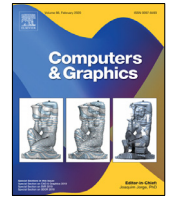




Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



Technical Section

COVIR: A virtual rendering of a novel NN architecture O-Net for COVID-19 Ct-scan automatic lung lesions segmentation[☆]



Kahina Amara^{a,*}, Ali Aouf^b, Hocine Kennouche^b, A. Oualid Djekoune^a, Nadia Zenati^a, Oussama Kerdjidi^a, Farid Ferguene^b

^a CDTA Centre for Development of Advanced Technologies, City 20 August 1956 Baba Hassen, Algiers, Algeria

^b USTHB University of science and technology Houari Boumediene, B.P 32 El Alia 16111 Bab Ezzouar, Algiers, Algeria

ARTICLE INFO

Article history:

Received 20 November 2021

Received in revised form 9 March 2022

Accepted 9 March 2022

Available online 15 March 2022

Keywords:

Automatic segmentation

COVID-19

O-Net

Deep learning architecture

3D reconstruction

Virtual reality

ABSTRACT

With the Coronavirus disease 2019 (COVID-19) spread, causing a world pandemic, and recently, the virus new variants continue to appear, making the situation more challenging and threatening, the visual assessment and quantification by expert radiologists have become costly and error-prone. Hence, there is a need to propose a model to predict the COVID-19 cases at the earliest possible to control the disease spread. In order to assist the medical professionals and reduce workload and the time the COVID-19 diagnosis cycle takes, this paper proposes a novel neural network architecture termed as O-Net to automatically segment chest Computerised Tomography Ct-scans infected by COVID-19 with optimised computing power and memory occupation. The O-Net consists of two convolutional autoencoders with an upsampling channel and a downsampling channel. Experimental tests show our proposal's effectiveness and potential, with a dice score of 0.86, pixel accuracy, precision, specificity of 0.99, 0.99, 0.98, respectively. Performance on the external dataset illustrates generalisation and scalability capabilities of the O-Net model to Ct-scan obtained from different scanners with different sizes. The second objective of this work is to introduce our virtual reality platform, COVIR, that visualises and manipulates 3D reconstructed lungs and segmented infected lesions caused by COVID-19. COVIR platform acts as a reading and visualisation support for medical practitioners to diagnose COVID-19 lung infection. The COVIR platform could be used for medical education professional practice and training. It was tested by Thirteen participants (medical staff, researchers, and collaborators), they conclude that the 3D VR visualisation of segmented Ct-Scan provides an aid diagnosis tool for better interpretation.

© 2022 Elsevier Ltd. All rights reserved.

1. Introduction

COVID-19 is a highly spreading contagious respiratory infection that has had a devastating impact on the world. Recently, new variants of COVID-19 continue to appear, making the situation more challenging and threatening [1]. The COVID-19 symptoms may vary from common like (fever, chill, dry, cough, tiredness, fatigue) to less common ones like (sore throat, aches and pains, conjunctivitis, diarrhoea, headache, a rash on the skin, or discolouration of fingers or toes, loss of taste or smell), or severe like (difficulty breathing or shortness of breath/chest pain or pressure/loss of speech or movement). The estimated incubation

period is between 2 and 14 days and has a median of 5 days. It is worthy to mention that some people become infected and do not develop any symptoms or feel unwell [2].

Other than the symptoms, there are ways to detect the disease. COVID-19 tests are available that can detect current infection or past infection:

- A viral test shows current infection. Two types of viral tests exist antigen tests that use polymerase chain reaction (PCR) [2], and nucleic acid amplification tests (NAATs).
- An antibody test (also known as a serology test) shows past infection. For current infection diagnosis, antibody tests should not be used.

PCR tests are widely used and considered the most accurate form available today; it takes time before the results are known CT scan images are preferred by an experienced doctor to detect the infected lesions in the lungs. CT scan benefits comprise less cost, provide valuable data, and wide availability [12]. Assessment and quantification of COVID-19 lung abnormalities based

[☆] This article was recommended for publication by B. Kozlikova.

* Corresponding author.

E-mail addresses: kamara@cdta.dz (K. Amara), ali.aouf1@etu.usthb.dz

(A. Aouf), hocine.kennouche@etu.usthb.dz (H. Kennouche), odjekoune@cdta.dz

(A.O. Djekoune), nzenati@cdta.dz (N. Zenati), okerdjidi@cdta.dz (O. Kerdjidi),

fferguene@yahoo.fr (F. Ferguene).

Table 1
Representative deep learning-based methods for Ct-scan COVID-19 diagnosis.

Reference	Methodology	Performance metrics	Task
Keno et al. [3]	Used transfer learning to train 3D-U-Net model using an 18-layer 3D ResNet	Dice 0.679	Segmentation
Adnan et al. [4]	Used SegNet and U-Net deep learning networks	PixAcc. 0.908 U-Net 0.907 SegNet	Segmentation
Chen et al. [5]	Used UNet++ to train in Keras in an image-to-image manner	PixAcc. 0.95	Detection
Li et al. [6]	Introduced COVNet	PixAcc. 0.96	Detection
Ying et al. [7]	Proposed DRE-Net	PixAcc. 0.86	Detection
Garain et al. [8]	Proposed SNNs	PixAcc. 0.91	Detection
Stefano et al. [9]	Proposed C-ENET	Dice. ~0.75	Segmentation
Zhang et al. [10]	Improved DCNN model	PixAcc. 0.93	Detection
Shouliang et al. [11]	Proposed DR-MIL method	PixAcc. 0.95	Detection

on chest Ct-scans can help determine disease stage, effectively allocate limited medical resources, and make informed treatment decisions. During the pandemic, expert radiologists' visual assessment and quantification of COVID-19 lung lesions have become costly and error-prone, urgently requiring the development of stand-alone practical solutions.

As a long-standing topic, accurate boundary segmentation is still a challenging and critical process in medical imaging. It is a crucial step in clinical treatment. It dates back to the first days of medical imagery when radiologists segmented objects in images manually. This process benefited primarily from technological advancements, which made it automatic and accelerated its execution time. The first generations of segmentation methods were based on pure mathematical concepts. The next generations used algorithms that can learn and adapt, which made them more accurate and precise [13].

Ct-scan images are regarded one of the most useful sensing approaches since it allows the physicians and radiologists to identify internal structures and see their shape, size, density and texture. Some examples of Ct-scan segmentation and classification methods in COVID-19 applications based deep learning approaches are summarised in Table 1 along with performance obtained and the used neural network model of each approach. The mathematical formula of Dice, pixel accuracy in addition to other metrics can be found in Section 3.4.1. The literature states that most of the techniques used pre-trained models without changes. Notably, the research done on U-Net [14] achieved positive accuracy on Ct scans [3–5]. In addition to that, it was noticed that UNet performs better when compared to deep learning other models, which are more complex and have a large number of parameters that need to be tuned. This motivated the use of UNet as it is less complex. More examples of deep learning-based segmentation methods are given in the next section. Ct-scans used in diagnosing COVID-19 may produce false-negative effects, especially early infection; it is one of its significant weaknesses [12]. Virtual reality visualisation can remedy false-negative detection at early-stage infection [12]. The user can interact and become immersed in a computer-generated environment in a realistic way via virtual reality technology. The key concepts that define VR are immersion, sense of presence, and the possibility to interact with the computer-generated environment [15]. Nowadays, medical imagery segmentation is the base where radiologists make their noted observations, and doctors provide their diagnoses. Moreover, this process is always under enhancement and improvement to make the error margin near zero. Researchers are developing applications for epidemic illnesses and diseases. Recently, virtual and augmented realities applications have shown great potential in the medical field and healthcare [16,17]. VR technology designs platforms to reduce the face to face interchange of doctors with the infected COVID-19 patients. Today VR systems overcome classical medical imagery problems with novel 3D imagery visualisation techniques. In the ongoing pandemic COVID-19, it has been shown that VR-developed techniques help healthcare-related applications [18]. Medical VR applications are

frequently used as surgery simulation, student teaching, doctors training, or generally as observation and analysis tools. With VR technology, doctors can visualise the data better with more details and less work fatigue. Ct-scans used in diagnosing COVID-19 may produce false-negative effects, especially early infection; it is one of its significant weaknesses [12]. Virtual reality visualisation can remedy false-negative detection at early-stage infection [12]. The user can interact and become immersed in a computer-generated environment in a realistic way via virtual reality technology. The key concepts that define VR are immersion, sense of presence, and the possibility to interact with the computer-generated environment [15]. Fig. 1 illustrates the main work parts. The main contributions of this paper are as follows:

- Firstly, we develop a novel neural network architecture O-Net to automatically segment lung lesions of chest Ct-scans infected by COVID-19 with optimised computing power and memory occupation;
- Secondly, the proposed O-Net improved architecture is based on U-Net, inspired by the Ki-U-Net, [19], double convolutional channels. It newly consists of two convolutional autoencoders with an upsampling channel and a downsampling channel. The O-Net enhanced the U-Net performance based on experimental results;
- Finally, a Virtual Reality platform, COVIR, is designed and deployed. It allows the 3D lungs and infected lung regions visualisation and 3D interaction. The fundamental usefulness of this system are (a) It allows medical practitioners (radiologists, medical students) to visualise the 2D medical data onto 3D models, interact, manipulate them, study and navigate the inside 3D reconstructed segmented Ct-scan data. (b) It offers a realistic view with stereoscopic depth perception, which give better insights and comprehension into medical imaging. It has the capacity for real-time interactivity and accurately visualises dynamic 3D volumetric data. (c) COVIR platform supplies the COVID-19 diagnosis by identifying and interpreting the lung damages caused by COVID-19. It reduces the face to face interaction of doctors with the infected COVID-19 patients. The developed system could be used for medical education professional training and as a tele-health VR platform. The COVIR platform was tested by a set of 13 participants composed of medical staff, computer science students, researchers, and collaborators. They volunteered to investigate and provide their opinions through user experience.

The remaining of this proposal is organised as follows: Section 2 is the related work about existing classical and deep learning Ct-scan segmentation based methods for COVID-19 lung lesions segmentation. A detailed description of the proposed architecture O-Net is given in Section 3; we introduce experiment settings, evaluation methods, dataset, and results. Section 4 presents the virtual reality platform for 3D lungs and lesions visualisation and manipulation. Section 5 closes this paper by highlighting the outcome achieved with a conclusion and future work.

