



An Encoder–Decoder-Based Method for Segmentation of COVID-19 Lung Infection in CT Images

Omar Elharrouss¹ · Nandhini Subramanian¹ · Somaya Al-Maadeed¹

Received: 12 June 2020 / Accepted: 2 September 2021

© The Author(s), under exclusive licence to Springer Nature Singapore Pte Ltd 2021

Abstract

The novelty of the COVID-19 Disease and the speed of spread, created colossal chaotic, impulse all the worldwide researchers to exploit all resources and capabilities to understand and analyze characteristics of the coronavirus in terms of spread ways and virus incubation time. For that, the existing medical features such as CT-scan and X-ray images are used. For example, CT-scan images can be used for the detection of lung infection. However, the quality of these images and infection characteristics limit the effectiveness of these features. Using artificial intelligence (AI) tools and computer vision algorithms, the accuracy of detection can be more accurate and can help to overcome these issues. In this paper, we propose a multi-task deep-learning-based method for lung infection segmentation on CT-scan images. Our proposed method starts by segmenting the lung regions that may be infected. Then, segmenting the infections in these regions. In addition, to perform a multi-class segmentation the proposed model is trained using the two-stream inputs. The multi-task learning used in this paper allows us to overcome the shortage of labeled data. In addition, the multi-input stream allows the model to learn from many features that can improve the results. To evaluate the proposed method, many metrics have been used including Sorensen–Dice similarity, Sensitivity, Specificity, Precision, and MAE metrics. As a result of experiments, the proposed method can segment lung infections with high performance even with the shortage of data and labeled images. In addition, comparing with the state-of-the-art method our method achieves good performance results. For example, the proposed method reached 78.6% for Dice, 71.1% for Sensitivity metric, 99.3% for Specificity 85.6% for Precision, and 0.062 for Mean Average Error metric, which demonstrates the effectiveness of the proposed method for lung infection segmentation.

Keywords Lung infection segmentation · COVID-19 · CT-scan image · Encoder–decoder network

Introduction

At the onset of coronavirus COVID19 pandemic in 2019, deaths due to this virus were a common sight. Wuhan City was confirmed to be the epicenter of the coronavirus which

then propagated to every other continent in the world. Thousands of people lost their lives [1]. Sever acute respiratory syndrome or with the acronym SARS-CoV-2 was named as the infectious disease caused by this virus. Similarity in appearance to the solar rings when viewed from electron microscopic analysis led to naming it corona virus [2]. The novelty of this coronavirus and the speed of spread created a rapid action for worldwide researchers to exploit all resources and capabilities to understand and analyze characteristics of coronavirus in terms of spread and virus incubation time as well as exploring new techniques. In addition, temporal procedures were placed to stop the speed of spread. The absence of vaccine prompted the governments to mitigate the impact of pandemic by imposing lockdowns and social distance measures. However, these policies were not efficient due to the delay in decision making by certain countries. To contribute to the implementation of these policies, scientists and technology experts attempted to find solutions

This article is part of the topical collection “Computer Aided Methods to Combat COVID-19 Pandemic” guest edited by David Clifton, Matthew Brown, Yuan-Ting Zhang and Tapabrata Chakraborty.

✉ Omar Elharrouss
elharrouss.omar@gmail.com

Nandhini Subramanian
nandhini.reborn@gmail.com

Somaya Al-Maadeed
s_alali@qu.edu.qa

¹ Department of Computer Science and Engineering, Qatar University, Doha, Qatar

to stop the speed of spread. An example of it being robots used to mitigate the contact between patients and hospital employees [2]. Drones were used for social distance monitoring and disinfecting public spaces. Scientists, specifically health scientist, placed medicines and cures for this virus as their priority research [3, 4]. Researchers in other domains such as computer sciences worked on techniques to detect infected persons using medical features such as CT and X-ray images. Artificial intelligence (AI) tools and computer vision algorithms enhanced the accuracy of detection with high precision. This aided in the early detection of disease [5]. This was chosen as an alternative to the laboratory analysis on blood samples.

Image processing, as well as computer vision combined with AI, is a multidisciplinary domain that can be used in different domains including medical, astronomy, agriculture [8–10, 12, 18]. In the medical field, medical imaging has been used for diagnosing diseases using X-ray and CT-scan images, also for surgery and therapy [5]. Good progress in this field has been reached due to the introduction of many techniques such as machine learning and deep learning algorithms. This improvement leads computer vision scientists to contribute to finding solutions for rapid diagnostics, prevention, and control. For the same purpose, several approaches have been proposed within a short time.

The detection of infection is the first step for the diagnostic of a disease. Using CT-scan images, the appearances of the infected regions are different from the normal regions, so the detection of and the extraction of this region automatically can help the doctor for diagnose within a short time [18]. For the same purpose, a deep learning method for segmentation lung infection for COVID-19 on CT-scan images is proposed. Before starting the learning process, each image has been composed to structure and texture components. The structure represents the homogeneous part of the images, where the texture component represented the texture of the image. The structure and texture features are introduced to the encoder–decoder model first for segmenting the regions of interest that can be affected. Then, segmenting the specific infected parts on the results of the first segmentation.

The remainder of this paper is organized as follows. The literature overview is presented in Sect. “[Related Works](#)”. The proposed method is presented in Sect. “[Proposed Method](#)”. Experiments performed to validate the proposed method are discussed in Sect. “[Experimental Results](#)”. The conclusion is provided in Sect. “[Conclusion](#)”.

Related Works

Recently, medical imaging has gained attention due to its importance to diagnose, monitor, and treat several medical problems. Radiography, a medical imaging technique, uses

[13–15], CT-scan [16–18], and gamma rays to create images of the body that requires internal viewing. analyzing images using different computer vision algorithms provides an alternative for rapid diagnostics and control of many diseases. For COVID-19, image analysis offers a good solution for early detection due to the complexity of laboratory analysis and the importance of early detection that can save lives. For the same purpose, many approaches have been proposed to COVID-19 detection and control using X-ray and CT-scans images of the lung. In addition, the use of artificial intelligence makes the precision and the performance of these results convincing. authors in [19] presented COVID-Net architecture that is based on DenseNet to diagnose the COVID-19 infections from X-rays and CT-scans images to decrease the turnaround time of the doctors and check more patients at that point of time. Using the proposed architecture, the infected regions are detected and marked.

Working on similar radiographic images, authors in [20] developed a classification and segmentation system for a real-time diagnostic of COVID-19. The presented model combines convolutional neural networks and encoder–decoder networks trained on CT-scan images. The proposed approach succeeds to detect the infected regions with an accuracy of 92%. In the same context, a learning architecture named Detail-Oriented Capsule Networks (DECAPS) has been proposed in [21]. To increase model stability, the authors use an Inverted Dynamic Routing mechanism model implementation. The proposed approach detected the infected regions without segmenting the regions edges. The same technique is used in [22] based on SqueezeNet architecture. The obtained results achieve 89% for performance accuracy, wherein [21] the accuracy reaches 87%. To demonstrate the efficiency and the automatic testes of COVID-19 infection, the authors in [23] proposed a machine learning scheme method for CT-scan image analysis for COVID-19 patients. Based on the FFT-Gabor scheme, the proposed method for predicting the state of the patient in real time. The performance accuracy achieves an average of 95.37% which represents a convincing rate.

