

# CovSeg-Unet: End-to-End Method-based Computer-Aided Decision Support System in Lung COVID-19 Detection on CT Images

Fatima Zahra EL BIACH, Imad IALA, Hicham LAANAYA, Khalid MINAOUI  
LRIT Associated Unit to the CNRST-URAC N29  
Faculty of Sciences, Mohammed V University in Rabat  
B.P.1014 RP, Rabat IT center  
Morocco

**Abstract**—COVID-19 epidemic continues to threaten public health with the appearance of new, more severe mutations, and given the delay in the vaccination process, the situation becomes more complex. Thus, the implementation of rapid solutions for the early detection of this virus is an immediate priority. To this end, we provide a deep learning method called CovSeg-Unet to diagnose COVID-19 from chest CT images. The CovSeg-Unet method consists in the first time of preprocessing the CT images to eliminate the noise and make all images in the same standard. Then, CovSeg-Unet uses an end-to-end architecture to form the network. Since CT images are not balanced, we propose a loss function to balance the pixel distribution of infected/uninfected regions. CovSeg-Unet achieved high performances in localizing COVID-19 lung infections compared to others methods. We performed qualitative and quantitative assessments on two public datasets (Dataset-1 and Dataset-2) annotated by expert radiologists. The experimental results prove that our method is a real solution that can better help in the COVID-19 diagnosis process.

**Keywords**—Deep learning; COVID-19; loss function; balanced data

## I. INTRODUCTION

In December 2019, a viral pneumonia epidemic of unknown etiology emerged in Wuhan city, Hubei province, China [1]. On January 9, 2020, the World Health Organization (WHO) and Chinese Health Authorities officially announced the discovery of a new coronavirus. This pneumonia is an infectious disease caused by a virus identified under the name SARS-CoV-2 (Severe Acute Respiratory Syndrome CoronaVirus-2) by the ICTV (International Committee on Taxonomy of Viruses) [2], and causing a disease called COVID-19 (COroNaVIrus Disease 2019). SARS-CoV-2 belongs to the coronavirus family. The reservoir of this virus is probably animal. Although SARS-CoV-2 closely resembles a virus detected in a bat, the animal that transmits it to humans has yet to be identified with certainty. Several research studies suggest that the pangolin, a small mammal eaten in southern China, could be involved as an intermediate host between bats and humans.

The new coronavirus has been confirmed to be transmitted between humans [3], and this is done mainly by air or by close

contact with a contagious subject. Smaller particles can also be emitted in the form of aerosols during speech or during coughing efforts, which would explain that the virus could persist suspended in the air in an unventilated room. Finally, the virus can retain infectivity for a few hours on inert surfaces from where it can be transported by the hands. According to data from the World Health Organization, updated up to 24 hours on June 18, 2021, COVID-19 has affected 220 countries and territories, causing 178,584,744 people to be infected and 3,866,607 deaths worldwide. The overall number of people recovered is 163,102,134. Currently, the active cases are 11,616,003 of which 99.3% in mild condition and 0.7% in serious or critical condition, which poses a great threat to international human health.

Due to the vaccination process slowness, the high rate of virus contamination, and the appearance of new dangerous COVID-19 mutations, it is essential to detect and identify the disease at an early stage so that suspected patients do not infect the healthy population. As a result, new requirements for the prevention and control strategy must be put in place. Reverse Transcription Polymerase Chain Reaction (RT-PCR), gene sequencing for respiratory, or blood samples confirm the diagnosis of COVID-19. However, the false negatives of the RT-PCR [4], the delay in obtaining the results, and the tests carried out on people not strongly suspected of being infected with COVID-19 imply that numerous COVID-19 patients would not be identified quickly to isolate them from others. In addition, given the rapid and contagious spread of the virus, they present a real threat to infect a larger population, especially in areas with high epidemics. On the other hand, chest examinations quickly established themselves as an interesting diagnostic tool, given the characteristic presentation of COVID-19 lesions [5]. These tests can identify lesions, underlying conditions and complications associated with acute airway conditions. Consequently, the use of CT in particular High Resolution Computed Tomography (HRCT) could provide enormous help to radiologists [7] for the diagnosis, follow-up or investigation of pulmonary complications in patients suspected or confirmed of COVID-19. Thus, the development of an artificial intelligence (AI) method based on deep learning could help them enormously to assess the degree of lung damage caused by COVID-19.

According to [6], the authors indicated that CT images can be used to detect COVID-19 even before certain clinical symptoms are observed. Typical signs of COVID-19 appear in CT images as unilateral, multifocal, and terminal Ground Glass Opacity (GGO). This is a hazy cloud above the lungs that indicates a variety of problems, and may mean that the lungs are partially filled with inflamed material, and there is thickening in lung tissue or partial breakdown of the alveoli and tiny air sacs of the lungs. Pleural effusion, lymphadenopathy, and condensation [3], which are air spaces in the lungs filled with a substance, usually pus, blood or water, surrounded by an opaque edge of frosted glass, and although this is a common feature of lung disease, it may be more characteristic of COVID-19. To detect COVID-19 disease at an early stage, it is necessary to detect and locate these pathological changes in a short time. The growing number of patients and the limited number of well-trained expert radiologists in most hospitals prevent and slow down the process of early detection. Indeed, the use of deep learning methods for the automatic segmentation of the COVID-19 CT model has become paramount, and may offer an effective solution to identify and locate signs of COVID-19 in CT images [8].