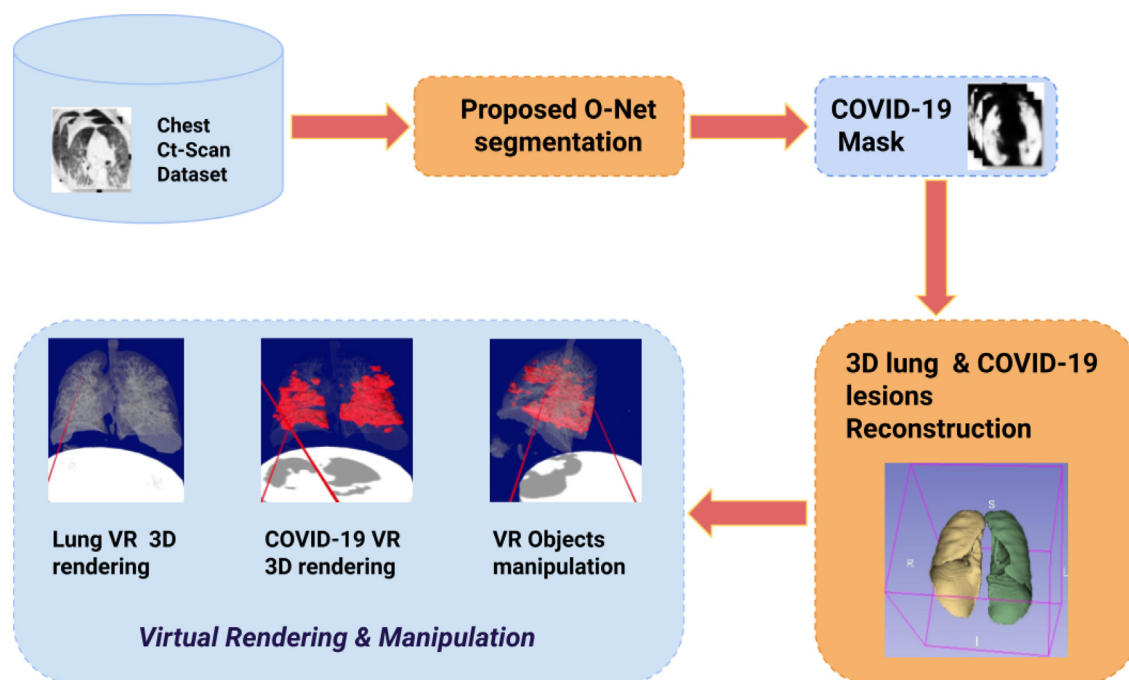


Fig. 1. The workflow of the proposed COVID-19: Virtual rendering of the proposed NN COVID-19 lung lesions segmentation.

2. Related works

An exhaustive visual diagnostic process includes critical components such as automatic detection and segmentation of abnormalities, qualitative and quantitative analysis, and interactive visualisation tools. Here, we restrict our discussion to work related to lungs segmentation, virtual 3D visualisation and COVID-19 diagnosis.

2.1. Lung and COVID-19 lesion segmentation

Image segmentation is the process of extracting one or a set of objects from a given image based on their class. The main idea behind this process is to generate an image where the pixels within the selected objects and the pixels that are not selected have different values. In this section, we will address both the classical and neural network-based segmentation methods. Furthermore, we will discuss the advantages and inconveniences of each category. Region-based segmentation methods are considered as the classical method; they consist of basic image processing operations. The most used procedure is thresholding; it sets one or two thresholds acting as boundaries for the pixel values. The process then proceeds into neutralising pixels that have values outside of the interval formed by the two thresholds [20]. Another well-known method is Regional-Growth segmentation [20]. This algorithm starts by selecting a set of pixels that will be considered as seeds. The process of adding neighbouring pixels into the already formed regions is called an iteration. The algorithm Region growing iterates until it converges, i.e. until all seeded regions have stopped growing [20]. We can also mention the Edge-Detection segmentation methods. The logic of these methods lies in the fact that there is an edge between every two neighbouring segments. The Edge detection segmentation method uses a convolution operation on a given image. The resulted image shows a high contrast difference between the edges and the background [20]. The type of filter is the main factor in this operation. The most used are the Sobel and Laplacian operators [20,21].

The Region-based segmentation methods make a well-suited introduction to the field of image segmentation. Morphological filters are based on the concept of linear filters, which are simple to implement and develop. Morphological filters are also less computationally complex, and their execution process does not occupy a large size in memory. Due to their simplicity, morphological filters have many weaknesses; they are hard to adapt to different imagery categories. On account of their limitations, morphological-filter based methods have poor usability in real-world medical imagery applications.

Clustering is an unsupervised learning technique that assigns labels to data based on measures of similarity or dissimilarity. It is used in medical image segmentation. While conventional clustering methods seek to group a set of items into clusters, they can be tuned to group a list of pixels into segments [22]. Its simplicity, dynamicity and the use of a Superpixel can accelerate the process and lower memory usage. Superpixel techniques segment an image into regions by considering similarity measures defined using perceptual features [23]. The drawback of using Superpixel is the loss of critical information, and details like pixel position. The outputted result is vague due to many objects grouped in one segment. The application of this approach stays limited.

To overcome the drawbacks mentioned earlier, we can either use a combination of these methods or switch to a more recent method. Recently, deep learning-based segmentation models in various tasks for automatic medical image segmentation have shown outstanding performance. Artificial neural networks got introduced in 1943. Since then, convolutional neural networks (CNNs) have evolved through many publications introducing new architectures and enhancing existing ones. The high accuracy and powerful generalisation allowed neural networks NN methods to become state of the art in many fields, especially for medical image segmentation.

In 2015, Ronneberger et al. in [14] developed a convolutional network architecture named U-Net for biomedical image segmentation. The U-Net architecture contains two parts; the encoder, which is used to extract the context of the image. The second part is a decoder used to enable precise localisation using transposed convolutions [24]. The decoder generates the mask based on the

contextual information extracted by the encoder. The encoder is a traditional stack of convolutional and max-pooling layers. The encoder and decoder are presented in a set of double convolutional layers, also known as stages; after each stage, a sampling method is added. U-Net has two limitations: its optimal depth is unknown ahead, requiring extensive architecture search or an inefficient ensemble of models of varying depths.

After the launch of U-Net architecture, several works were proposed based on its architecture. Ki-U-Net is a method that combines the U-Net architecture with the Ki-Net architecture. The Ki-U-Net extends the U-Net architecture with an upsampling convolutional channel, it resulted in significant improvements regarding segmenting tiny segments while staying resilient to noise. Furthermore, this architecture introduces a high number of inter-connections between its channels leading to a drastic increase in the number of calculation operations during the back-propagation step therefore sacrificing time for more accurate segmentation. But it sacrifices the time and size of the work to get better results [19]. Ki-U-Net has an improved efficiency caused by the high amount of interconnections between layers. But, it requires heavy computation and significant space, which makes it unsuitable for light machines.

SegNet, [25], is a semantic segmentation model submitted by the University of Cambridge in 2017, this core trainable segmentation architecture consists of an encoder-decoder network, followed by a pixel-wise classification layer [25]. The encoder is formed based on the VGG-16 architecture, which contains, 13 convolutional layers and a 2 \times 2 max pooling layer between each two convolutional layers [26]. While the decoder is composed of a set of transposed-convolutional layers which perform the upsampling operations. In the end, to predict the class for each pixel, there is a K-class softmax classifier used [25]. SegNet is an efficient segmenting approach. The model size is much smaller than other approaches, but even with the decent result, it still cannot compete with the recent approaches, including U-Net models and its variants, for medical image segmentation. It has a low memory requirement during both training and testing.

Transformers-based U-Net is considered the alternative architecture to U-Nets. Trans-U-Net uses attention mechanisms that rely heavily on convolutional layers [27]. Transformers designed for sequence-to-sequence prediction shows exceptional performance in various machine learning tasks. Combined with U-Net to enhance details by recovering localised spatial information, recently proposed, TransU-Net gave promising results [27]. It achieves more accurate results than a variety of similar architectures, including CNN-based self-paying methods. Yet, using transformers to improve the results added complexity, which requires much more data to tune the whole model.

As previously shown, architectures like SegNet and U-Net (with their variations) are considered state of the art medical images segmentation. Once healthcare centres published the first segmented images of COVID-19 [28], multiple researchers trained these previously mentioned architectures to segment COVID-19 medical scans. One of the first papers that performed segmentation on COVID-19 Ct-scan is a work published by Deng-Ping Fan et al. [29] on 22 April 2020. This work proposed an architecture based on SegNet, with the only difference being the use of attention mechanisms instead of convolutional layers. The authors named this architecture “Inf-Net” as an abbreviation of Infection-network. The model they used had 33.122 million parameters. After training it on the Italian dataset [28], and comparing their results with results obtained from U-Net based models, they deduced that Inf-Net models succeeded in detecting most infection segments while U-Net based models failed. However, they also showed that In-Net models are detecting unexciting segments. Another work by Keno K. Bressemer et al. [3] focused on 3D Ct-scan

segmentation. The authors developed a 3D-U-Net; their proposed model consisted of using a 3D convolutional layer instead of 2D layers. They increased the number of blocks to 5 and half the number of filters. They used a pre-trained 18-layer 3D ResNet encoder. The authors trained their model three times, each time on a different dataset. The datasets they used are the Chinese Coronacases dataset [30], the Russian MosMed dataset [28], and the RICORD dataset [31]. The authors of [3] defended the use of 3D based model rather than 2D by stating that 2D slices may introduce “selection bias into the data by excluding slices that do not show lung or infiltrate area”. The researchers also argued that 3D models preserve spatial information and allow the model to see the entirety of the lung rather than just a slice. However, training a 3D model comes with multiple obstacles, including the large memory consumption, the need for a large dataset, and the long execution time. The paper showed these obstacles when the authors used an Nvidia Quadro P6000 GPU with more than 22 Gigabytes of VRAM (“video random access memory”) to train the model for more than 130 h. In [4], Adnan Saood et al. made a performance comparison between U-Net and SegNet on the task of segmenting COVID-19 Ct-scan. Both models were trained on the Chinese Coronacases dataset [30], the authors also affirmed that by defining the segmentation as binary-class segmentation (infected or not infected) rather than a multi-class segmentation (background, infected areas and perhaps lung tissue), the pixel accuracy marginally improved by 0.05% [4].