For developing an accurate method for real-time diagnosing of COVID-19 using CT-scan images, deep learning models require a large-scale data set to train these models, which might be difficult to obtain at this moment. Hence, authors in [24], build a public CT-scan data set people detected positive for COVID-19. In addition, a deep-learning-based method has been proposed for classifying the COVID or NON-COVID images. The approach showed an accuracy of 72%, which means that it is not a very accurate method for COVID-19 testing. A multitask deep learning model to identify COVID-19 patients and segmentation of infected regions from chest CT images has been proposed in [25]. The authors used an encoder–decoder model for segmentation and perceptron for classification. The proposed method

Table 1 Categories of COVID-19 methods

Task	Method	Architecture
Classification	Ozturk et al. [8]	DarkNet, CNN
	Minaee et al. [9]	ResNet18, CNN
	Apostolopoulos et al. [10]	MobileNet, CNN
	Apostolopoulos et al. [11]	ResNet18, VGG19, Inception, Xception
	Abbas et al. [12]	ResNet18, CNN
	Adhikari et al. [19]	DenseNet-based CNN
	Mobiny et al. [21]	CNN
	Polsinelli et al. [22]	SqueezeNet-based CNN
	Al-karawi et al. [23]	Machine learning, FFT-Gabor
	He et al. [25]	CNN
	Talha et al. [26]	CNN
Class and segmentation	Wu et al. [20]	CNN, encoder–decoder
	Amyar et al. [25]	CNN, encoder–decoder
Segmentation	Fan et al. [27]	CNN, partial decoder

has the same technic as the method proposed in [20] but using different architectures. The obtained performance accuracy of the proposed method reaches 86%. With an accuracy of 89%, EfficientNet a neural network architecture is proposed for the detection of COVID-19 patients using CT-scan images [26]. For segmenting the infected region in a CT-scan image for COVID-19 patient identification, the authors in [27] proposed a convolutional neural network (CNN) model with a partial decoder. The proposed model performance achieves an accuracy of 73% for detecting the infected regions in a CT-scan image. The authors also train other architectures including Unet [28], Unet++ [29], Attention-Unet [30], Gated-Unet [31], and Dense-Unet [32] on the same data set.

All the proposed method attempts to develop a timely and effective method for testing the coronavirus patients. The proposed methods can be classified into two general categories: classification methods [19, 21–23] and segmentation [27] of infected region methods. Some of the presented approaches work on two tasks like [20, 25]. Table 1 presents the proposed method for each category as well as the architectures used in each one of them.

Proposed Method

In this section, the proposed approach for lung infection segmentation is provided. The method starts by splitting the texture and structure component of the image, before starting the training of the proposed model for regions of interest extraction. The extraction of these regions is the operation of separation or segmentation of the regions that can contain the infection which is the intern region of the lung. After the extraction of regions of interest, which is performed using the proposed encoder–decoder network,

the segmentation model that uses the output of the previous operation, is the same is used this time for detecting and segmenting the infected part of the lung. A description is provided for each step including structure–texture decomposition, region of interest extraction, and lung infection segmentation.

Structure and Texture Component Extraction

Each image can contain information that includes the form of the object (or the homogenous part) in the image as well as a texture that also contains information that can be useful in some computer vision tasks. Here, the texture component is used as input of the proposed encoder–decoder neural networks. for that, preprocessing is performed to extract the texture component. The structure–texture decomposition method proposed in [36] is adopted using the interval gradient, for adaptive gradient smoothing. Given an image f , It is a technique that splits the image into $S + T$ ($f = S + T$) of a bounded variation component (Structure) and a component that contains the oscillating part (Texture/Noise) of the image [37, 38]. We applied this technique on the CT-scan images. Then, we use the texture and structure components as an input of the proposed encoder–decoder model. The extraction of the structure component uses gradient rescaling with interval gradients followed by a color handling operation.

To produce a texture-free signal from an input signal I , the gradients within texture regions should be suppressed. Furthermore, the signal should be either increasing or decreasing for all local windows Ω_r . With these objectives, we use the following equation to rescale the gradients of the input signal with the corresponding interval gradients:

$$(\nabla' I)_p = \begin{cases} (\nabla I)_p \cdot w_p & \text{if } \text{sign}((\nabla I)_p) = \text{sign}((\nabla_{\Omega} I)_p) \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

where $(\nabla' I)_p$ represents the rescaled gradient, and w_p is the rescaling weight:

$$w_p = \min\left(1, \frac{|(\nabla_{\Omega} I)_p| + \epsilon_s}{|(\nabla I)_p| + \epsilon_s}\right) \quad (2)$$

where ϵ_s is a small constant to prevent numerical instability. Too small values of ϵ_s would make the algorithm sensitive to noise, introducing unwanted artifacts to filtering results. The sensitivity to noise can be reduced by increasing s but textures may not be completely filtered if ϵ_s is too big. We set $\epsilon_s = 10^{-4}$ in our implementation.

For filtering color images, we use the gradient sums of color channels in the gradient rescaling step (Eqs. (1) and (2)), that is:

$$w_p = \min\left(1, \frac{\sum_{c \in \{r, g, b\}} |(\nabla_{\Omega} I^c)_p| + \epsilon_s}{\sum_{c \in \{r, g, b\}} |(\nabla I^c)_p| + \epsilon_s}\right) \quad (3)$$

Figure 3 shows filtering examples to demonstrate the results of structure and texture extraction using the same parameters $\epsilon \in [0.0001, 0.0009]$.

Encoder–Decoder Architecture

Dense pixelwise classification is a required operation for semantic labeling of images. To achieve an effective semantic segmentation, many architectures have been proposed including FCN [39], SegNet [40] architectures. In this paper, we based on the SegNet model which is an encoder–decoder architecture that provides an image output with the same image input size.

The encoder part of SegNet is based on the VGG-16 [41] convolutional layers that composed 5 blocks, where each one contains 2–3 convolutional layers with 3×3 kernels, 1 padding, ReLU, and a batch normalization (BN) [42]. The convolution block is followed by a max-pooling layer of size 2×2 . At the end of the encoder, each feature map has $H/32$, $W/32$, where the original image resolution is $H \times W$.

The decoder part proceeds the operation of upsampling and classification. The main role of this part is the learning of the method of spatial resolution restoration by transforming of encoder features maps into the final labels. The decoder structure is symmetrical with encoder, while the pooling layers in the encoder part are replaced by unpooling layers in the decoder part. The role of convolutional blocks after unpooling layer is to densify the sparse feature maps. This procedure is looped until the feature map reaches the resolution of the input image.

The final layer in SegNet as well as in most of the proposed architecture in the same context, SoftMax is used to compute the multinomial loss:

$$\text{loss} = \frac{1}{N} \sum_{N=1}^N \sum_{i=1}^k y_i^i \log\left(\frac{\exp(p_j^i)}{\sum_{l=1}^k \exp(p_l^i)}\right) \quad (4)$$

where N is the number of pixels in the input image, k the number of classes and, for a specified pixel i ; y_i denote its label and PPP the prediction vector. This means that we only minimize the average pixelwise classification loss without any spatial regularization, as it will be learnt by the network during training.