In this paper, a new efficient method of COVID-19 diagnostic using Deep Learning network is proposed. Section II presents the related works, Section III shows problem statement, and Section IV explains in detail the proposed method. Section V describes simulations experiments. Section VI discusses the obtained results. Finally, the summary and future works are delineated in Section VII.

## II. RELATED WORK

In the literature, numerous methods of segmentation based on deep learning networks have been used to process and analyze chest X-ray or CT images for the COVID-19 diagnosis [9]. These methods mainly consist of delineating the regions of interest in these images, such as lobes, bronchopulmonary segments, lung, and infected regions or lesions for further quantification and evaluation.

CT provides detailed and high definition three dimensional images to detect COVID-19. Among the segmentation methods used for the diagnosis of COVID-19 we cite, U-Net [10], UNet++ [11], V-Net [12]. The authors of [9] have proposed a 3D architecture of U-Net using inter-slice information; this method consists in replacing the conventional layers of U-Net by a 3D version. In [12], the authors proposed the V-Net architecture, in which they used the residual blocks as the basic convolutional block, and optimized the network by a loss of dice. In [13], the authors proposed the Attention U-Net method, which captures fine structures to locate lesions and pulmonary nodules in medical images. Generally, the large number of well-labeled images is the key to forming an efficient and robust segmentation network. In the case of COVID-19 image segmentation, the data used during the training phase is limited and often unavailable because manual lesion delineation is a difficult operation and requires a lot of time.

Several other research works obtaining reasonable segmentation results have been proposed in this context. The lung segmentation field is experiencing a lack of labeled medical images, as a result, semi supervised and unsupervised methods are very favorable and recommended in studies on COVID-19, as in [10], the authors used an unsupervised method to generate pseudo-segmentation masks for the images. In [14], the authors proposed a new COVID-19 lung infection segmentation network called Inf-Net to detect infected regions from chest CT images. This method uses a parallel decoder for aggregating high-level features and generating a global map. Then, it uses a semi-supervised segmentation framework based on a propagation strategy chosen at random to overcome the lack of labeled data. In [15], the authors proposed a computer-assisted diagnosis (CAD) system based on the YOLO predictor to detect and diagnose COVID-19. The CAD method calls the data balancing regularizations, transfer learning, and augmentation to improve the overall diagnostic performance for COVID-19. The authors of [16] proposed a synergistic approach based on deep meta-learning to accelerate the detection of COVID-19 cases. This approach uses contrastive learning with a pre-trained ConvNet encoder for the classification of COVID-19 cases. In [17], the authors proposed a computer-aided detection (CAD) method to assist radiologists to automatically detect COVID-19 on the chest X-ray images. The proposed method uses the DLs: the Discrimination-DL to extract lung features from chest X-ray images, and the Localization-DL to localize and assign the infected lung region. In [18], the authors built prognosis models to predict the patients' severity outcomes. The proposed method is based on deep learning in the CT image segmentation process for COVID-19 pneumonia, and it uses datasets from multiple institutions worldwide to validate the proposed models. In [19], the authors propose a CovFrameNet framework to detect COVID-19 cases using CT images, which incorporates an image preprocessing mechanism and a deep learning model for smoothing, denoising, feature extraction, classification, and performance measurement.

## III. PROBLEM STATEMENT

Our main goal through this paper is to diagnose COVID-19 lung infection in chest CT images. In this regard, we use an architecture similar to that of the U-Net [10] approach called CovSeg-Unet method, U-Net is considered as the most commonly used algorithm in medical image segmentation. U-Net is a symmetrical encoder/decoder architecture consisting of several stages. Each stage of the encoder performs a set of operations such as, convolution, normalization, max pooling, activation, concatenation etc. In parallel, each stage of the decoder performs deconvolution operations. U-Net method uses jump connections allowing the exploitation of local and global information. These connections concatenate the subsampling characteristics of the contraction path with those of the up-sampling of the expanding path.

The general idea of the CovSeg-Unet approach is described as follows. Let  $S$  be a learning space which contains a set of  $n$  images  $X = X_1 \dots X_n$ , and  $n$  corresponding ground-truth masks  $Y = Y_1 \dots Y_n$ . From the  $Y$  ground-truth masks, the network learns the lung infections distribution of the  $X$

learning images to establish an image-to-image mapping relationship between  $X$  and  $Y$ , this map is defined as follows:  $\Phi = f_{dec} \circ f_{enc}$ .

$f_{enc}$  is the encoder function that learns the characteristic vectors of infected lung regions to establish a functional space. While  $f_{dec}$  is the decoder function, which learns the spatial localization of features to better locate the infected/uninfected region. And  $\Phi$  is the set of probabilities extracted from an input biomedical image, and is represented by  $\Phi(f) = P_1, P_2, \dots, P_n$ , where  $P_i = \{0,1\}$  are the probabilities given to the last convolutional layer of classification by the nonlinear function SoftMax. The key idea here focuses on the spatiotemporal modeling of characteristic points that can represent lung infections.

#### IV. PROPOSED METHOD

In this section, we present the CovSeg-UNet approach that we propose to diagnose COVID-19 from CT images. The CovSeg-UNet approach architecture is shown in Fig. 1. Generally, the CovSeg-UNet approach architecture's is made up of two blocks; the first one consists of preprocessing the CT images. While the second one performs all encoder/decoder operations to learn high level features from training sets, and locate the spatial information. The CovSeg-UNet network details are shown.