2.1.1. Discussion

The radiological imaging may help support early screening of COVID-19. While being recently introduced, NN methods showed great potential at extracting and learning patterns from an inputted sample. This feature allowed NN models to succeed in complex tasks like detection, classification, and image segmentation. Nevertheless, NN models are subject to multiple drawbacks, including the necessity of a large dataset, long training time, high computation power consumption, and the potential collapse due to data bias selection, poor data flow, or wrong sequencing of layers. Building a reliable NN model is still a challenging task requiring familiarity with multiple concepts like overfitting, underfitting, tuning, and the overall role of functional layers, activation, and loss functions.

Based on the results of the studies mentioned above dealing with the COVID-19 Ct-scan segmentation and our experimentation, we noted that the U-Net based models perform better than SegNet based models while having a relatively equal number of parameters. Attention-based models like Inf-Net and U-Net variations have tens of millions of parameters that help the model encode more information. The drawbacks of such models are the long time to converge and the necessity of a large dataset to help tune their large number of parameters. Some researchers developed architectures that could provide more accurate segmentation, with the most notable one being Ki-U-Net. This architecture uses two convolutional channels. The authors stated that Ki-U-Net is more accurate and can detect fine details and shapes of tumours and abnormalities better than the U-Net architecture. Yet, the cost of developing and deploying this architecture is tremendously higher than the vanilla U-Net. Ki-U-Net models are relatively small, with around one million parameters and an enormous forward/backwards pass size that grows exponentially relative to the input resolution. The main drawbacks of current NN methods are scalability, memory consumption, and long convergence time. Our work is about developing a new architecture based on U-Net, inspired by the Ki-U-Net double convolutional channels, yet lighter and with less data and computing-power consuming. Inspired by U-Net appellation [14], the term O-Net; Net refers to network; is due to the architecture

of upsampling and downsampling. The term O-Net was previously used in [32]. The proposed work in this paper consists of two convolutional channels that coordinate to segment the same class of objects. However, in the cited paper, [32], each convolutional channel is working on segmenting a different class (blood vessels, blood oxygenation). Another key difference is the fact that in their model, the channels are homogeneous, while in ours, one channel does the downsampling, and the other does the upsampling. To conclude, the only mutual thing between the two works is the, O-Net, name given to the networks.

As shown in Fig. 2, the architecture is in the form of the letter 'O'.

2.2. VR systems during COVID-19 pandemic

VR could be viewed as an practical solution for 3D visualisation of medical images since it could provide efficient disease analysis and diagnosis regarding the classic approaches. VR gives an chance to immerse users in a fully artificial digital medical environment that involves the human anatomy described in 3D models. The authors of [33] suggested that virtual reality when simulating the clinical environment, can overcome the significant disruption to in-hospital medical training and can be particularly useful to supplement the traditional in-hospital medical training during the COVID-19 pandemic. Their user study evaluation results affirmed the positive effect of virtual reality training realistic for the initial clinical assessment. The medical students were offered to complete their practical training online including the access to a virtual reality platform with a variety of clinical case-based scenarios of different types and complexity. [18] presented a study on Virtual Reality and its applications for the COVID-19 pandemic. They concluded that VR technology develops platforms to reduce the face to face interaction of doctors with the infected COVID-19 patients. VR has a number of advantages over traditional rehabilitation approaches for Cognitive Rehabilitation during the COVID-19 pandemic [15] by achieving adequate cognitive stimulation in the era of social distancing related to COVID-19 pandemic. Few works addressed the 3D virtual visualisation of segmented COVID-19 lung lesions to the best of our knowledge. Recently, [12] has addressed the virtual reality visualisation for computerised COVID-19 lesion segmentation. They combined CT imaging tools and VR technology to generate a system for accurately screening COVID-19 disease and navigating 3D visualisations of medical scenes to visualise dynamic 3D volumetric data. They are using a threshold based method. A DICOM imagery stack was converted to OBJ file and/or slt formats. In their configuration, they used Blender software framework to import OBJ files and generated FBX format directly used to provide 3 lung visualisations. In [34], an automatic lung segmentation using a deep learning model is proposed. The authors presented a 2D and 3D visualisation application specially tailored for radiologists to diagnose COVID-19 from chest CT data. They implemented their visualisation tool; COVID-view; ground-up using VTK and Qt with a simplified and essential interface integrating lungs and lesions automatic segmentation, and COVID-19 classification.

3. O-Net: new architecture for COVID-19 Ct-scan lung lesions segmentation

3.1. O-Net description

This paper proposes a novel architecture, termed as O-Net, that uses U-Net as a base model and introduces several improvements and changes that attempt to render the model more accurate with less memory consumption than Ki-U-Net. Our proposed architecture consists of two convolutional auto-encoders.

Similar to Ki-U-Net, it has an upsampling channel and a downsampling channel. Each channel is composed of building blocks called stages. A stage is a sequence of two convolution layers, where each convolutional layer is followed by a batch normalisation layer and a ReLU activation layer. Our proposal uses a decreasing number of filters. In our proposal, we used 64 filters in the first stage Fig. 2, followed by 32 in the second stage. Unlike Ki-U-Net, our architecture uses two bottlenecks and no interconnections between the channels. These changes allowed us to have a model lighter than Ki-U-Net, with more parameters, a faster convergence rate, and less computation and memory footprint based on the results obtained. According to the authors of Ki-U-Net, U-Net showed a considerable performance drop in the case of detecting smaller anatomical landmarks with blurred, noisy boundaries. Their solution includes adding a new convolutional channel that projects the data into higher dimensions. However, the authors of the Ki-U-Net introduced skip connections between each level of the channels. This change increased the convergence time and memory finger of the model drastically. In our work, we kept the channel that projects the data into a higher dimension since it had a very interesting concept. However, we dropped the skip connections between the channels since it was not cost-effective. We tested both increasing and decreasing numbers, and based on our experiments, the network with the decreasing number of kernels performed better. For the kernel size, we used the standard. The previous sections show that introducing a second convolutional channel as an upsampling channel increases the model accuracy. However, this may increase the memory footprint size. Moreover, the high number of interconnections between the convolutional channels caused a large memory consumption and slowed the training time ten times.

3.2. Used COVID-19 Ct-scan datasets, pre-treatment and data augmentation

Throughout our experiments, we used the first and second versions of the Italian dataset [28] to learn the fundamentals of segmentation and test basic U-Net models. This dataset, however, was shown less precise segmentation that led to low performance. Eventually, we had to opt for other datasets. The accessible alternatives were the Chinese Coronocases datasets [30] and the Russian MedMod dataset [35]. We extracted ten scans from the Coronocases datasets [30], since the other ten were noisy and did not match the desired shape.

Moreover, we could only use fifty scans from the MedMod dataset [35] considering they were the only segmented scans. The total number of scans was sixty, with each scan having a number of slices between 63 and 300. The total number of slices was approximately 4100, with only 2000 slices that contained infected regions. Before starting any deep learning process, the collected data had to be adjusted and formatted to fit the inference requirements. First and foremost, we resized the slices from both datasets to 512×512 , 256×256 , and 160×160 resolutions. The role of this step is to give us the flexibility of testing the performance of architectures on images with small resolutions (160×160). Then, we increase the resolution gradually as we proceed in the experiments. Furthermore, we opted not to use de-noising processes as the Coronocases dataset provided clean slices. In contrast, the scans from the MedMod dataset had a noisy background, probably formed by the capturing machine. The MedMode dataset's noise was mainly harmless since it did not interfere with the pixel-value intervals of infected tissue, and it was surrounding the lung rather than being inside it.

Regarding the augmentation techniques, we mainly used methods that do not affect or alter the shape of the slices, which

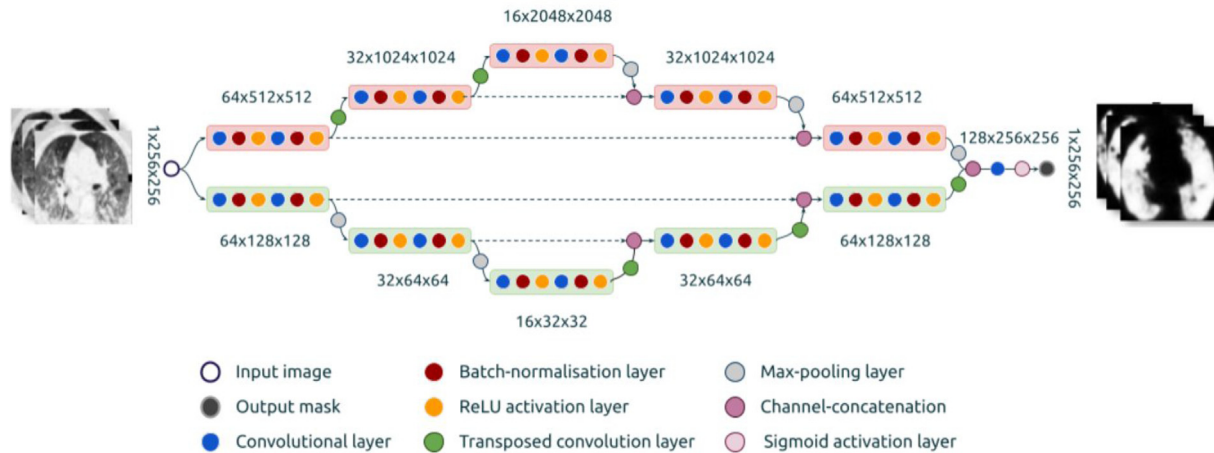


Fig. 2. The O-Net architecture, where the upsampling channel is on the top and the downsampling channel is in the bottom.

can be translated to transformations that do not include elastic effects like shearing. Affine transformations were the primary techniques we used. Through experimentation, we noticed that rotation transformation help NN models adapt to lung position and rotation. Moreover, we saw that combining horizontal or vertical flips with angular rotation help reduce the probability of performing the same transformations on the same image multiple times. Cropping a border of up to 25 pixels off the images is harmless as long as the lungs, bones, and other tissues are fully included. Furthermore, we have tested the effect of normalising the pixel values of images using U-Net and O-Net models, and we found that normalisation does not improve the performance. Models that were trained on normalised data took more epochs to converge and were less stable. On that note, blurring the input produced less accurate models, and adding white noise resulted in a total collapse of the models. Regardless of the used dataset or the number of training epochs, the outputted masks were black images.