In this paper, we followed the same scenarios, but with two streams as input. The encoder part of the proposed model contains two components. The encoder for the texture component and the encoder for the structure component. The encoder features map is formed by the concatenation of the two encoders. In this architecture, we have 5 encoder blocks. Each block of the encoder is composed of a *conv + BN + PReLU + pooling*, where the decoder blocks composed of *upsampling + Conv + BN + PReLU*. The filter size of each convolutional layer in the encoder part is in the range of 32, 64, 128, 256, Fig. 2 represents the proposed architecture.

The model is trained using *binary_crossentropy* as loss with a batch size of 16 examples and Adam as optimizer like illustrated in Table 2.

Lung Infection Segmentation

The current CT-scan data set of COVID-19 is very limited in terms of the number of labeled images. In addition, the manual segmentation of the infected regions of the lung is a challenging task due to the needs of an expert(doctor), as well as the time that this operation can take. To solve the current problem of limited data, we augment this data using rotation and translation of a large number of images from the labeled data. Using the presented encoder–decoder architecture, we first segment the region of interest or the region that can contain the infection. The inside part of the lung is segmented before being used in the stage of lung infection segmentation.

For segmenting the lung infected region using black and white colors, where the white color represents the infected region, we use the segmented region which is the output of the first step of segmentation, and the original image as input of the model in the second stage, as illustrated in Fig. 1. Multi-task segmentation is an effective solution regarding the limitation of the size of the data discussed above. In addition, the segmentation of the infected region

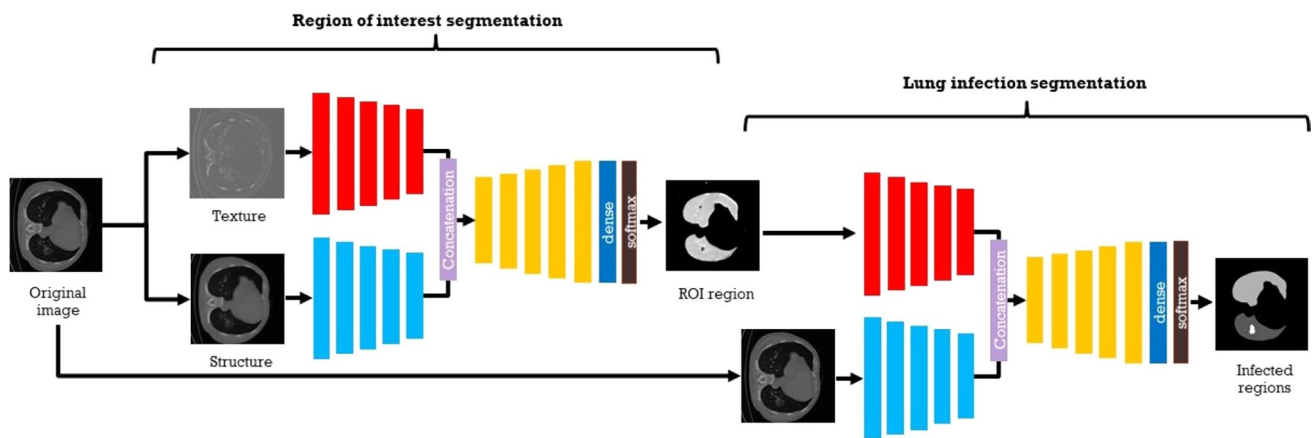


Fig. 1 Flowchart of the proposed method

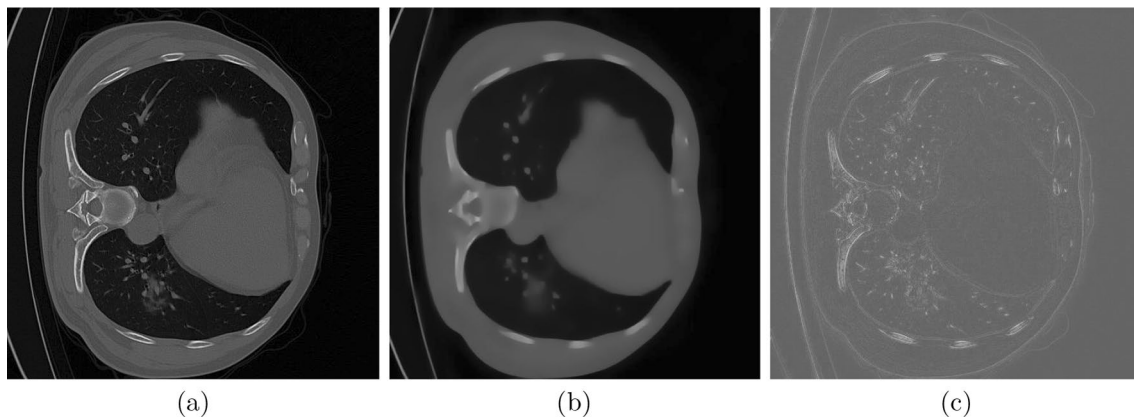


Fig. 2 Structure and texture extraction results: (a) original image, (b) Structure component, (c) Texture component

can be used for segmenting these regions using multiclass which represents the next task in this paper.

Multi-class Lung Infection Segmentation

The segmentation using multiclass or many colors to represent the lung infection can be more helpful for the diagnostic of COVID-19 also more practical. For that, using the previous segmentation of lung infection and the proposed deep learning model, multiclass segmentation is performed. Two-stream inputs of the deep learning model are represented by the lung infection segmentation results and the results of the first segmentation which is the region of interest. This technique allows learning on the specific region and can be more accurate according to the data limitation. Figure 3 illustrates the input and the output of the deep learning model in this stage for multiclass segmentation.

Experimental Results

In this section, we demonstrate the relevance of our proposed method by providing the experimental results. The evaluation has been performed on two segmentation categories including simple segmentation of infected regions with black and white, and multi-class labeling using colors. For the first category, we compare our results with a set of state-of-the-art methods, including Unet [28], Unet++ [29], Attention-Unet [30], Gated-Unet [31], Dense-Unet [32] and Semi-Inf-Net [27]. For multi-class labeling, the obtained results are compared with four state-of-the-art methods such as Semi-Inf-Net [27], multi-class U-Net [33], FCN8s [34] and DeepLabV3+ [35]. In addition, the results are visualized by presenting some segmented examples using the proposed method as well as state-of-the-art methods.