##### A. Preprocessing

One of the major problems encountered in the deep learning is processing of biomedical images that coming from several machines with different acquisition parameters. To deal with this problem, we apply a preprocessing step on these images to improve the learning of the COVID-19 suspect recognition model despite of the data heterogeneity. This step of preprocessing consists of two steps; the first one concerns the normalization of the signal that processes the intensity of each scanner CT voxel. In this step, we chose the pulmonary window value from Table I to separate the lungs from the other organs. Since each CT scanner has its own Hounsfield (HU) units, consequently, the data collected in different hospitals will have different HUs. For this reason, we are

using a multi-valued window, ( $W_L$  is the window center value, and  $W_W$  is the window width), the  $W_L$  value is randomly assigned from -600 to -500, while the  $W_W$  value is fixed at 1200. In the second preprocess step, we normalizing the CT images to be in [0, 255]. Through this process we normalize multiple images of different scanners to the same standard, by separating the lungs from other organs, removing unnecessary information (CT features, etc.) or/and noise, increasing images, and improved accuracy.

##### B. CovSeg-UNet Architecture

In order to detect COVID-19 lung infections using CT images, we propose the CovSeg-UNet approach, which characterized by an end-to-end architecture based on one of the most robust approaches in biomedical image segmentation that is U-Net. The network of the CovSeg-UNet approach is supplied as input by pre-processed CT images and their ground-truth masks. Our proposed method relies on an encoder to extract contextual feature maps from pre-processed images to reduce the dimensions of CT images, and a decoder to locate feature map information in the image. Table II shows the detailed architecture of our encoder/decoder.

Generally, the encoder is made up of four residual blocks (ResBlock), already pre-formed on the ImageNet database, the use of these blocks allows us to avoid the disappearance/explosion gradient, to preserve the local information, to improve precision by increasing the depth of the network, and to optimize the formation of layers. Each residual block is made up of two blocks (conv\_bloc and identity\_bloc), and is expressed by  $R(I) + I$ , where  $I$  is the input vector, and  $R(\cdot)$  represents the mapping from input  $I$  to the output of the residual unit. However, the decoder has two inputs, one input from the parallel layer of the encoder and a second from the previous layer of the decoder. Finally, the decoder output is sent to the SoftMax activation function (see equation (1)) for the prediction of the region infected with COVID-19. Algorithm 1 describes the training steps of the proposed model.

$$\sigma(\vec{z})_i = \frac{e^{x_i}}{\sum_{j=1}^k e^{z_j}} \quad (1)$$

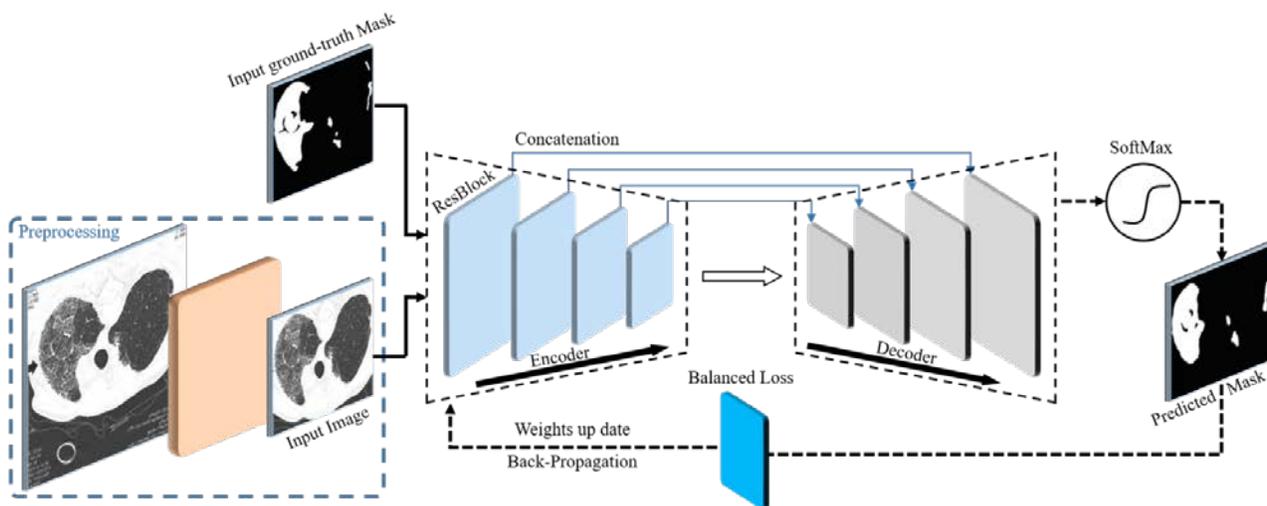


Fig. 1. Overview of the Proposed Framework for Diagnosing COVID-19 from CT Images.