3.3. Training cycles and training process

To train the models on the full-size dataset, we used a workstation that contains two Intel Xeon (R) Silver 4114 (20 cores @ 2.199 GHz) processor, 256 GB ram, and NVIDIA Quadro P2000 (5 GB VRAM) GPU, and a 2TB HDD storage. Although this machine had less VRAM, limiting the size of models that we can have, it came with a large ram and fast access to local storage, which meant that loading and treating all the slices became possible. Based on the work done by [3,29] on segmenting COVID-19 CT Scans, they both concluded that a two cycles training process is more effective than one cycle. Multi-cycle training refers to the process of training a model multiple times on different hyperparameters. Usually, the first cycles use a large batch size, a high learning rate, and a small slice resolution. The batch size and learning rate are then decreased while the slice resolution is increased gradually during the cycles. To find the optimal combination of hyperparameters, we trained a U-Net model on 2000 CT scan slices. We varied the batch size (8, 16, and 32) and the learning rate (10^{-5} , 10^{-4} , and 10^{-3}) during this experiment. Fig. 3 shows the results we obtained.

As shown in Fig. 3, a learning rate of 10^{-3} and 10^{-4} achieved the best performance while 10^{-5} resulted in relatively lower performance. On the other hand, varying the batch size did not affect the performance of the final models. One more thing to note is the execution time. A large batch size means a large number of iterations per epoch which increases the execution

time and leads to high memory consumption. Moreover, A small learning rate leads to a slow convergence time because the steps made in each iteration are small. Oppositely, a high learning rate renders the model unstable, which leads to unpredictable training results or a model collapse. Finally, a benefit of using a small batch size and a medium learning rate is reducing memory usage (fewer samples are loaded into the memory) and stabilising the training process. The data was split into 30% for testing and 70% for training.

3.4. Training, validation and test data setting

3.4.1. Evaluation metrics

For our models' performance assessment and comparison, we used seven well know evaluation metrics: Pixel-Accuracy "Pix-Acc.", Sørensen–Dice coefficient also known as F1-Score "Dice.", Intersection over Union, also known as Jaccard score "IoU.", Precision "Prec.", Sensitivity "Sens.", Specificity "Spec.", and "G-means". The Eqs. (1) to (7) describe these metrics. Where (tp) stands for True Positive; (fp) stands for False Positives; (fn) stands for False Negatives; (tn) stands for True Negatives.

$$PixAcc. = \frac{tp + tn}{tp + fp + tn + fn} \quad (1)$$

$$Dice. = \frac{2 \times tp}{2 \times tp + fp + fn} \quad (2)$$

$$IoU = \frac{tp + fn + fp}{tp} \quad (3)$$

$$Prec. = \frac{tp}{tp + fp} \quad (4)$$

$$Sens. = \frac{tp}{tp + fn} \quad (5)$$

$$Spec. = \frac{tn}{tn + fp} \quad (6)$$

$$G - mean. = \sqrt{sensitivity * specificity} \quad (7)$$

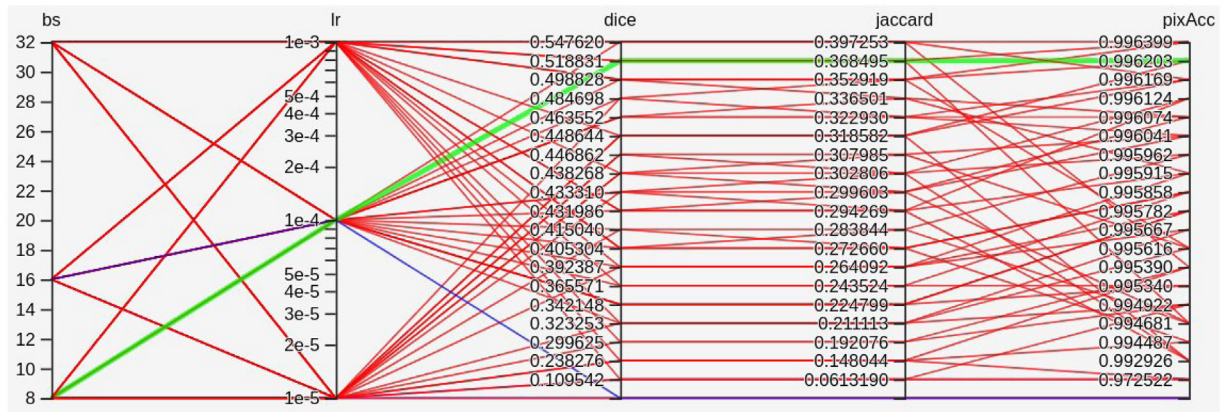


Fig. 3. A plot showing the performance of different hyperparameters combinations.



Fig. 4. O-Net Dice coefficient: validation vs. training.

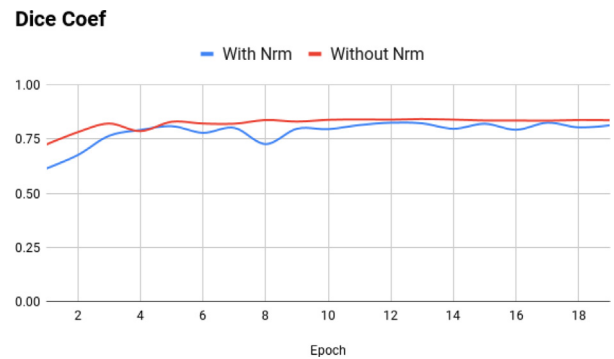


Fig. 5. The effect of normalisation on the convergence and adaptation, Nrm: normalisation.

3.4.2. Experiments, results and evaluation

We trained our models on the Coronacases dataset [30] for two cycles with each cycle having 20 epochs. Fig. 4 shows the Dice coefficient at each epoch of an O-Net model trained on 160×160 slices. The figure illustrates that the model takes approximately three epochs to reach the highest Dice score value (around 0.80). Moreover, we notice that both the validation and training plots are flat, suggesting that our model is not facing an Overfitting problem. Furthermore, we wanted to test the effect of normalising the inputted slices on the model convergence rate and the adaptation to unseen data. We applied two experiments, one with normalised data and the other with regular data. Fig. 4 serves the role of showing that the model is not overfitting. For that, we plotted the data from the first experiment. For Fig. 5, we showed the data from both experiments. The validation set was only used to test the model's performance and was not included in weights adjustment. The Fig. 5 shows a comparison between a model trained on normal data and another on normalised data.

Based on Fig. 5, we can confidently note that the normalisation did not help the model to converge faster or become more stable. Models trained on normalised data took four epochs to reach the highest Dice score, and throughout the training process, these models failed to outperform models trained on regular data. Fig. 6 shows a list of samples and infection masks predicted by O-Net. Figs. 6 and 7 illustrate the O-Net and U-Net visual assessment results respectively using Coronacases dataset [30].

For statistical evaluation, we calculated the COVID-19 infection rate (Eq. (8)): with N_{lung} : number of pixels segmented as lung, $N_{COVID-19}$: number of pixels segmented as COVID-19 lesions.

$$COVID - Rate. = \frac{N_{COVID-19}}{N_{lung} + N_{COVID-19}} \quad (8)$$

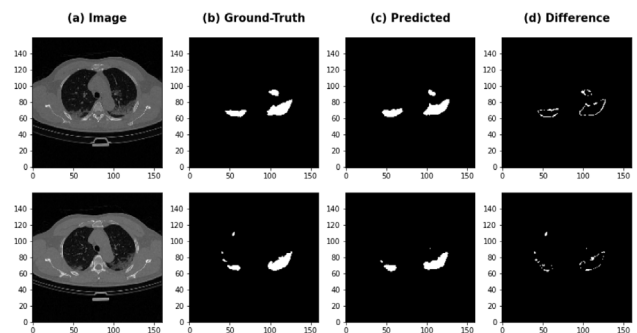


Fig. 6. COVID-19 infection masks predicted by O-Net using Coronacases dataset [30].

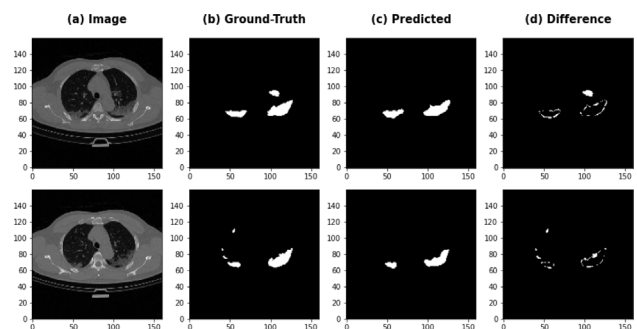


Fig. 7. COVID-19 infection masks predicted by U-Net using Coronacases dataset [30].

Table 2
Performance comparison between normalised and normal O-Net models.

O-Net	PixAcc.	Dice.	IoU.	Prec.	Sen.	Spec.	G-mean.
Normalised	0.996	0.818	0.695	0.998	0.991	0.722	0.845
Normal	0.995	0.866	0.764	0.998	0.984	0.708	0.834

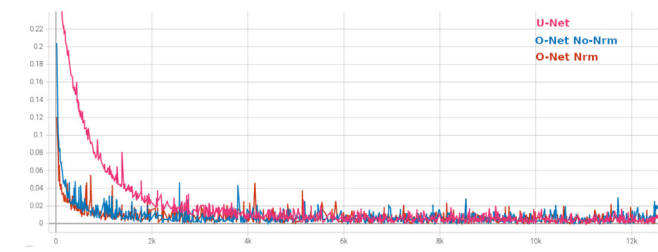


Fig. 8. Comparison between U-Net [14] and O-Net on convergence rate; Orange: O-Net with normalised data; Blue: O-Net with normal data; Pink: U-Net.

Table 3
Adaption comparison between U-Net [14] and O-Net.

Model	PixAcc.	Dice.	IoU.	Prec.	Sen.	Spec.	G-mean.
U-Net [14]	0.997	0.507	0.372	0.999	0.997	0.834	0.911
O-Net	0.997	0.584	0.428	0.999	0.997	0.817	0.901

Table 2 compares normalised models to normal models.