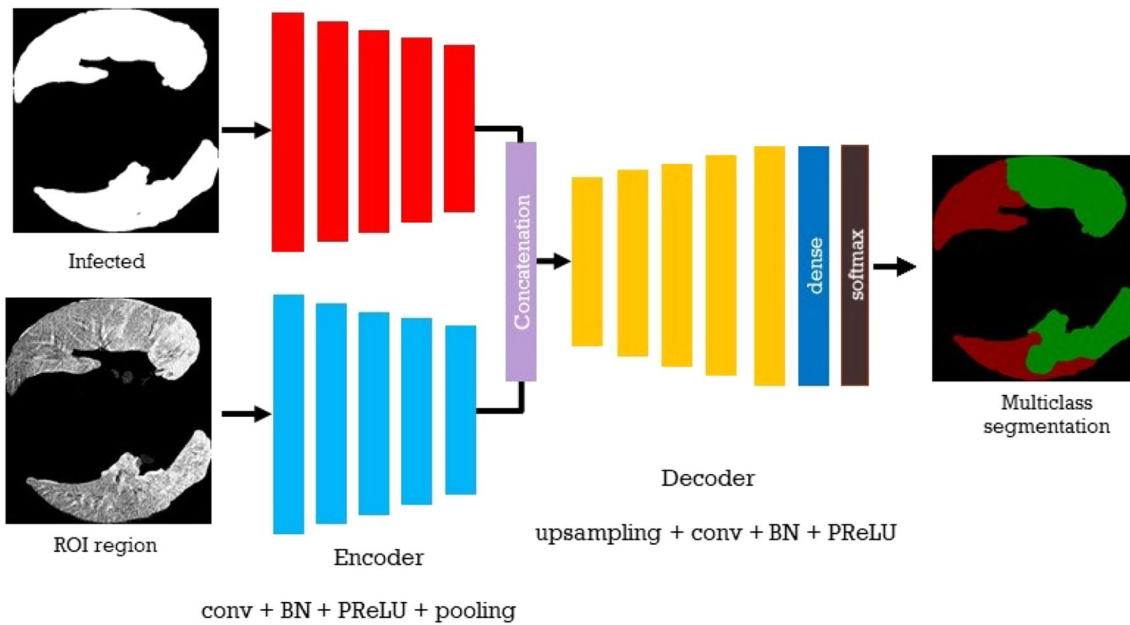


Fig. 3 Multi-class segmentation of the lung infected regions

Table 2 Hyperparameters of the proposed model

Method	Loss	Optimizer	epochs	Batch size
Proposed model	Binary_crossentropy	Adam	100	16

Data Sets

The only segmentation data set of CT-scan images available for COVID-19¹. The data set consists of 100 axial CT images for 20 COVID-19 patients, collected by the Italian Society of Medical and Interventional Radiology. The data set contains CT images labeled. There are two categories of labels, the first one labels the regions of interest, where the infection can be located. The other one is the specific infected regions labeled with two colored red and green colors. This part contains cropped images for training and testing. The training images are composed of 50 images, where the infections are labeled with one color (one class) and multi-class (2 colors). The test folder contains 48 images labeled with the same labels as the training. In this paper, we train our multitask model on the two categories of images. We train the first part of our model on the 2000 images for region segmentation. For infection segmentation, we use the labeled data and we augment the data by rescaling and rotating the images for getting more data.

¹ <https://medicalsegmentation.com/covid19/>.

Evaluation Metrics

To evaluate the segmentation results of the proposed method, a set of measures has been exploited including Sorensen–Dice similarity, Sensitivity, Specificity, Precision, and MAE used in [45–47]. For calculation of these three parameters, four measures are required, namely, true positive (TP), false positive (FP), true negative (TN), and false-negative (FN). The true positive (TP) represents the number of pixels being correctly identified. True negative (TN) describes the number of non-lung infection pixels being correctly identified as non-lung infection. False-positive (FP) denotes the number of non-lung infections being wrongly classified as lung infections, whereas false negative (FN) means the lung infection pixels being wrongly classified as non-lung infection.

Sorensen–Dice similarity: Let consider that A is segmented regions that we need to assess the quality, B is ground truth. The Sorensen–Dice similarity [45] is computed as follows:

$$Dice = \frac{2|A \cap B|}{|A| + |B|} \quad (5)$$

The value of the Sorensen–Dice similarity metric is between 0 and 1. The higher the Sorensen–Dice value, the better the segmentation result.

Sensitivity: Sensitivity, defined as the ratio of correctly identified lung infection to the total number of lung infection pixels, is computed as

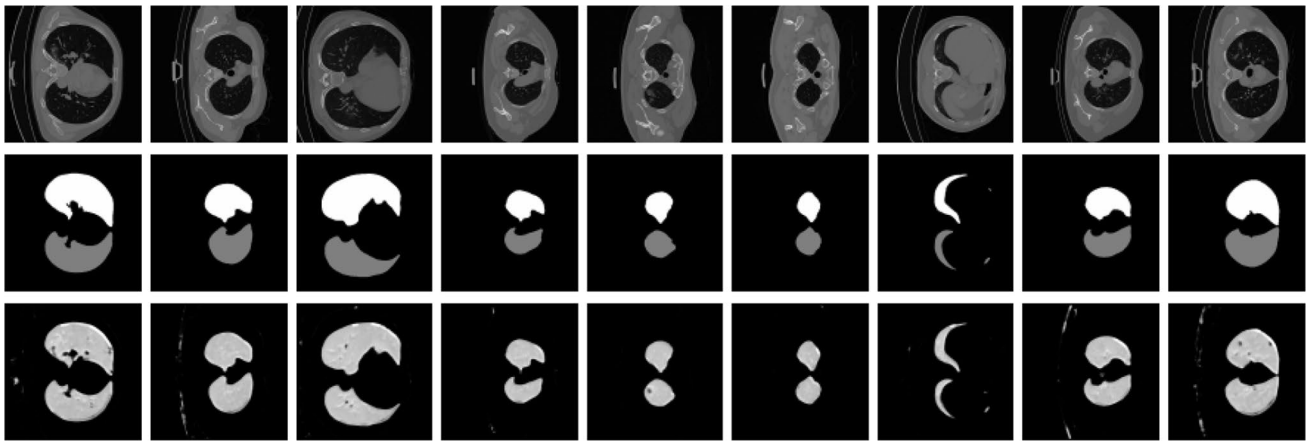


Fig. 4 Region of interest segmentation results. First row: original images. Second row: groundtruth. third row: segmentation using the proposed method

$$\text{Sensitivity} = \frac{TP}{TP + FN} \tag{6}$$

Specificity: Specificity, defined as the ratio of correctly detected non-lung infection to the total number of non-lung infection pixels, is measured as

$$\text{Specificity} = \frac{TN}{TN + FP} \tag{7}$$

Precision: Precision gives the percentage of unnecessary positives through the compared total number of positive pixels in the detected binary objects mask [47]

$$\text{Precision} = \frac{TP}{TP + FP} \tag{8}$$

Mean Absolute Error (MAE) This measures the pixelwise error between S_p and G , which is defined as:

$$\text{MAE} = \frac{1}{w \times h} \sum_i^w \sum_j^h |S(i,j) - G(i,j)|. \tag{9}$$

Discussion

The first step of our proposed model is to segment the regions that can contain the infection from lung images. This step is sued for segmenting the lung infection regions in the second part of the proposed method. For that, this step is an accurate preprocessing operation for segmenting the lung infection region after. Figure 4 illustrates some obtained results for the region of interest segmentation by presenting the original image with the ground truth as well as the obtained results. We can observe that the proposed model gives promising segmentation results. In addition,

comparing with the ground truth, we obtain a segmentation without false segmented regions. For example, even the variations between the shapes of segmented regions like in Fig. 4 first column, fifth column, and seventh column, the proposed method segment these regions with good quality and quantity comparing with the ground truth. For Fig. 4 first column, second column, third column, fourth column, and the last two columns the shape of the region of interests are close but they are different from the others, but the segmentation is effective and this is due to the use of two-stream using structure and texture component of each image which allow learning from different features.

