflow. https://github.com/matterport/Mask_RCNN (2017)
3. Badrinarayanan, V., Kendall, A., Cipolla, R.: Segnet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 39(12), pp. 2481–2495 (2017) doi: 10.1109/TPAMI.2016.2644615.
4. Chan, J.F.W., Yuan, S., Kok, K.H.: A Familial Cluster of Pneumonia Associated with the 2019 Novel Coronavirus Indicating Person-to-Person Transmission: A Study of a Family Cluster. The Lancet, 395(10223), pp. 514–523 (2020) doi: 10.1016/S0140-6736(20) 30154-9.
5. Chowdhury, M.E.H., Rahman, T., Khandakar, A.: Can AI Help in Screening Viral and COVID-19 Pneumonia?. IEEE Access 8, pp. 132665–132676 (2020) doi: 10.1109/ ACCESS.2020.3010287.
6. Gatys, L.A., Ecker, A.S., Bethge, M.: Texture and Art with Deep Neural Networks. Current Opinion in Neurobiology, 46, pp. 178−186 (2017) doi: 10.1016/j.conb.2017.08.019.
7. Hofmanninger, J., Prayer, F., Pan, J.: Automatic Lung Segmentation in Routine Imaging is Primarily a Data Diversity Problem, not a Methodology Problem. European Radiology Experimental, 4(1) (2020) doi: 10.1186/s41747-020-00173-2.
8. Jiang, Z.P., Liu, Y.Y., Shao, Z.E.: An Improved VGG16 Model for Pneumonia Image Classification. Applied Sciences, 11(23) (2021) doi: 10.3390/app112311185.
9. LeCun, Y., Bengio, Y., Hinton, G.: Deep Learning. Nature, 521(7553), pp. 436 (2015) doi: 10.1038/nature14539.
10. Litjens, G., Kooi, T., Bejnordi, B.E.: A Survey on Deep Learning in Medical Image Analysis. Medical Image Analysis, 42, pp. 60–88 (2017) doi: 10.1016/j.media.2017.07.005.
11. Liu, Z., Jin, C., Wu, C.C.: Association Between Initial Chest CT or Clinical Features and Clinical Course in Patients with Coronavirus Disease 2019 Pneumonia. Korean J Radiology, 21(6), pp. 736−745 (2020) doi: 10.3348/kjr.2020.0171.
12. Minaee, S., Boykov, Y., Porikli, F.: Image Segmentation Using Deep Learning: A Survey. IEEE Transactions on Pattern Analysis and Machine Intelligence, 44(7), pp. 3523–3542 (2022) doi: 10.1109/TPAMI.2021.3059968.
13. Orioli, L., Hermans, M.P., Thissen, J.P.: COVID-19 in Diabetic Patients: Related Risks and Specifics of Management. Annales d'Endocrinologie, 81(2), pp. 101–109 (2020) doi: 10.1016/j.ando.2020.05.001.
14. Ren, S., He, K., Girshick, R.: Faster R-CNN: Towards Real-Time Object Detection with Region Prop Osal Networks. IEEE Transactions on Pattern Analysis and Machine Intelligence, 39(6), pp. 1137–1149 (2017) doi: 10.1109/TPAMI.2016.2577031.
15. Ronneberger, O., Fischer, P., Brox, T.: U-Net: Convolutional Networks for Biomedical Image Segmentation. In: Navab, N., Hornegger, J., Wells, W.M.: Medical Image Computing and Computer-Assisted Intervention. MICCAI 2015. pp. 234–241. Springer International Publishing (2015)

16. Salazar-Urbina, A., Ventura-Molina, E.J., Yáñez-Márquez, C.: Minicovid-unet: CT-Scan Lung Images Segmentation for COVID-19 Identification. Computación y Sistemas, 28(1), pp. 75–84 (2024) doi: 10.13053/CyS-28-1-4697.
17. Shi, H., Han, X., Jiang, N.: Radiological Findings from 81 Patients with COVID-19 Pneumonia in Wuhan, China: A Descriptive Study. The Lancet Infectious Diseases, 20(4), pp. 425–434 (2020) doi: 10.1016/S1473-3099(20)30086-4.
18. Wang, L., Lin, Z.Q., Wong, A.: COVID-net: A Tailored Deep Convolutional Neural Network Design for Detection of COVID-19 Cases from Chest X-Ray Images. Scientific Reports, 10(1), pp. 19549 (2020) doi: 10.1038/s41598-020-76550-z.
19. Ye, Z., Zhang, Y., Wang, Y.: Chest CT Manifestations of New Coronavirus Disease 2019 (COVID-19): A Pictorial Review. European Radiology, 30(8), pp. 4381–4389 (2020) doi: 10.1007/s00330-020-06801-0.
20. Zhang, Y., Yukun, Y., Wang, D.: Clinical Application of Image Processing and Neural Network in Cytopathological Diagnosis of Lung Cancer. Chinese Journal of Thoracic and Cardiovascular Surgery, 6(04) (2003)
21. Zhang, Z., Ni, X., Huo, G.: Novel Coronavirus Pneumonia Detection and Segmentation Based on the Deep-Learning Method. Annals of Translational Medicine, 9(11) (2021) doi: 10.21037/atm-21-1156.