TABLE I. DIFFERENT HOUNSFIELD VALUES OF DIFFERENT SUBSTANCES

Substance	Hounsfield Unit (HU)
Air	-1000
Bone	+700 to +3000
Lungs	-500
Water	0
Kidney	30
Blood	+30 to +45
Grey matter	+37 to +45
liver	+40 to +60
White matter	+20 to +30
Muscle	+10 to +40
Soft Tissue	+100 to +300
Fat	-100 to -50
Cerebrospinal fluid(csf)	15

TABLE II. DETAILED ARCHITECTURE OF THE PROPOSED METHOD. WHERE THE ENCODER IS THE RESIDUELLE BLOCK SIMILAIRE TO RESNET50[20] ARCHITECTURE, AND THE DECODER BLOCK IS A SUCCESSION OF BATCHNORMALISATION, RELU, CONV2D, BATCHNORMALISATION, AND RELU OPERATIONS

Stages	Layers	Output size
Initial	Conv2d(7 x 7), stride = 2 BatchNormalisation ReLu MaxPoling	(64 x 64 x 64)
Encoder	Stage 1 : x2 ResBlock(64 x 64 x 256) Stage 2 : x3 ResBlock(128 x128 x512) Stage 3 : x5 ResBlock(256x256 x1024) Stage 4 : x2 ResBlock(512x512 x2048)	(64 x 64 x 256) (32 x 32 x 512) (16 x16 x 1024) (8 x 8 x 2048)
Decoder	Stage 1 : UpSampling(16 x 16 x 256) Stage 2 : UpSampling(32 x 32 x 128) Stage 3 : UpSampling(64 x 64 x 64) Stage 4 : UpSampling(128 x 128 x 32)	(32 x 32 x128) (64 x 64 x 64) (128 x128 x32) (256 x 256 x16)
Final	Conv2d(256 x 256 x 16) Softmax Activation(256 x 256 x 1)	(256 x 256 x 1) (256 x 256 x 1)

Where  $\vec{z}$  is the input vector of the softmax function ( $z_0, \dots, z_k$ ), all  $z_i$  values can take any real value.  $e^{z_j}$  is applied to have a positive value for each element of the input vector. The term at the formula bottom is the normalization term, which makes it possible to have a  $Sum = 1$  of all the output values (are each in the range (0, 1)) of the function, thus constituting a valid probability distribution,  $k$  represents the number of classes in the multiclass classifier.

**Algorithm 1:** Training procedure for CovSeg-UNet method

```

input :  $X$ : Image CT;  $Y$ : Label;  $N$ : Batch-size;  $\lambda_1$ ;  $\lambda_2$ ;
 $Lr$ : Learning-rate;  $W_0$ : Initial weights;
output:  $P_L$ : Predicted mask;
1 begin
2    $(X_{train}, Y_{train}), (X_{valid}, Y_{valid}) \leftarrow \text{split}((X, Y),$ 
    $\text{split-size}=0.2)$ 
3   while epoch  $\leq 200$  do
4     for mini batch sample  $x_k \{x_{train}, x_{valid}\}_{k=1}^N$  do
5        $z_1 = f_{encoder}(x_k)$ 
6        $z_2 = f_{decoder}(z_1)$ 
7       predict( $z_2$ ) with SoftMax equation
8       Compute  $\Delta_w$  the stochastic gradient by
       minimizing the loss function eq. 2
9       Update weights
10       $w \leftarrow w + \alpha \cdot \text{AdamOptimiser}(w, \Delta_w)$ 

```

C. Loss

The imbalanced class problem is considered as the major challenge in the detection process of lung infections because the distribution of infected/uninfected regions is highly skewed (the infected regions vary between 0% and 20% of the pixels of the lung image). So, if the loss function does not consider this problem, the model will classify the majority of pixels as uninfected regions, and become overfit. For this reason, we use a class-balanced cross-loss function and the penalty factor  $\lambda$ .

The loss function is defined as a weighted sum of two loss functions; the balanced cross loss  $L_{BCE}$  and the inverse cross loss  $L_{ICE}$ :

$$L_B = \lambda_1 \cdot L_{BCE} + \lambda_2 \cdot L_{ICE} \quad (2)$$

Respecting to the cross-loss, we use the balanced cross-entropy to overcome noise in biomedical images, we also use the balance parameter  $W$  to balance the pixel distribution of infected/uninfected regions in  $L_{BCE}$ :

$$L_{BCE} = -w_c \sum_{i \in y_c} q(y_i = 1 | s) \log(p(y_i = 1 | s)) - w_{cn} \sum_{i \in y_{nc}} q(y_i = 0 | s) \log(p(y_i = 0 | s)) \quad (3)$$

Where  $q(y_i = i | s)$  is the ground-truth mask of the sample  $S$ .  $Y_i = 0, 1$ ,  $p(y | s)$  is the probability map produced by the function softmax,  $w$  represents the balancing parameter  $w(k) = S/KS(k)$ .  $S$  is the samples number in the training set, and  $S(K)$  represents the samples number in the class  $K$ . However, cross-entropy relies heavily on the accuracy of the annotation. When the data is mislabeled,  $q(k | s)$  will not be able to represent the true class distribution, which will lead the cross-entropy  $p(k | x)$  to learn this incorrect distribution type. To deal with this problem, we use reverse learning to know which classes the input  $x$  does not belong to. Inverse cross-entropy is defined as follows:

$$L_{ICE} = \sum_{i=1}^k p(y_i = 1 | s) \log q(y_i = 1 | s) \quad (4)$$

The weighted combination of entropies ( $L_{ICE}$  and  $L_{BCE}$ ) in the loss function allowed a good convergence of the gradient and relevant learning. The  $L_{BCE}$  term efficiency exhibits in the distribution balance of infected and uninfected classes, while the  $L_{ICE}$  term strength appears in the resistance against noise caused by scanner settings. As a result, the balanced symmetric entropy function obtained high performance in the COVID-19 lung infections localization on the test data set.