To visualise the enhancement in convergence slop, we registered the loss values of O-Net and compared them to loss values from U-Net training (Eq. (9)). We tested both models on normalised and non-normalised data. Fig. 8 illustrated the results. We used binary cross-entropy, as shown in Eq. (9). We calculate the entropy H of a q distribution, where y is a label, N is a number of points, p is the predicted probability.

$$H_p(q) = -\frac{1}{N} \sum_{i=1}^N y_i \cdot \log(p(y_i)) + (1 - y_i) \cdot \log(1 - p(y_i)) \quad (9)$$

Fig. 8 shows O-Net models converge faster due to having fewer parameters than U-Net models.

3.5. Generalisation: Assessment on an external dataset, and scalability.

One of the essential factors of a well developed NN model is adaptation. An adaptive NN model does not rely only on the dataset it was trained on but should generalise its understanding to similar datasets. To evaluate the adaptation of O-Net models, we took a model trained on the Coronacases dataset [30], and tested it on MosMed dataset [35]. We compared the values we obtained to a similar test performed on a U-Net model. Table 3 shows the results.

From Table 3, we can conclude that adaptation is still an intricate task. O-Net model witnessed a drop in performance. However, when we compare the O-Net performance to the U-Net performance on an unseen dataset, O-Net models adapt better to the new data. Figs. 9 and 10 present the visual evaluation for adaptation performance. As mentioned before in this work, we had access to three publicly available datasets. The publishers of these datasets had different kinds of scanning equipment which generated data with different ranges and values of noise. CT-Scans use a Hounsfield scale which has a wide range (−1000 to +20,000). Figs. 9-d and 10-d present the visual assessment of misclassified pixels of the unseen data. However, we wanted to enhance the U-Net model by giving it another channel to further extract and encode more data. While it might appear that both

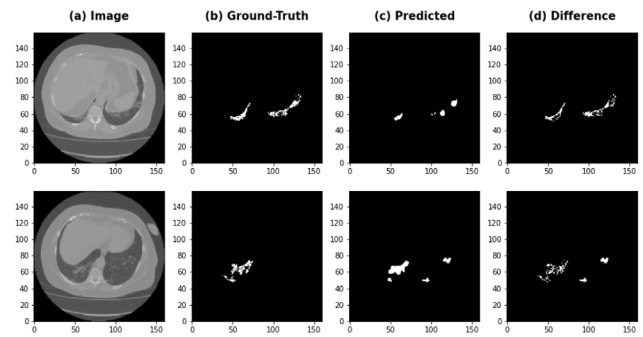


Fig. 9. COVID-19 infection masks predicted by O-Net using MosMed dataset [35].

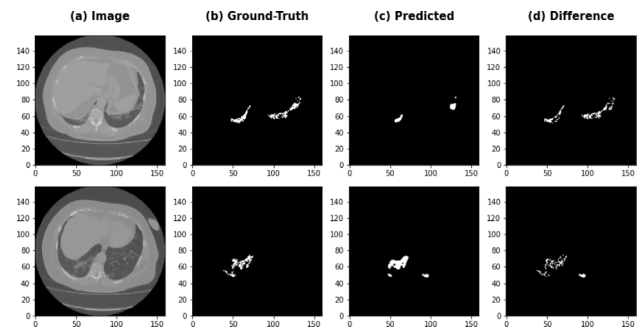


Fig. 10. COVID-19 infection masks predicted by U-Net using MosMed dataset [35].

Table 4
Scalability: Testing U-Net and O-Net on 256 × 256 slices.

Model	PixAcc.	Dice	IoU	Prec	Sen.	Spec.	G-mean
U-Net [14]	0.996	0.779	0.641	0.999	0.992	0.821	0.901
O-Net	0.997	0.844	0.731	0.998	0.991	0.547	0.734

models scored best at two metrics each, however, the G-means is calculated by multiplying the sensitivity by the specificity, which gave the U-Net a slightly higher score.

Our final test evolved around the ability to well-perform on higher-resolution slices. For this test, we used an O-Net model trained on 160 × 160 slices and tested it on 256 × 256 slices. Moreover, we performed the same test on a U-Net model, and the results are presented in Table 4.

Based on the Table 4, we can see that the O-Net model performed well when tested on higher resolution slices. This note means that an O-Net model can be trained on low-resolution data and properly perform when tested on high-resolution data. U-Net scored better than O-Net in three matrices. Two of which were with an increase of 0.001. The remaining one is specificity which measures the correctly negatively classified pixels. Moreover, O-Net scores better on dice coef and IoU (Jaccard). This remark means that O-Net is more concerned about finding the right shapes of the infected regions. This makes the model a bit conservative when detecting the outer layer or pixels within infected regions.

3.6. Performance comparison

Multiple characteristics must be taken into consideration to initiate a comparison between NN models. Furthermore, we collected the number of parameters and the pass-size of various models. Table 5 compares O-Net with U-Net [19] and Ki-U-Net [19]. As can be seen in Table 5, O-Net stands in the middle

Table 5
Comparison between O-Net, U-Net [14], and Ki-U-Net [19].

Model	Number of parameters	Pass size	Minutes per epoch
U-Net [14]	31,036,481	363 MB	7
Ki-U-Net [19]	213,828	3767 MB	73
O-Net	476,321	1643 MB	11

Table 6
Comparison of performance between U-Net [14], 3D-U-Net [3], SegNet [25], and O-Net on the task of segmenting COVID-19 Ct-scan.

Model	PixAcc	Dice	IoU	Prec	Sen.	Spec.	G-mean
U-Net [14]	0.949	0.749	–	–	0.956	0.9542	0.955
3D-U-Net [3]	–	0.648	–	–	–	–	–
SegNet [25]	0.954	0.583	–	–	0.964	0.948	0.956
O-Net	0.995	0.866	0.764	0.998	0.984	0.708	0.834

ground between U-Net, which has a high number of parameters but a smaller pass size, and Ki-U-Net, which has a small number of parameters and a gigantic pass size.

Moreover, Table 6 shows the evaluation results of different models on the Coronacases dataset [30]. Unfortunately, due to the Ki-U-Net enormous pass size, we could not train it or test it. Regarding the 3D-UNET the authors trained the model for 210 epochs. In comparison, the U-Net and SegNet needed 160 epochs. Our O-Net was trained on a total of 40 epochs. Our model scored lower on specificity compared to other models because it attempts to match the shape of the infected regions as much as possible. This led our model to consider the outer layer of pixels surrounding the infected regions as infected. Since specificity is calculated based on the true negative (not infected), it caused the model to drop by about 0.2 compared to other models.

Based on the comparison made in the previous section, it is fair to say that O-Net performed better than all the previously mentioned state-of-the-art NN models. The strength of O-Net lies in its double convolutional layers. The Up/Down sampling method showed the potential of adapting to new datasets. After each stage, the data is downsampled in the encoder and upsampled in the decoder. The low number of parameters means that O-Net can converge faster than U-Net, and it is less likely to overfit due to its simplicity. The major drawback of O-Net architecture is relatively its high pass-size. This problem limits the resolution to which O-Net can be trained. However, as we previously saw, O-Net can be trained on a low-resolution input and perform relatively equal to models trained on high-resolution inputs.

4. COVIR: Our COVID-19 lung lesions virtual rendering

Our VR platform, COVIR, acts as a reading and visualisation support for medical practitioners to diagnose the COVID-19 lung infection. In this part, we dealt with improving the medical diagnostic proposed method. We provide radiologists with an immersive and interactive VR platform for diagnosis process enhancement to visualise and interact with 3D lesions and 3D lungs infected by COVID-19. Another benefit of VR diagnosis support is the possibility to offer valuable learnings for medical students and learners for COVID-19 management situations in hospitals and clinics [33]. This section presents the COVIR platform for virtual visualisation and manipulation of 3D lungs and lesions caused by COVID-19. Making a VR visualisation takes many steps to completion, starting by treating the data, 3D reconstruction, treating the 3D module, and lastly, build a VR environment to show and present the result. To accomplish this, we used 3D-Slicer [36], which allows the treating of the data, showing the Ct-scan in detail while enabling the representation of the segmented parts in 3D (Fig. 11). Most importantly, it reconstructs

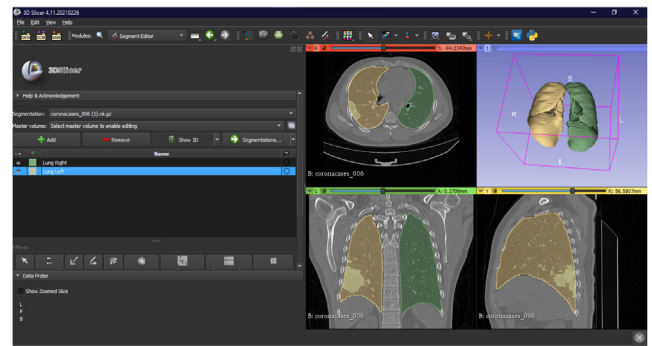


Fig. 11. Screenshot of 3D-Slicer application.

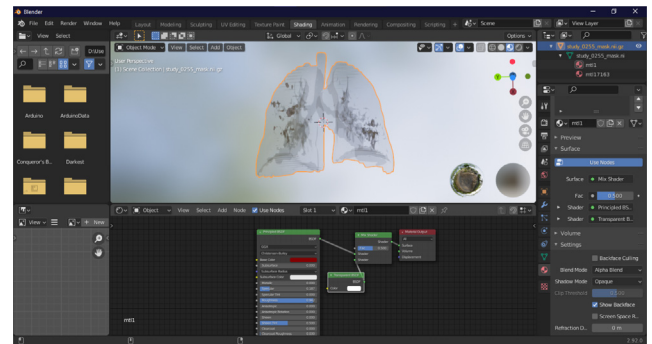


Fig. 12. Screenshot of Blender application.

the segmentation creating a 3D model file (.OBJ), which allows us to use it in Blender [37]. Blender supports the entirety of the 3D pipeline from modelling, simulation, rigging, rendering, animation, video editing and 2D animation pipeline, compositing and motion tracking. Lastly, we use Unity to create the virtual environment [38]; it gives complete access to any object created and imports 3D modules (.OBJ), which we needed to import our lung 3D model in the VR scene. Fig. 12 shows the interface of the software. The VR pipeline is summarised into the following points:

- Load the segmented CT scan in 3D slicer;
- Export the segmentations (lungs and COVID-19 lesions) in one file;
- Load it in Blender and fix any export problems (optional);
- Use the VR application, Unity 3D, to load the 3D object.