To evaluate the performance of the proposed method for lung infection segmentation, Fig. 5 shows some examples of the obtained results. For the visualized results we can observe that the proposed model can detect the infection effectively with some erroneous that can be considered negligible. In addition, the segmentation results are close to the ground truth, as shown in the second row of Fig. 5. The infected parts in the images of the first column, the second, the third, the fourth and the fifth columns have different shapes. For example, we can find that the infected parts in the fourth column contain one small part and another big part, the same observation for all the other column images. However, the proposed method segment these parts in an effective way, and without post-processing that can remove some erroneous segmented pixels. The success of the proposed method to label the infections is owed to the proposed architecture that uses two-stream input which allows robust learning. In addition, the use of the results of the region of interest as input beside the original image gives the model a specification of the region that can contain the infection. The robustness of the presented approach is shown also in the presented metrics that demonstrate the performance evaluation including Dice, Sensitivity, specificity, precision, and

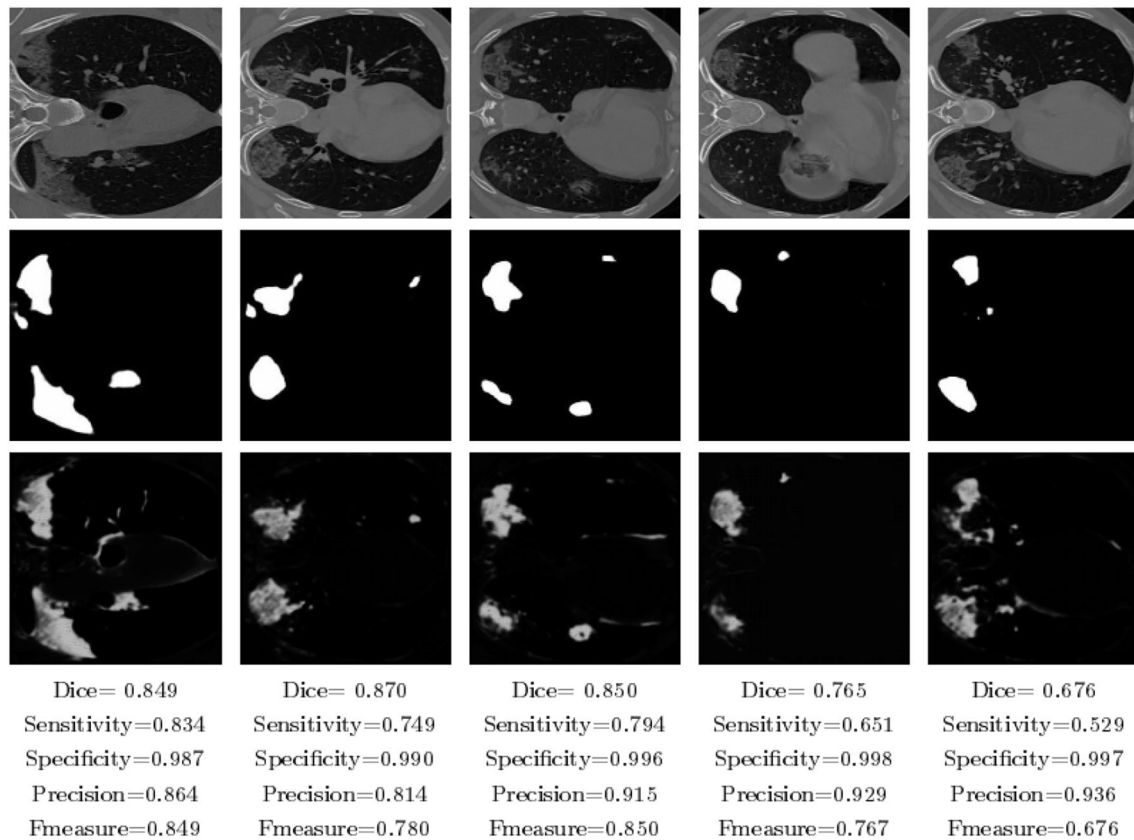


Fig. 5 Infection region segmentation. First row: original images. Second row: groundtruth. Third row: segmentation using the proposed method. Fourth row: evaluation results

F-measure. For example, for the first CT-scan image, our approach can segment with a high precision value achieves 84% for the F-measure metric which is a convincing value. The present results can be improved using a preprocessing on the results images like the morphology operations.

To assert the results represented by the binary image in Fig. 5, the qualitative and quantitative results obtained by each method are represented in Table 3. From these results, we can see that the proposed method gives high performance comparing with state-of-art methods. While the proposed method reached the best MAE, Precision, specificity, and dice values than the state-of-the-art methods. The specificity value reached using the proposed method is 99.3 which overperforms Semi-Inf-Net by 3%, U-Net by 13.5%, Attention-UNet by 7.2%, and Gated-UNet by 9.1%. The effectiveness of the proposed method comes from the use of multi-task learning and the use of two-stream input for our model.

The multi-class infection labeling results are also illustrated in Fig. 6. As shown in the figure the proposed method performs an accurate segmentation of the lung infection using multi-class labeling. The best result comes from the succession of tasks for performing the multi-class segmentation. The use of the results of unit-class segmentation with

the original image leads to a precise segmentation of the lung infection. In addition, the segmentation is performed on various images with different infected regions like in the first column images that contain a big infected part. The same observation is for the image in the seventh column. In addition, the shape of each infected part, like in the second and sixth column, can differentiate from small to big regions with can make any technique not able to segments these parts with a good performance. But for the obtained results using the proposed method, the performance of segmented regions using multi-labeling is in general good and the small part are segmented. The evaluation using different metric in Table 4 also demonstrate the advantage of the proposed method compared with the other existing methods. For example, For Dice and sensitivity and precision metrics, the proposed method reached the best results of 64% for dice, 63% for sensitivity, and 56.1% for precision. In addition, the proposed method reached the second-best results after the semi-inf-Net method for specificity and MAE metric by a value of 95.3% and 0.060, respectively. The semi-inf-Net method succeeds to obtain close results due to the multi-task learning model. While we can find that the difference in terms of MAE metric is 0.003, 1.4% for specificity. In