V. EXPERIMENT

In this section, we explain the different experiments carried out to evaluate the performance of our proposed approach under different simulation scenarios. We start by describing the used dataset and experimental environment of the simulation. Subsequently, we perform quantitative study of the hyper-parameters to show their effects on the model learning. Finally, we present the different performance metrics used to assess the efficiency and robustness of our approach.

A. Datasets

In this work, we have applied our method on two Datasets.

Dataset-1: this dataset contains two versions [21]. The first version is published on April 2, 2020, comprising 100 CT images (of 40 COVID-19 patients) of size  $512 \times 512$  labeled by radiologists, these radiologists have defined three tags: pleural effusion in frosted glass (= 3), consolidation (= 2), and mask value (= 1). The second version is released on April 14, 2020, comprising 829 CT images (of 9 COVID-19 patients) sized  $630 \times 630$ . Radiologists have tagged 373 images with COVID-19 pulmonary symptoms and the rest of the images as normal cases.

Dataset-2: this dataset is publicly available, contains 20 CT volumes with more than 1800 slices collected from 40 different COVID patients [22]. Each CT slice is of size  $512 \times 512$  labeled by expert radiologists to mark regions for infections.

The first column images of Fig. 2 represent the original images, while the second column images represent their corresponding ground-truth masks. Two examples of images of normal people are shown in the third and fourth rows. The first and second rows include two COVID-19 images, where the COVID-19 regions are the white and gray regions of the ground-truth masks, while the healthy regions are the black pixels (note that if a person is in good health his ground-truth mask will be completely black).

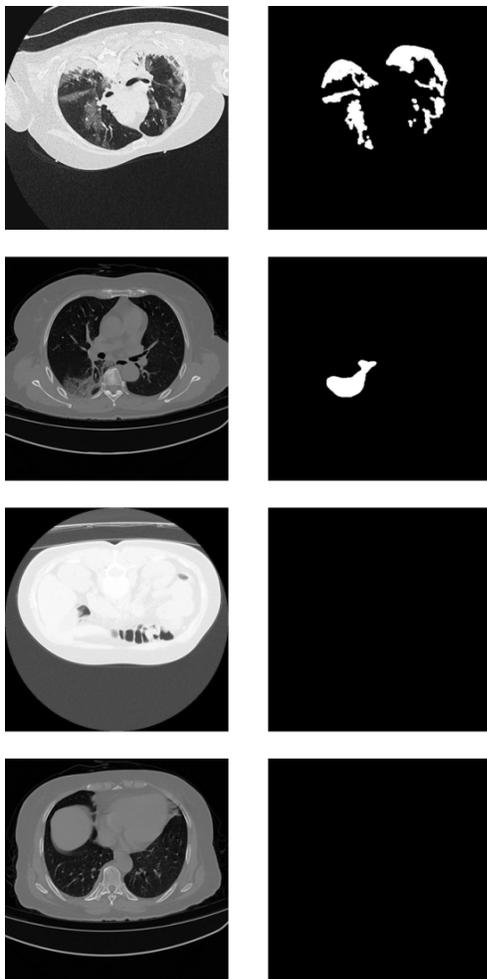


Fig. 2. Example of Images belong to Dataset-1 and Dataset-2.

In our simulations, we merged the two versions of dataset-1 to form a new dataset, while keeping 60% of these images for the training, 20% for the validation, and 20% for the test. We implemented our method in the Keras simulation environment with the TensorFlow back-end using the Python 2.7 programming language. Simulations were run on an infrastructure equipped with a Tesla P-100 GPU card, and 16 GB RAM memory.

### B. Hyper-parameters Setting

In deep learning, hyper-parameters of the deep neural network crucially influence the performance of the network. In this part, we carried out several experiments to choose the best values of the Hyper-parameters allowing improving the performances of our method. In this regard, we fixed the number of epochs and the batch size respectively at 100 and 64. We simulated and compared the different precision obtained values with different optimizers, different learning rates, and different values of  $\lambda_1, \lambda_2$  of the  $L_B$  loss function.

TABLE III. THE CHOICE OF THE BEST COMBINATION FOR INPUT HYPERPARAMETERS, BEST ACCURACY SHOWN IN BOLD FONT

Optimizer	Learning-rate	Loss (LB)	Accuracy
ADAM SGD Adadelta	Lr = 0.0001	$\lambda_1 = 0.5, \lambda_2 = 0.5$	<b>0.9445</b> 0.8921 0.9032
ADAM	Lr = 0.001 Lr = 0.00001 <b>Lr = 0.000001</b>	$\lambda_1 = 0.5, \lambda_2 = 0.5$	0.944 0.924 <b>0.964</b>
ADAM	Lr=0.000001	$\lambda_1 = 0.4, \lambda_2 = 0.2$ $\lambda_1 = 0.5, \lambda_2 = 0.5$ <b><math>\lambda_1 = 0.3, \lambda_2 = 0.5</math></b>	0.957 0.964 <b>0.991</b>

From Table III, we observe that the values of the hyper-parameters  $\lambda_1 = 0.3, \lambda_2 = 0.5$ , and the use of the ADAM [23] optimizer with a learning rate equal to 0.000001 showed the best performance in term of accuracy.