4.1. 3D Reconstruction, virtual visualisation and 3D interaction

We integrated different packages, including the interaction package for human–computer interaction management, the packages for 3D lesions, 3D lungs design, data manager package for data exchange between packages and 3D scene updates.

The hardware necessary to develop the COVIR platform should support VR rendering capacity for devices connected to the Oculus Rift S Head Mounted Display (HMD), [39] display. We used an MSI personnel computer PC with Core i9-10900KF CPU, 32 Go RAM and AMD Radeon RX 5700 XT. After treating the data in 3D-slicer and blender, we started the work in Unity. We created two scenes: The first scene represents the application description with some information on how to use it. Fig. 13 shows the COVIR platform interface. The second scene created allows the Control of 3D objects (Fig. 14).

The possible 3D interactions that users can use to interact with the application are:

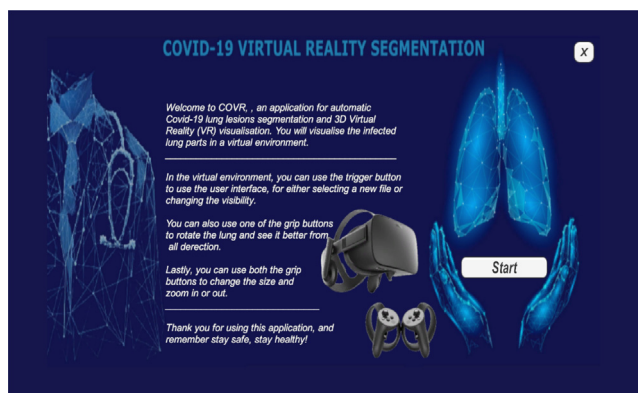


Fig. 13. COVIR platform: Description interface.

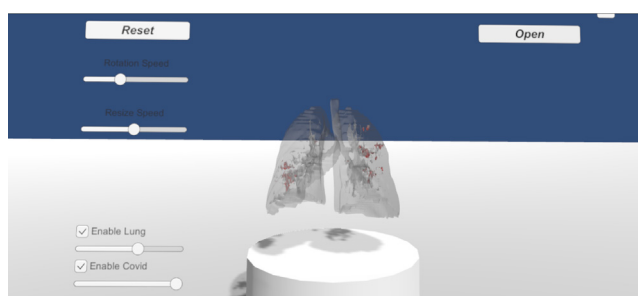


Fig. 14. COVIR platform: Control interface.

Size control. The option of controlling the size of the 3D module is to see the small parts better, with the possibility of controlling the speed of the resizing from slow to fast.

Rotation control. Controlling the rotation of the 3D module permits us to see all the sides with different speeds from slow to fast. If the rotation and the size get out of control, there is a reset button to reset to the default start.

Lung visibility. There is the lung visibility control, with the option to remove the lung to see the infection alone.

Infection visibility. There is the visibility of the infection, with the option to remove it to see the lung alone. COVIR platform functionalities with different interactions are presented on Fig. 15 for better explanation.

4.2. COVIR platform results

We developed a VR platform that allows 3D data generation, visualisation, and 3D interaction from medical Ct-scan images. Fig. 15 shows the different 3D interaction and above-mentioned functionalities. Fig. 15-a illustrates a user wearing the HMD Oculus Rift S Head and interacting with the COVIR platform. Fig. 15-b presents the virtual rendering of lungs and infected COVID-19 regions. The different 3D VR interactions are illustrated on Fig. 15-c, -d, and -e. Fig. 15-f and -e showcase the lung lesions caused by COVID-19, in addition to the COVID-19 infection rate, *COVID – Rate*, deduced from the lesions segmentation results (Eq. (8)). Fig. 15-h presents the 3D virtual view of two infected patients with *COVID – Rate* equals to 20.21% and 3.81% respectively.

4.3. COVIR platform: user study

The second contribution of this paper is to develop a VR platform for 3D visualisation and 3D interaction of automated COVID-19 lung lesion segmentation. We performed preliminary tests to

provide guidance. Medical staff, computer science students and collaborators volunteered to explore our COVIR platform. They provide their statements through user experience in a subjective questionnaire.

4.3.1. Participants

We made a subjective evaluation to explore our proposed VR COVID-19 Aid diagnosis system's efficiency. A set of 13 participants composed of medical staff, computer science students, researchers, and collaborators volunteered to investigate our COVIR platform and provide their opinions through user experience. We used Oculus Rift S Head Mounted Display (HMD) to track the participant's head and hands movement in our subjective evaluation. 46.2% of participants are used to employing VR equipment in different context. Fig. 16 displays the percentage of VR use equipment context; 50% were used to use the VR tools in gaming activities, 37.5% in professional activities, and 12.5% in leisure activity (watching films).

4.3.2. Procedure

The main procedures and scripts of the experiment are introduced in this subsection. As a brief introduction, the researcher, one of the authors, explains the global project concept and the different VR platform functionalities to the participant, and he was asked to read fulfil the first part of the questionnaire, including information about the demographic profile and the VR use experience ever before. Before starting the evaluation, the participant was able to ask any questions. To start, the experimenter shows a Ct-Scan simple viewing of lungs infected by COVID-19 to the participant. Then, the participant wears the Oculus Rift S Head Mounted Display (HMD) for the virtual visualisation and interact with the virtual lung and lesions via the triggers. The experimenter initiates the test session and lunches the VR visualisation application. After completing the trials, the participant fulfilled the second part of the questionnaire.

4.3.3. Questionnaire

Once the participant completed his trials, the researcher asked him to fill a questionnaire. Responses are recorded on a five-point-Likert scale from Strongly agree to Strongly disagree (1: Strongly Disagree, 2: Disagree, 3: Neutral, 4: Agree, and 5: Strongly Agree). The fifteen participants' survey questions concern: Utility, Ease of learning, Ease of use, and Satisfaction. The category questions are listed in Table 7.

4.3.4. Discussion and analysis

Participants were asked to state the degree to which they consented or contradicted.

Our user study yielded the stacked chart (Fig. 17). It is prevalent that the most dominant answer is "agree". Moreover, many candidates felt "strongly agree" in questions related to learning difficulty and satisfaction from using the application.

According to questions concerning the utility of the COVIR platform (Q1, Q2, Q3, Q4), 62.7% of the participants agreed on the system utility, and 23.5% provided a strongly agreed opinion. Question 2 showed that 46.2% strongly agreed and found the application helps to visualise the lungs and Covid-19 lesions better. Based on question 3, 76.9% of the participants arranged on the realistic view of the VR rendering. In question 5, 61.5% of the participants judged the application easy to learn to use; they strongly agreed, whereas 15.4% were neutral. 69.2% of the participants believed that they easily remembered how to use the platform. About the presence in the virtual environment, Q8, 92.3% of the participants affirmed that the COVIR platform was responsive to actions that they initiated. For the engagement, Q9, 84.6% of the participant asserted that the sense of moving

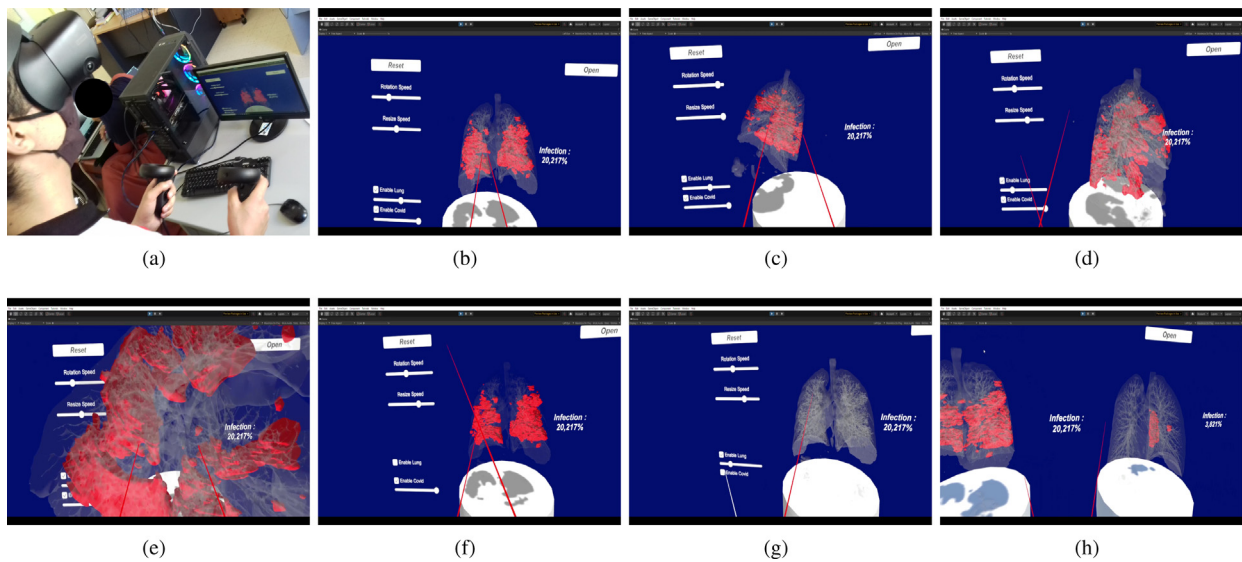


Fig. 15. COVR platform: VR functionalities, (a) VR used devices, (b) Lungs and COVID-19 lesions VR view, (c) VR interaction: 1st Rotation, (d) VR interaction: 2nd Rotation, (e) VR interaction: Resizing, (f) Infection visibility, (g) Lung visibility, (h) Two lungs VR view.

Table 7
Evaluation study's questionnaires.

Evaluation criteria	Utility	Ease of learning	Ease of use	Satisfaction
Questions	Q1. Is it useful? Q2. Does it help to visualise the lungs and Covid-19 lesions better? Q3. Is the view of the lung and the infected regions realistic? Q4. Does it everything I would expect it to do?	Q5. Is it easy to learn to use it? Q6. I learned to use it quickly? Q7. I easily remember how to use it?	Q8. Presence: The virtual environment was responsive to actions that I initiated? Q9. Engagement: The sense of moving around inside the virtual environment was compelling? Q10. Immersion: I felt stimulated by the virtual environment. Q11. Flow: I felt I could perfectly control my actions. Q12. Usability: I found it easy to use the apparatus (Oculus headset, trigger). Q13. Technology adoption: If I use again the same virtual environment, my interaction with the environment would be clear and understandable for me.	Q14. I enjoyed being in the virtual environment? Q15. I enjoyed using the COVRplatform?