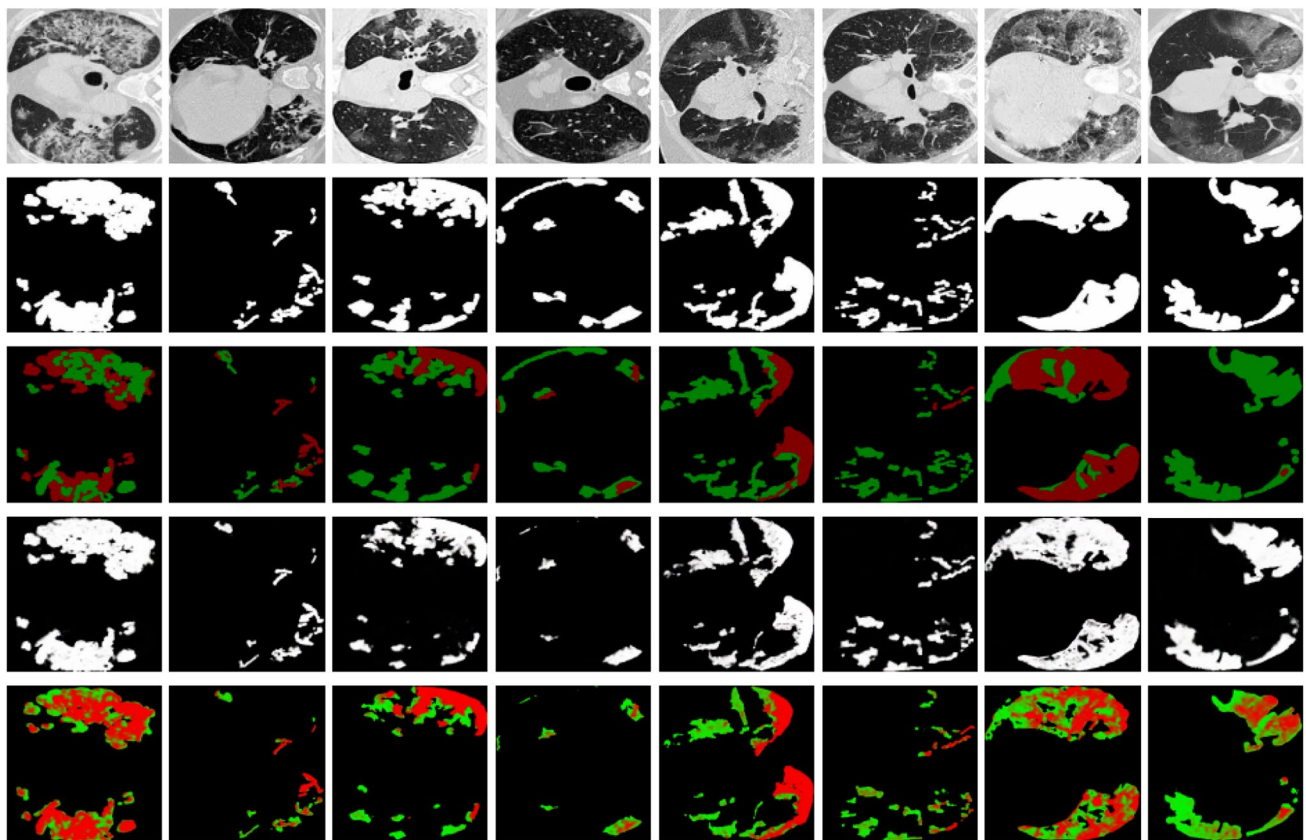


Fig. 6 Segmentation results using the proposed method. First row: original images. Second row: binary ground truth. Third row: multi-class ground truth. Fifth row: Our binary segmentation results. Sixth row: Our multi-class segmentation results

Table 3 Quantitative results of binary infection regions segmentation on COVID-SemiSeg data set

Method	Dice	Sensitivity	Specificity	Precision	MAE
U-Net [28]	0.439	0.534	0.858	-	0.286
Attention-UNet [30]	0.583	0.637	0.921	-	0.112
Gated-UNet [31]	0.623	0.658	0.926	-	0.102
Dense-UNet [32]	0.515	0.594	0.840	-	0.184
U-Net++ [29]	0.581	0.672	0.902	-	0.120
Semi-Inf-Net [27]	<i>0.739</i>	0.725	<i>0.960</i>	-	0.064
Proposed method	0.786	<i>0.711</i>	0.993	0.856	0.062

Bold, and italic indicates the first and the second best results, respectively

contrast to the other model that is used as it is lime UNet of FC8s models.

Ablation Study

In this subsection, we conduct several experiments to validate the performance of different components of the proposed method, including the results of binary infected results

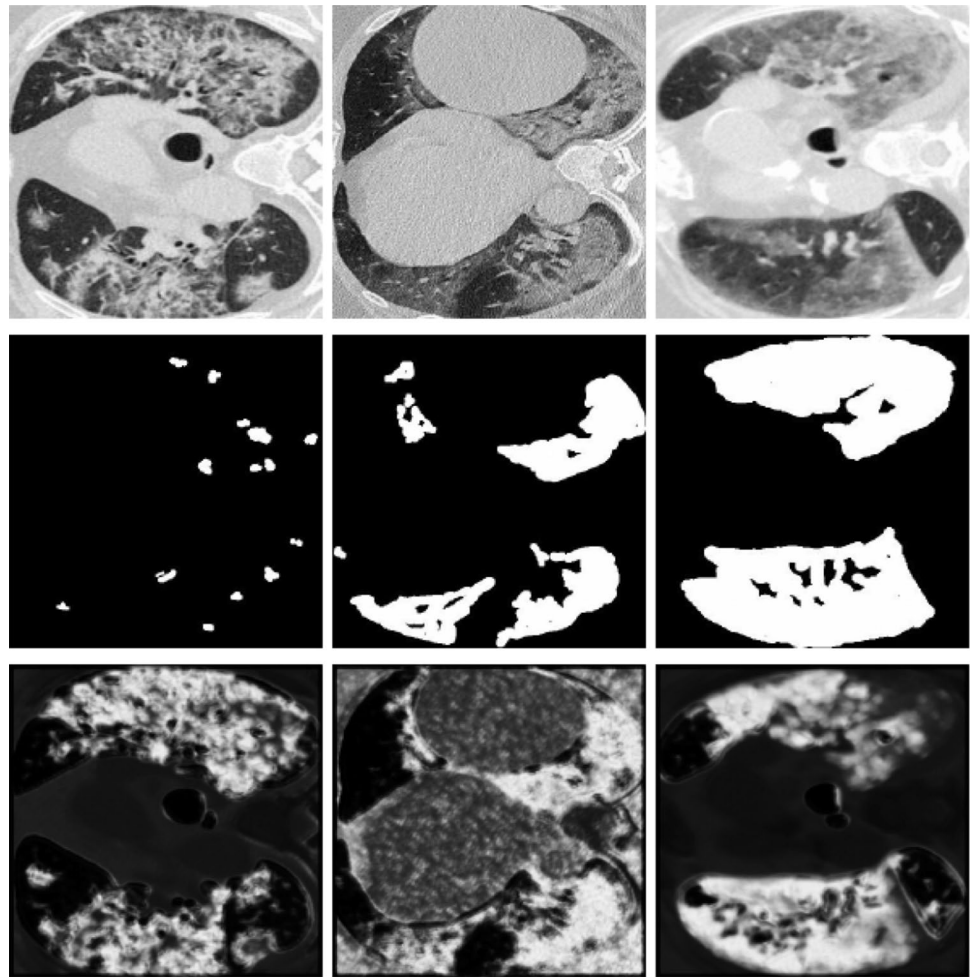
Table 4 Quantitative results of infection regions using multi-class segmentation on COVID-SemiSeg data set

Method	Dice	Sensitivity	Specificity	Precision	MAE
Multi-class U-Net [33]	0.581	0.672	0.902	-	0.066
DeepLabV3+ [35]	0.341	0.512	0.766	-	0.117
FC8s [34]	0.375	0.403	0.811	-	0.076
Semi-Inf-Net [27]	<i>0.541</i>	<i>0.564</i>	0.967	-	0.057
Proposed method	0.640	0.630	<i>0.953</i>	0.561	<i>0.060</i>

Bold, and italic indicates the first and the second best results, respectively

without structure–texture decomposition, as well as multi-class infection segmentation using the original images without ROI region segmentation and structure–texture decomposition. Figures 7 and 8 illustrate the results of the obtained results for the two scenarios compared by the ground truth. From the figure, we can find that the structure–texture decomposition introduced into the proposed model has an impact on the segmentation results. For example, In Fig. 7 first column, the proposed method without structure–texture decomposition can not segment the small infected parts. The

Fig. 7 Segmentation results using ROI binary segmentation without structure–texture preprocessing



same thing for the second and the third column, while just the big regions are segmented without precision. For Fig. 8 third column the proposed method segments the infected regions but with false segmented regions. the same for the first and second column with the segmentation is not effective and without specification of the small parts. When we compare the obtained results with the results in Fig. 6 we can find that ROI region segmentation on structure–texture components used as input of the model followed by infected segmentation gives convincing results. The Quantitative results presented in Table 6 also demonstrate that the processing phases provide good performance in terms of quantitative and qualitative results (Table 5).