### C. Performance Metrics Evaluation

To assess the efficiency and robustness of our proposed approach, we use the following performance metrics: Accuracy [24], Sensitivity [25], Matthews Correlation Coefficient [25], and Dice [26]. By definition, higher values of these metrics imply a better segmentation quality. The mathematical formulas of these metrics are respectively expressed below:

$$Accuracy = \frac{TP + TN}{TP + TN + FN + FP} \quad (5)$$

$$Sensitivity = \frac{TP}{TP + FN} \quad (6)$$

$$MCC = \frac{TP * TN - FP * FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \quad (7)$$

$$Dice = \frac{2 * TP}{2 * (TP + FN + FP)} \quad (8)$$

## VI. RESULTS

In this section, we evaluate the effectiveness of the proposed method by performing qualitative and quantitative studies on an open-source benchmark. The obtained results of our method are compared with those of existing methods in the state of the art.

A. Ablation Study

To show the importance of each component of our approach, we did an ablation study on the dataset-1 by evaluating the following performance metrics; Accuracy, Dice, Sensitivity, and Precision. The study of ablations was subdivided into three possible cases. In the first case, we added the preprocessing block without using the loss function ( $L_B$ ) during the learning phase, whereas in the second case, in the learning phase we introduced the loss function without using the preprocessing block. In the latter case, we used the preprocessing block and the loss function ( $L_B$ ) (see Fig. 3). From these simulation cases, we notice that the data preprocessing step has a remarkable effect on the model performances. The first case shows the overfitting phenomenon that occurred when it exceeds epoch 40, and which led to a degradation in the performance of the model (see Fig. 4). The results of this study are illustrated in Table IV.

On the other hand, the result of the third case shows that the use of the loss function with the preprocessing block helps to avoid the overfitting problem, and consequently, to improve the performances of the model and to have good results. Fig. 5 and Fig. 6 show the qualitative results of our method on different test samples, the first column (1) represents the lung images, the second column (2) shows the ground-truth mask, and the last column (3) corresponds to the predicted lung infection mask of COVID-19. These qualitative results prove the robustness and the efficiency of our diagnostic method of the region infected by COVID-19.

TABLE IV. ABLATION STUDY ON THE DATASET-1

Cases	Accuracy	Dice	MCC	Sensitivity
Preprocessing, BinaryCrossEntropy	0.890	0.514	0.432	0.741
$L_B$ , without Preprocessing,	0.934	0.632	0.590	0.842
<b>Preprocessing, <math>L_B</math></b>	<b>0.991</b>	<b>0.833</b>	<b>0.851</b>	<b>0.982</b>

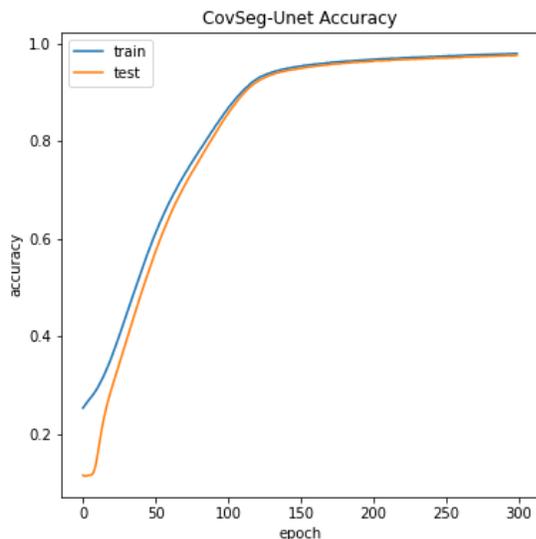


Fig. 3. Accuracy of Simulation with Data Preprocessing Step.

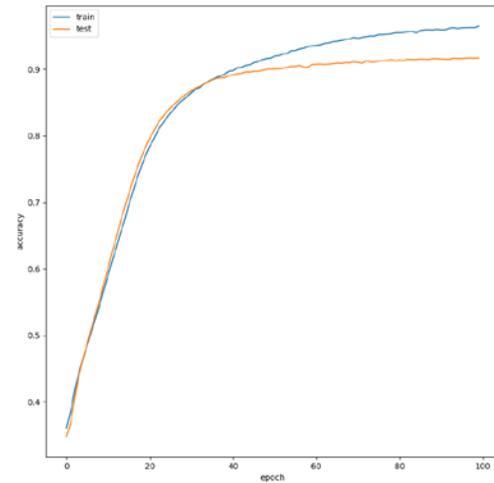


Fig. 4. Accuracy of Simulation without Data Preprocessing Step.

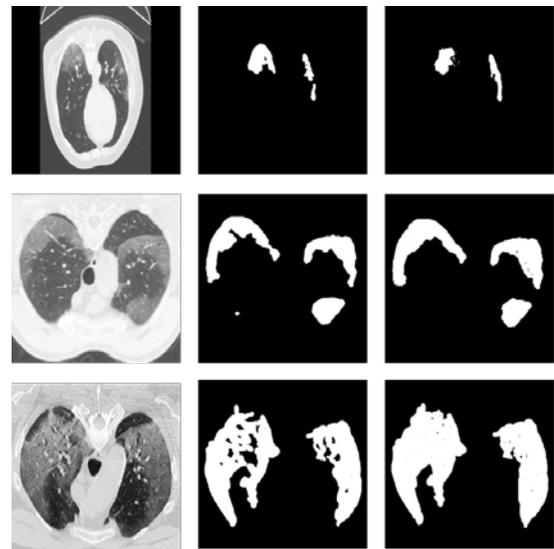


Fig. 5. Results Examples obtained by CovSeg-Unet on Dataset-1.