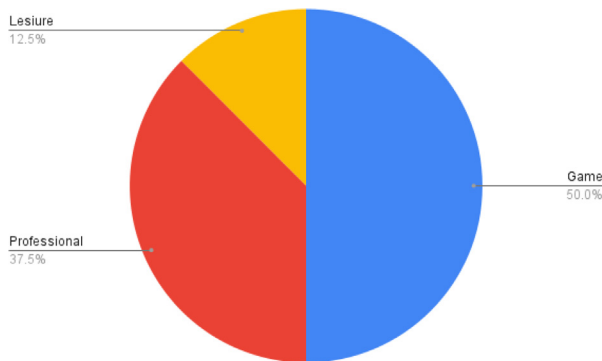


Fig. 16. VR equipment use context.

around inside the virtual environment was compelling. Based on question 12, 30% of the participants were neutral to the ease of use of the VR apparatus (Oculus headset, trigger). For the VR immersion, 66.7% of the participants felt stimulated by the virtual environment, whereas others, 8.3%, were neutral. 84.6% of the participants enjoyed being in this virtual environment based on question 14. Question 15 showed that all the participants enjoyed using the COVR platform (46.2% of the participant strongly agreed, and 53.8% agreed).

However, we noted that many researchers and students wanted to include more diseases and show more organs, as they felt this application has educational potential. One doctor expressed less interest in the visualisation and requested displaying more technical details regarding the patient and the Ct-Scan. Another doctor found information like the percentage more critical than the displaying Ct-Scan. Regardless, he proposed that information about which lung is more infected may also help. This user study does not intend to generalise the findings due to the target application. Therefore, the need for more in-depth analysis is revealed and required with practical evaluations and feedback.

5. Conclusion and future work

COVID-19 has been considered the world's most threatening challenge of the current century. Due to new emerging COVID-19 variants, the number of COVID-19 daily cases is still increasing in many areas. It is, consequently, imperative for healthcare experts and authorities across the globe to find practical solutions to manage the COVID-19 pandemic. With the COVID-19 spread, radiologists faced massive overwork, which may cause fatal errors in patients diagnoses. This work presented a new NN architecture to automatically segment infected Ct-scans caused by COVID-19, which gives fast and accurate results based on the results obtained with optimised computing power and memory occupation.

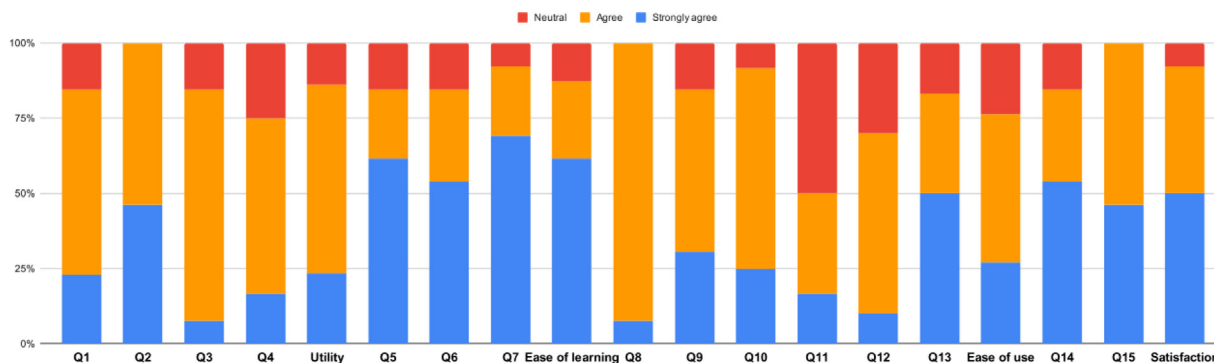


Fig. 17. Evaluation study: Utility, Ease of use and satisfaction.

Our work started by collecting segmented COVID-19 datasets we could only get access to publicly published data. However, we still faced a data shortage. To solve that issue, we used different data augmentation methods to increase the size of our dataset. Using the augmented data, we tested NN models known for being well-performing in the medical imagery segmentation. After comparing the results and weighing the advantages and drawbacks of each NN model, we proposed a novel NN model that aims to enhance these state of the art models performance. Termed as O-Net, our improved architecture is based on U-Net, inspired by the Ki-U-Net double convolutional channels. It consists of two convolutional autoencoders with an upsampling channel and a downsampling channel. The O-Net enhanced the U-Net performance based on experimental results. Evaluating our model gave remarkable results compared to the other state-of-the-art models. Our O-Net model scored 0.86 on the Dice coefficient. After building and testing multiple variations of O-Net, we noted the excellent performance it showed even with small networks sizes. However, in future work, we are interested in replacing the Up/Down sampling architecture with transformer encoders that may be more robust and optimal. Once we were confident with our segmentation system, we developed a COVIR, which is a VR platform to visualise the segmented Ct-scan lung lesions caused by COVID-19. We created a user-friendly interface, and we integrated a function that loads and shows Ct-scan in the virtual environment, with many interactions like rotation and scaling, which gave better visualisation of the results. Our second contribution is the VR visualisation and 3D interaction of COVID-19 lesions in an immersive virtual environment. We provide radiologists with an aid COVID-19 diagnosis tool with a clear and realistic view to provide preliminary analysis and interpretation. Based on the user evaluation study results, the COVIR platform provides a clear and realistic view of lungs and COVID-19 lesions for preliminary analysis and interpretation. Future works plan to optimise the application and expand its use case to include other pulmonary infections diseases. We plan to include more details and informations about the infected regions.

CRedit authorship contribution statement

Kahina Amara: Conceptualisation, Methodology, Investigation, Resources, Writing – original draft, Data acquisition. **Ali Aouf:** Methodology, Software, Deep learning model implementation, Formal analysis, Reviewing and editing. **Hoceine Kennouche:** Methodology, 3D reconstruction, VR visualisation, 3D interaction. **A. Oualid Djekoune:** Resources, Supervision. **Nadia Zenati:** Resources, Supervision. **Oussama Kerdjadj:** Methodology, Resources, Investigation, Reviewing and editing, Software, Validation. **Farid Ferguene:** Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- [1] WHOrganization. Tracking sars-cov-2 variants. 2021, (accessed July 29, 2021). URL <https://www.who.int/en/activities/tracking-SARS-CoV-2-variants/>.
- [2] WHOrganization. Coronavirus, health topics. 2021, (accessed October 12, 2021). URL https://www.who.int/health-topics/coronavirus#tab=tab_1.
- [3] Bressem KK, Niehues SM, Engel G, Hamm B, Makowski MR, Vahldiek JL, Adams LC. 3D u-net for segmentation of covid-19 associated pulmonary infiltrates using transfer learning: State-of-the-art results on affordable hardware. 2021, arXiv:2101.09976.
- [4] Saood A, Hatem I. Covid-19 lung ct image segmentation using deep learning methods: U-net versus segnet. 2021.
- [5] Chen J, Wu L, Zhang J, Zhang L, Gong D, Zhao Y, Hu S, Wang Y, Hu X, Zheng B, Zhang K, Wu H, Dong Z, Xu Y, Zhu Y, Chen X, Yu L, Yu H. Deep learning-based model for detecting 2019 novel coronavirus pneumonia on high-resolution computed tomography: a prospective study in 27 patients. 2020, <http://dx.doi.org/10.1101/2020.02.25.20021568>.
- [6] Li L, Qin L, Xu Z, Yin Y, Wang X, Kong B, Bai J, Lu Y, Fang Z, Song Q, Cao K, Liu D, Wang G, Xu Q, Fang X, Zhang S, Xia J, Xia J. Using artificial intelligence to detect covid-19 and community-acquired pneumonia based on pulmonary ct: Evaluation of the diagnostic accuracy. *Radiology* 2020;296:200905.
- [7] Ying S, Zheng S, Li L, Zhang X, Zhang X, Huang Z, Chen J, Zhao H, Wang R, Chong Y, Shen J, Zha Y, Yang Y. Deep learning enables accurate diagnosis of novel coronavirus (covid-19) with ct images. 2020, <http://dx.doi.org/10.1101/2020.02.23.20026930>, medRxiv: [https://www.medrxiv.org/content/early/2020/02/25/2020.02.23.20026930](https://www.medrxiv.org/content/early/2020/02/25/2020.02.23.20026930.full.pdf). URL <https://www.medrxiv.org/content/early/2020/02/25/2020.02.23.20026930>.
- [8] Garain A, Basu A, Giampaolo F, Velasquez J, Sarkar R. Detection of covid-19 from ct scan images: A spiking neural network-based approach. *Neural Comput Appl* 2021. <http://dx.doi.org/10.1007/s00521-021-05910-1>.
- [9] Stefano A, Comelli A. Customized efficient neural network for covid-19 infected region identification in ct images. *J Imaging* 2021;7(8). <http://dx.doi.org/10.3390/jimaging7080131>, URL <https://www.mdpi.com/2313-433X/7/8/131>.
- [10] Zhang Y-D, Satapathy S, Shuaiqi L, Li G-R. A five-layer deep convolutional neural network with stochastic pooling for chest ct-based covid-19 diagnosis. *Mach Vis Appl* 2021;32. <http://dx.doi.org/10.1007/s00138-020-01128-8>.
- [11] Qi S, Xu C, Li C, Tian B, Xia S, Ren J, Yang L, Wang H, Yu H. Dr-mil: deep represented multiple instance learning distinguishes covid-19 from community-acquired pneumonia in ct images. *Comput Methods Programs Biomed* 2021;211:106406. <http://dx.doi.org/10.1016/j.cmpb.2021.106406>, URL <https://www.sciencedirect.com/science/article/pii/S0169260721004806>.
- [12] Oulefki A, Aгаian S, Trongtirakul T, Benbelkacem S, Aouam D, Zenati-Henda N, Abdelli M-L. Virtual reality visualization for computerized covid-19 lesion segmentation and interpretation. *Biomed Signal Process Control* 2022;73:103371.
- [13] Li C, Tan Y, Chen W, Luo X, He Y, Gao Y, Li F. Anu-net: Attention-based nested u-net to exploit full resolution features for medical image segmentation. *Comput Graph* 2020;90:11–20. <http://dx.doi.org/10.1016/j.cag.2020.05.003>, URL <https://www.sciencedirect.com/science/article/pii/S0097849320300546>.