Conclusion

In this paper, a lung infection segmentation method for COVID-19 has been proposed. Based on encoder–decoder networks on CT-scan images exploit the computer vision techniques to identify the lung infected regions for COVID-19 patients. Due to the shortage of data at this moment, multi-task learning has been performed including the use of two-stream as inputs of the deep learning model. In addition, using computer vision features like structure and texture component of the images that help for good extraction of the region of interest that can contain infections. Different segmentation has been performed including the binary segmentation and multi-class segmentation of lung infected regions. Comparing the proposed approach with the state-of-the-art method, the experiment shows an accurate segmentation of the lung infection region in both binary and multiclass segmentation. The obtained results can be improved using more data for training and more labeled data for multi-class segmentation, which represents our future works.

Fig. 8 Segmentation results using the proposed method without region infection segmentation and structure–texture preprocessing

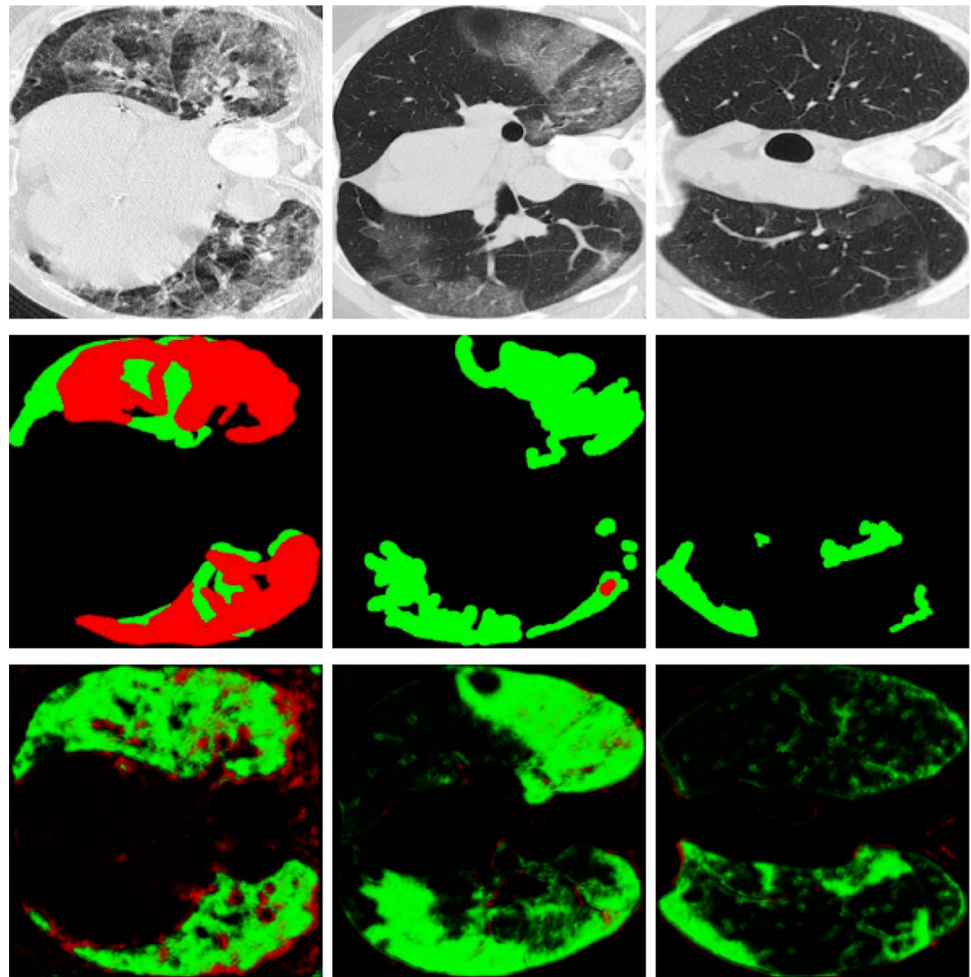


Table 5 Quantitative results of binary and multi-class segmentation on COVID-SemiSeg DATASET without region infection segmentation and structure–texture preprocessing

Method	Dice	Sensitivity	Specificity	Precision	MAE
ROI segmentation without ST(binary)	0.365	0.254	0.987	0.644	0.278
Model without ST and ROI Region (multi-class)	0.586	0.448	0.947	0.850	0.186
Proposed method (multi-class)	0.640	0.630	0.953	0.561	0.060

Acknowledgements This publication was jointly supported by Qatar University ERG-250. The findings achieved herein are solely the responsibility of the authors.

References

1. NGUYEN, Thanh Thi. . Artificial intelligence in the battle against coronavirus (COVID-19): a survey and future research directions. Preprint, DOI. 2020;10.
2. Ulhaq A, Khan A, Gomes D, et al. Computer vision for COVID-19 control: a survey. arXiv preprint [arXiv:2004.09420](https://arxiv.org/abs/2004.09420), 2020.
3. Sarrouti M, El Alaoui SO. A yes/no answer generator based on sentiment-word scores in biomedical question answering. *Int J Healthcare Inform Syst Inform (IJHISI)*. 2017;12(3):62–74.
4. Sarrouti M, Abacha AB, Demner-Fushman D. Visual question generation from radiology images. In 2020. p 12–8.
5. LATIF, Siddique, USMAN, Muhammad, MANZOOR, Sanaulah, , et al. Leveraging Data Science To Combat COVID-19: A Comprehensive. Review. 2020.
6. Arora V, Ng EYK, Leekha RS, Darshan M, Singh A (2021) Transfer learning-based approach for detecting COVID-19 ailment in lung CT scan. *Comput Biol Med*, p 104575.
7. Kalkreuth, Roman et Kaufmann, Paul COVID-19: A Survey on Public Medical Imaging Data Resources. arXiv preprint [arXiv:2004.04569](https://arxiv.org/abs/2004.04569), 2020.
8. Ozturk T, Talo M, Yildirim EA, et al. Automated detection of COVID-19 cases using deep neural networks with X-ray images. *Comput Biol Med*. 2020: 103792.