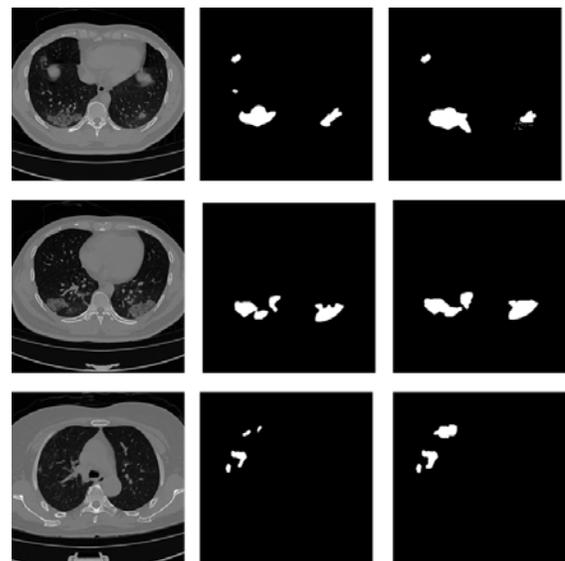


Fig. 6. Results Examples obtained by CovSeg-Unet on Dataset-2.

### B. Comparison with Baseline Methods

In this part, we compare our segmentation method of CT images with the reference segmentation methods such as U-Net basic [10], DenseUNet [27], Attention U-Net [13], and UNet++ [11]. Table V represents obtained results by the different methods on dataset-1 and dataset-2. As results, the proposed method outperforms other methods in terms of performance measures. Dice, Sensitivity, and Accuracy metrics reach 83.3%, 98.2% and 99.1% respectively on dataset-1, and 83.4% 89.1% 98.3% on dataset-2.

TABLE V. PERFORMANCES COMPARISON AGAINST BASELINE ARCHITECTURES ON DATASET-1 AND DATASET-2

	Methods	Dice	Sensitivity	Accuracy
Dataset-1	U-Net [10]	0.708	0.678	0.865
	DenseUNet [27]	0.660	0.607	0.651
	Attention U-Net [13]	0.560	0.623	0.632
	UNet++ [11]	0.815	0.857	0.903
	<b>CovSeg-Unet</b>	<b>0.833</b>	<b>0.982</b>	<b>0.991</b>
Dataset-2	UNet [10]	0.712	0.665	0.747
	DenseUNet [27]	0.610	0.607	0.715
	Attention U-Net [13]	0.631	0.723	0.890
	UNet++ [11]	0.815	0.887	0.968
	<b>CovSeg-Unet</b>	<b>0.834</b>	<b>0.891</b>	<b>0.983</b>

### C. Comparison with other Methods

Many studies have been done to diagnose COVID-19. To prove the robustness of our method, we carried out a comparative study with different approaches such as Inf-Net [14] and Automatic [17]. We simulated these approaches using their open-source implementation. The quantitative results obtained by the different methods are shown in Table VI. Dice, Sensitivity, and Accuracy are the performance metrics to be evaluated in this benchmarking study. From Table VI we notice that the CovSeg-Unet method reaches an Accuracy = 0.991 on Datasets-1, and an Accuracy = 0.983 on Datasets-2. The obtained values in these simulations prove that the CovSeg-Unet method outperforms other approaches in terms of Dice, Sensitivity, and Accuracy.

TABLE VI. PERFORMANCES COMPARISON AGAINST EXISTING APPROACHES ON DATASETS-2.

Methods	Dice	Sensitivity	Accuracy
Inf-Net [14]	0.579	0.877	—
Automatic [17]	0.714	0.733	0.739
<b>CovSeg-Unet</b>	<b>0.834</b>	<b>0.891</b>	<b>0.983</b>

## VII. CONCLUSION

In this work, we address a more difficult task in the segmentation of limited and unbalanced biomedical images. To cope with this task, we have proposed an end-to-end architecture similar to U-Net, the proposed method network learns the discriminating features of lung infections from CT images to establish an image-to-image mapping relationship. We used the ResNet50 architecture to preserve local information and avoid the issue of fading gradients. To improve the learning of the discriminating features of the network, we introduced a preprocessing block to remove noise and unnecessary information which blows up the performance of the network. In order to strengthen the model to be learned

from non-equilibrium data, we have proposed a loss function  $L_B$ . Experimental results on two datasets demonstrated the effectiveness of the CovSeg-Unet method in locating COVID-19 infected regions. The quantitative and qualitative results obtained by comparing CovSeg-Unet method with the others methods prove the efficiency of our method, which can be a real solution to detect, diagnose and locate regions infected with COVID-19.