- [14] Ronneberger O, Fischer P, Brox eN, Thomas[†], Hornegger J, Wells WM, Frangi AF. U-net: Convolutional networks for biomedical image segmentation. In: *Medical image computing and computer-assisted intervention – MICCAI 2015*. Cham: Springer International Publishing; 2015, p. 234–41.
- [15] Mantovani E, Zucchella C, Bottirolti S, Federico A, Giugno R, Sandrini G, Chiamulera C, Tamburin S. Telemedicine and virtual reality for cognitive rehabilitation: A roadmap for the covid-19 pandemic. *Front Neurology* 2020;11:926. <http://dx.doi.org/10.3389/fneur.2020.00926>, URL <https://www.frontiersin.org/article/10.3389/fneur.2020.00926>.
- [16] Zorzal ER, Sousa M, Mendes D, dos Anjos RK, Medeiros D, Paulo SF, Rodrigues P, Mendes JJ, Delmas V, Uhl J-F, Mogorrón J, Jorge JA, Lopes DS. Anatomy studio: A tool for virtual dissection through augmented 3d reconstruction. *Comput Graph* 2019;85:74–84. <http://dx.doi.org/10.1016/j.cag.2019.09.006>, URL <https://www.sciencedirect.com/science/article/pii/S0097849319301578>.
- [17] Fischer J, Eichler M, Bartz D, Straßer W. A hybrid tracking method for surgical augmented reality. *Comput Graph* 2007;31(1):39–52. <http://dx.doi.org/10.1016/j.cag.2006.09.007>, URL <https://www.sciencedirect.com/science/article/pii/S0097849306001841>.
- [18] Singh RP, Javaid M, Kataria R, Tyagi M, Haleem A, Suman R. Significant applications of virtual reality for covid-19 pandemic. *Diabetes Metabolic Syndrome: Clin Res Rev* 2020;14(4):661–4. <http://dx.doi.org/10.1016/j.dsx.2020.05.011>, URL <https://www.sciencedirect.com/science/article/pii/S1871402120301302>.
- [19] Jose JM, Sindagi V, Hacihaliloglu I, Patel VM. Kiu-net: Towards accurate segmentation of biomedical images using over-complete representations. 2020. [arXiv:2006.04878](https://arxiv.org/abs/2006.04878).
- [20] Yuheng S, Hao Y. Image segmentation algorithms overview. 2017. [CoRR abs/1707.02051](https://arxiv.org/abs/1707.02051). URL <https://arxiv.org/abs/1707.02051>.
- [21] Uchida S. Image processing and recognition for biological images, development, growth & differentiation. 2013, p. 55. <http://dx.doi.org/10.1111/dgd.12054>.
- [22] Chakraborty S, Mali K. Sufmofpa: A superpixel and meta-heuristic based fuzzy image segmentation approach to explicate covid-19 radiological images. *Expert Syst Appl* 2021;167:114142. <http://dx.doi.org/10.1016/j.eswa.2020.114142>, URL <https://www.sciencedirect.com/science/article/pii/S0957417420308897>.
- [23] Nixon MS, Aguado AS. 8 - region-based analysis. In: Nixon MS, Aguado AS, editors. *Feature extraction and image processing for computer vision* (fourth edition). 4th ed.. Academic Press; 2020, p. 399–432. <http://dx.doi.org/10.1016/B978-0-12-814976-8.00008-7>, URL <https://www.sciencedirect.com/science/article/pii/B9780128149768000087>.
- [24] Lamba H. Understanding semantic segmentation with unet. *Towards Data Sci* 2019. URL <https://towardsdatascience.com/understanding-semantic-segmentation-with-unet-6be4f42d4b47>.
- [25] Badrinarayanan V, Kendall A, Cipolla R. Segnet: A deep convolutional encoder–decoder architecture for image segmentation. 2017. <http://dx.doi.org/10.1109/TPAMI.2016.2644615>.
- [26] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. 2015. [CoRR arXiv:1409.1556](https://arxiv.org/abs/1409.1556).
- [27] Lakshmanamoorthy R. Hands-on transunet: Transformers for medical image segmentation. 2021. <https://analyticindiamag.com/hands-on-transunet-transformers-for-medical-image-segmentation/>.
- [28] M. s. COVID-19, [link]. URL <http://medicalsegmentation.com/covid19/>.
- [29] Fan D-P, Zhou T, Zhou G-PJY, Chen G, Fu H, Shen J, Shao L. Inf-net: Automatic covid-19 lung infection segmentation from ct images. 2020. [arXiv:2004.14133](https://arxiv.org/abs/2004.14133).
- [30] Jun M, Cheng G, Yixin W, Xingle A, Jiantao G, Ziqi Y, Mingqing Z, Xin L, Xueyuan D, Shucheng C, Hao W, Sen M, Xiaoyu Y, Ziwei N, Chen L, Lu T, Yuntao Z, Qiongjie Z, Guoqiang D, Jian H. Covid-19 ct lung and infection segmentation dataset. 2020. <http://dx.doi.org/10.5281/zenodo.3757476>.
- [31] Tsai EB, Simpson S, Lungren MP, Hershman M, Roshkovan L, Colak E, Erickson BJ, Shih G, Stein A, Kalpathy-Cramer J, Shen J, Hafez M, John S, Rajiah P, Pogatchnik BP, Mongan J, Altinmakas E, Ranschaert ER, Kitamura FC, Topff L, Moy L, Kanne JP, Wu CC. The rsna international covid-19 open radiology database (ricord). *Radiology* 2021;299(1):E204–13. <http://dx.doi.org/10.1148/radiol.2021203957>, PMID: 33399506. [arXiv:https://doi.org/10.1148/radiol.2021203957](https://doi.org/10.1148/radiol.2021203957).
- [32] Luke GP, Hoffer-Hawlik K, Namen ACV, Shang R. O-net: A convolutional neural network for quantitative photoacoustic image segmentation and oximetry. 2019. [arXiv:1911.01935](https://arxiv.org/abs/1911.01935).
- [33] De Ponti Roberto, Marazzato Jacopo, Maresca Andrea M, Rovera Francesca, Carcano Giulio, Ferrario Marco M. Pre-graduation medical training including virtual reality during covid-19 pandemic: a report on students' perception. *BMC Med Educ* 2020. <http://dx.doi.org/10.1186/s12909-020-02245-8>.
- [34] Jadhav S, Deng G, Zawin M, Kaufman AE. Covid-view: Diagnosis of covid-19 using chest ct. *IEEE Trans Vis Comput Graphics* 2022;28(1):227–37. <http://dx.doi.org/10.1109/TVCG.2021.3114851>.
- [35] Morozov S, Andreychenko A, Blokhin I, Gelezhe P, Gonchar A, Nikolaev A, Pavlov N, Chernina V, Gombolevskiy V. Mosmeddata: data set of 1110 chest ct scans performed during the covid-19 epidemic. *Digit Diagn* 2021;1:49–59. <http://dx.doi.org/10.17816/DD46826>.
- [36] 3d slicer. 3D slicer. 2021. URL <https://www.slicer.org>.
- [37] Blender. Blender. 2021. URL <https://www.blender.org>.
- [38] Start creating with unity, Unity Store. URL https://store.unity.com/download-nuo?_ga=2.149713022.959971313.1632923638-2054984004.1632923638.
- [39] Oculus rift s features, Oculus. URL <https://www.oculus.com/rift-s/features/>.

Kahina Amara obtained her PhD degree at the university of science and technology Houari Boumediene USTHB, Algiers, Algeria, in 2018 and her master's in control and robotics at USTHB in 2011. She works at the Centre for Development of Advanced Technologies CDTA with the IRVA team. Her main research includes: Augmented and Virtual Reality, 3D Interaction, Affective Computing, Computer Vision, Emotion Recognition, Healthcare.

Ali Aouf received his Master's degree of Smart Computational Systems (SII) in 2021 at university of science and technology Houari Boumediene USTHB, Algiers, Algeria. He is Machine-learning engineer with professional experience in web and software development. His research interests includes Machine Learning, Computational Neuroscience, and Reinforcement Learning.

Hocine Kennouche received his Master's degree of Smart Computational Systems (SII) in 2021 at university of science and technology Houari Boumediene USTHB, Algiers, Algeria. He is has professional experience in web and software development. He is interested in virtual reality, serious game and machine learning.

Oualid Djekoune received the PhD degree in Process Control and Robotics from the University of Science and Technology Houari Boumediene (USTHB), Algiers, in 2010, and the University Habilitation degree in 2017. Since 1998, he has worked as a permanent researcher at the Center for the Development of Advanced Technologies (CDTA) and Head of the Robotics and Industriel Automation Division since 2018. His current research interests include Computer Vision for Automation, Robotics and Virtual and Augmented Reality applications. He is the author or co-author of many scientific documents.

Nadia Zenati is a research director at the Centre for Development of Advanced Technologies, CDTA, Algiers, Algeria. She is head of Human Machine interaction, Virtual and Augmented Reality - IRVA- team. She obtained her PhD degree in automation and computer science in 2008 at Franche-Comté university, France. She has authored many publications, articles, conference papers, and book chapters. She coordinates multidisciplinary research projects in computer vision, augmented and virtual reality and human-machine interaction focusing on 3D interaction and Augmented and virtual reality for medical and healthcare applications.

Oussama Kerdjdidj obtained his PhD degree at the University of Laghouat, Algeria, in 2019. His research interests involve embedded systems, high-performance computing using FPGAs, Hardware and Software implementation, signal processing, and Artificial Intelligence applied to Healthcare applications. He is currently a permanent researcher at the Center for Development of Advanced Technologies (CDTA)

Farid Ferguene received his engineer degree from the University of the Sciences and Technology, Houari Boumediene (USTHB), Algeria, in 1993. In 1994 he joined the LRPE laboratory of the same university, where he received his M.Sc. and PhD degrees in 1997 and 2008, respectively. All his degrees are in electronic engineering. He is a USTHB university professor and team leader at LRPE Laboratory. His research interests include advanced systems control and robotics.