9. Minaee S, Kafieh R, Sonka M, et al. Deep-covid: Predicting covid-19 from chest x-ray images using deep transfer learning. arXiv preprint [arXiv:2004.09363](https://arxiv.org/abs/2004.09363), 2020.
10. Apostolopoulos ID, Aznaouridis SI, et Tzani MA Extracting possibly representative COVID-19 biomarkers from X-ray images with deep learning approach and image data related to pulmonary diseases. *J Med Biol Eng* 2020;1–8.
11. Apostolopoulos ID, et Mpesiana TA Covid-19: automatic detection from x-ray images utilizing transfer learning with convolutional neural networks. *Phys Eng Sci Med*. 2020;1.
12. Abbas A, Abdelsamea MM, et Gaber MM Classification of COVID-19 in chest X-ray images using DeTraC deep convolutional neural network. arXiv preprint [arXiv:2003.13815](https://arxiv.org/abs/2003.13815), 2020.
13. Cohen JP, Morrison P, et Dao L COVID-19 image data collection. arXiv preprint [arXiv:2003.11597](https://arxiv.org/abs/2003.11597), 2020.
14. “COVID-19 Patients Lungs X Ray Images 10000,” <https://www.kaggle.com/nabeelsajid917/covid-19-x-ray-10000-images>, accessed: 2020-04-11.
15. Chowdhury MEH, Rahman T et al. “Can AI help in screening Viral and COVID-19 pneumonia?” arXiv, 2020.
16. Zhao J, Zhang Y, He X, Xie P “COVID-CT-Dataset: a CT scan dataset about COVID-19,” arXiv, 2020.
17. “COVID-19 CT segmentation dataset,” <https://medicalsegmentation.com/covid19/>. Accessed 2020-04-11.
18. Riahi A, Elharrouss O, Almaadeed N, Al-Maadeed S (2021) BEMD-3DCNN-based method for COVID-19 detection.
19. Adhikari NCD. Infection severity detection of CoVID19 from X-rays and CT scans using Artificial intelligence. *Int J Comput (IJC)*. 2020;38(1):73–92.
20. Wu Y-H, Gao S-H, Mei J et al. JCS: an explainable COVID-19 diagnosis system by joint classification and segmentation. arXiv preprint [arXiv:2004.07054](https://arxiv.org/abs/2004.07054), 2020.
21. Mobiny A, Cicalese PA, Zare S et al. Radiologist-Level COVID-19 Detection using CT scans with detail-oriented capsule networks. arXiv preprint [arXiv:2004.07407](https://arxiv.org/abs/2004.07407), 2020.
22. Polsinelli M, Cinque L, et Placidi G A Light CNN for detecting COVID-19 from CT scans of the chest. arXiv preprint [arXiv:2004.12837](https://arxiv.org/abs/2004.12837), 2020.
23. Al-Karawi D, Al-Zaidi S, Polus N et al. Machine learning analysis of chest CT scan images as a complementary digital test of coronavirus (COVID-19). *Patients medRxiv*. 2020.
24. He X, Yang X, Zhang S et al. Sample-efficient deep Learning for COVID-19 diagnosis based on CT scans. *medRxiv*, 2020.
25. Amyar A, Modzelewski R, et Ruan S Multi-task deep learning based CT imaging analysis For COVID-19. Classification and Segmentation *medRxiv*. 2020.
26. Anwar, Talha et ZAKIR, Seemab. Deep learning based diagnosis of COVID-19 using chest CT-scan images. 2020.
27. Fan D-P, Zhou T, Ge-Peng JI et al. Inf-Net: Automatic COVID-19 lung infection segmentation from CT images. *IEEE Transactions on Medical Imaging*. 2020.
28. Ronneberger O, Fischer P, Brox T “U-Net: convolutional networks for biomedical image segmentation,” in MICCAI. Springer, 2015, pp 234–241.
29. Zhou Z, Siddiquee MMR, Tajbakhsh N, Liang J. UNet++: A nested U-Net architecture for medical image segmentation. *IEEE Transactions on Medical Imaging*. 2019;3–11.
30. Oktay O, Schlemper J, et al. Attention U-Net: learning where to look for the pancreas. In 2018.
31. Schlemper J, Oktay O, Schaap M, Heinrich M, Kainz B, Glocker B, Rueckert D. Attention gated networks: learning to leverage salient regions in medical images. *Med Image Anal*. 2019;53:197–207.
32. Li X, Chen H, Qi X, Dou Q, Fu C, Heng P. H-DenseUNet: hybrid densely connected UNet for liver and tumor segmentation from CT volumes. *IEEE Trans Med Imaging*. 2018;37(12):2663–74.
33. Ronneberger O, Fischer P, Brox T “U-Net: Convolutional networks for biomedical image segmentation,” in MICCAI. Springer, 2015, pp 234–241
34. Long J, Shelhamer E, Darrell T Fully convolutional networks for semantic segmentation. *CVPR*. 2015;3431–40.
35. Chen L-C, Zhu Y, Papandreou G, Schroff F, Adam H Encoder-decoder with atrous separable convolution for semantic image segmentation. *ECCV*. 2018;801–18.
36. Jan-Mark Geusebroek, R. van den Boomgaard, A.W.M. Smeulders, A. Dev, Color and scale: the spatial structure of color images, in: *Proceeding of the 6th European Conference on Computer Vision*, vol. 1, 2000, pp. 331–341.
37. Elharrouss O, Moujahid D, Tairi H. Motion detection based on the combining of the background subtraction and the structure–texture decomposition. *Optik-Int J Light Electron Opt*. 2015;126(24):5992–7.
38. Elharrouss O, Moujahid D, Elkah S, Tairi H. Moving object detection using a background modeling based on entropy theory and quad-tree decomposition. *Journal of Electronic Imaging*. 2016;25(6): 061615.
39. Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation. In 2015. p 3431–40.
40. Badrinarayanan V, Kendall A, Cipolla R. Segnet: a deep convolutional encoder-decoder architecture for image segmentation. *IEEE Trans Pattern Anal Mach Intell*. 2017;39(12):2481–95.
41. Simonyan, Karen et Zisserman, Andrew Very deep convolutional networks for large-scale image recognition. arXiv preprint [arXiv:1409.1556](https://arxiv.org/abs/1409.1556), 2014.
42. Ioffe S, Szegedy C Batch normalization: accelerating deep network training by reducing internal covariate shift. In 2015. p 448–56.
43. Shi F, Xia L, Shan F et al. Large-scale screening of covid-19 from community acquired pneumonia using infection size-aware classification. arXiv preprint [arXiv:2003.09860](https://arxiv.org/abs/2003.09860), 2020.
44. Shan F, Gao Y, Wang J, Shi W, Shi N, Han M, Xue Z, Shi Y “Lung infection quantification of covid-19 in ct images with deep learning.” arXiv preprint [arXiv:2003.04655](https://arxiv.org/abs/2003.04655) (2020).
45. Thanh T, Dang NH et al. Blood vessels segmentation method for retinal fundus images based on adaptive principal curvature and image derivative operators. In 2019.
46. Khan MA, Khan TM, Soomro TA, Mir N, Gao J. Boosting sensitivity of a retinal vessel segmentation algorithm. *Pattern Anal Appl*. 2019;22(2):583–99.
47. Elharrouss Omar, Al-Maadeed Noor, Al-Maadeed Somaya. Video summarization based on motion detection for surveillance systems. In 2019.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.