### REFERENCES

- [1] C. Huang, Y. Wang, X. Li, L. Ren, J. Zhao, Y. Hu, L. Zhang, G. Fan, J. Xu, X. Gu et al., "Clinical features of patients infected with 2019 novel coronavirus in wuhan, china," *The lancet*, vol. 395, no. 10223, pp. 497-506, 2020.
- [2] N. Zhu, D. Zhang, W. Wang, X. Li, B. Yang, J. Song, X. Zhao, B. Huang, W. Shi, R. Lu et al., "A novel coronavirus from patients with pneumonia in china, 2019," *New England journal of medicine*, 2020.
- [3] J. T. Wu, K. Leung, and G. M. Leung, "Nowcasting and forecasting the potential domestic and international spread of the 2019-ncov outbreak originating in wuhan, china: a modelling study," *The Lancet*, vol. 395, no. 10225, pp. 689-697, 2020.
- [4] T. Liang, H. Cai, Y. Chen, Z. Chen, Q. Fang, W. Han et al., "Handbook of covid-19 prevention and treatment. 2020," *The First Affiliated Hospital, Zhejiang University School of Medicine. Compiled According to Clinical Experience*, 2020.
- [5] F. Pan, T. Ye, P. Sun et al., "Time course of lung changes on chest ct during recovery from 2019 novel coronavirus (covid-19) pneumonia [e-pub ahead of print], *radiology*," 2020.
- [6] S. Salehi, A. Abedi, S. Balakrishnan, and A. Gholamrezanezhad, "Coronavirus disease 2019 (covid-19): a systematic review of imaging findings in 919 patients," *American Journal of Roentgenology*, vol. 215, no. 1, pp. 87-93, 2020.
- [7] J. P. Kanne, "Chest ct findings in 2019 novel coronavirus (2019-ncov) infections from wuhan, china: key points for the radiologist," 2020.
- [8] F. Shan, Y. Gao, J. Wang, W. Shi, N. Shi, M. Han, Z. Xue, D. Shen, and Y. Shi, "Lung infection quantification of covid-19 in ct images with deep learning," *arXiv preprint arXiv:2003.04655*, 2020.
- [9] Ö. Çiçek, Abdulkadir, S. S. Lienkamp, T. Brox, and O. Ronneberger, "3d u-net: learning dense volumetric segmentation from sparse annotation," in *International conference on medical image computing and computer-assisted intervention*. Springer, 2016, pp. 424-432.
- [10] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*. Springer, 2015, pp. 234-241.
- [11] Z. Zhou, M. M. R. Siddiquee, N. Tajbakhsh, and J. Liang, "Unet++: A nested u-net architecture for medical image segmentation," in *Deep learning in medical image analysis and multimodal learning for clinical decision support*. Springer, 2018, pp. 3-11.
- [12] F. Milletari, N. Navab, and S.-A. Ahmadi, "V-net: Fully convolutional neural networks for volumetric medical image segmentation," in *2016 fourth international conference on 3D vision (3DV)*. IEEE, 2016, pp. 565-571.
- [13] O. Oktay, J. Schlemper, L. L. Folgoc, M. Lee, M. Heinrich, K. Misawa, K. Mori, S. McDonagh, N. Y. Hammerla, B. Kainz et al., "Attention u-net: Learning where to look for the pancreas," *arXiv preprint arXiv:1804.03999*, 2018.
- [14] D.-P. Fan, T. Zhou, G.-P. Ji, Y. Zhou, G. Chen, H. Fu, J. Shen, and L. Shao, "Inf-net: Automatic covid-19 lung infection segmentation from ct images," *IEEE Transactions on Medical Imaging*, vol. 39, no. 8, pp. 2626-2637, 2020.
- [15] M. A. Al-antari, C.-H. Hua, J. Bang, and S. Lee, "Fast deep learning computer-aided diagnosis of covid-19 based on digital chest x-ray images," *Applied Intelligence*, vol. 51, no. 5, pp. 2890-2907, 2021.
- [16] M. Shorfuzzaman and M. S. Hossain, "Metacovid: A siamese neural network framework with contrastive loss for n-shot diagnosis of covid-19 patients," *Pattern recognition*, vol. 113, p. 107700, 2021.

- [17] Z. Wang, Y. Xiao, Y. Li, J. Zhang, F. Lu, M. Hou, and X. Liu, "Automatically discriminating and localizing covid-19 from community acquired pneumonia on chest x-rays," *Pattern recognition*, vol. 110, p.107613, 2021.
- [18] K. Gong, D. Wu, C. D. Arru, F. Homayounieh, N. Neumark, J. Guan, V. Buch, K. Kim, B. C. Bizzo, H. Ren et al., "A multi-center study of covid-19 patient prognosis using deep learning-based ct image analysis and electronic health records," *European journal of radiology*, vol. 139, p. 109583, 2021.
- [19] O. N. Oyelade, A. E. Ezugwu, and H. Chiroma, "Covframenet: An enhanced deep learning framework for covid-19 detection," *IEEE Access*, 2021.
- [20] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770-778.
- [21] "covid-19 ct segmentation dataset. available: <https://medical-segmentation.com/covid19/>,"2020.
- [22] M. Jun, G. Cheng, W. Yixin, A. Xingle, G. Jiantao, Y. Ziqi et al., "Covid-19 ct lung and infection segmentation dataset 2020," 2020.
- [23] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.
- [24] E. Fernandez-Moral, R. Martins, D. Wolf, and P. Rives, "A new metric for evaluating semantic segmentation: leveraging global and contour accuracy," in *2018 IEEE intelligent vehicles symposium (iv)*. IEEE, 2018, pp. 1051-1056.
- [25] S. Boughorbel, F. Jarray, and M. El-Anbari, "Optimal classifier for imbalanced data using matthews correlation Coefficient metric," *PloS one*, vol. 12, no. 6, p. e0177678, 2017.
- [26] T. Trongtirakul, A. Oulefki, S. Agaian, and W. Chiracharit, "Enhancement and segmentation of breast thermograms," in *Mobile Multimedia/Image Processing, Security, and Applications 2020*, vol. 11399. International Society for Optics and Photonics, 2020, p.113990F.
- [27] X. Li, H. Chen, X. Qi, Q. Dou, C.-W. Fu, and P.-A. Heng, "H-denseunet: hybrid densely connected unet for liver and tumor segmentation from ct volumes," *IEEE transactions on medical imaging*, vol. 37, no. 12, pp. 2663-2674, 2018.