Received: 4 March 2021 | Revised: 28 April 2021 | Accepted: 30 April 2021

DOI: 10.1111/exsy.12742

ORIGINAL ARTICLE

Expert Systems WILEY

A deep learning semantic segmentation architecture for COVID-19 lesions discovery in limited chest CT datasets

Nour Eldeen M. Khalifa[1] | Gunasekaran Manogaran[2,3] | Mohamed Hamed N. Taha[1] | Mohamed Loey[4]

[1]Department of Information Technology, Faculty of Computers & Artificial Intelligence, Cairo University, Cairo, Egypt

[2]University of California, Davis, California, USA

[3]College of Information and Electrical Engineering, Asia University, Taichung, Taiwan

[4]Department of Computer Science, Faculty of Computers and Artificial Intelligence, Benha University, Benha, Egypt

Correspondence
Mohamed Loey, Department of Computer Science, Faculty of Computers and Artificial Intelligence, Benha University, Benha 13518, Egypt.
Email: mloey@fci.bu.edu.eg

Abstract

During the epidemic of COVID-19, Computed Tomography (CT) is used to help in the diagnosis of patients. Most current studies on this subject appear to be focused on broad and private annotated data which are impractical to access from an organization, particularly while radiologists are fighting the coronavirus disease. It is challenging to equate these techniques since they were built on separate datasets, educated on various training sets, and tested using different metrics. In this research, a deep learning semantic segmentation architecture for COVID-19 lesions detection in limited chest CT datasets will be presented. The proposed model architecture consists of the encoder and the decoder components. The encoder component contains three layers of convolution and pooling, while the decoder contains three layers of deconvolutional and upsampling. The dataset consists of 20 CT scans of lungs belongs to 20 patients from two sources of data. The total number of images in the dataset is 3520 CT scans with its labelled images. The dataset is split into 70% for the training phase and 30% for the testing phase. Images of the dataset are passed through the pre-processing phase to be resized and normalized. Five experimental trials are conducted through the research with different images selected for the training and the testing phases for every trial. The proposed model achieves 0.993 in the global accuracy, and 0.987, 0.799, 0.874 for weighted IoU, mean IoU and mean BF score accordingly. The performance metrics such as precision, sensitivity, specificity and F1 score strengthens the obtained results. The proposed model outperforms the related works which use the same dataset in terms of performance and IoU metrics.

KEYWORDS
COVID-19, CT images, deep learning, semantic segmentation, transfer learning

1 | INTRODUCTION

Based on the World Health Organization (WHO), the coronavirus (COVID-19) pandemic puts healthcare services worldwide under immense and growing pressure (WHO) (Ghazy et al., 2020). Coronavirus Disease 2019 (COVID-19) has become a global pandemic with an unprecedented rate of development and a mechanism of dissemination that is not well known. Governments have begun to focus on new space utilization techniques, psychological distancing, and supplies for medical workers and regular persons (Loey et al., 2021b). This COVID-19 will induce a respiratory ailment to those that have it, and it is infectious, transmitted by droplets in the air. The infection is most commonly spread through close contact and through the inhalation of the nose and upper respiratory droplets which are delivered when a person coughs or sneezes. The ugly symptoms

this infection produces can include coughing, breathing difficulties, and fever (Wang, Wang, et al., 2020). Trouble in breathing is a warning that you might be developing pneumonia which needs immediate health consideration. Efficient screening of contaminated patients is a crucial phase in the battle against COVID-19 (Vandenberg et al., 2020).

Deep learning has demonstrated immense promise in many real-life implementations in several different areas (Khalifa, Taha, et al., 2020). All these potential applications include object recognition. This new method of object recognition is based on deep learning algorithms. This type of deep learning has obtained impressive results in identifying an object in images (Loey et al., 2021a). Chest CTs also emerged in clinical practice as an invaluable method for the diagnosis of lung disease in COVID-19 related individuals (Kwee & Kwee, 2020). When working with lung images, segmentation is necessary to analyse the lung images, for example, to detect disease. COVID-19 identification is a difficult task that regularly involves taking a gander at patient clinical photos (Perumal et al., 2021). Via deep learning and semantic image segmentation, the main objective of this research is to identify and localize the COVID-19 contaminated areas in the patient's lungs.

Every area of work involves the development of goals, which helps in solving the cases and problems facing the organization. The field of Artificial Intelligence involves both studying and constructing intelligent entities which include both machine learning and deep learning. Deep neural networks (DNNs) are composed of artificial neurons that are built up in layers of data processing (Loey et al., 2021a) (Khalifa et al., 2021). These networks are structured in a deep architecture referred to as a 'deep neural network', a collection of multiple layers, that translate information into decisions. A subtype is CNN, whose known uses are exemplified by images and videos (Khalifa, Loey, et al., 2020). While information interchange at times gets hard to manage, convolutional neural networks (CNNs) have been a subtype that has flourished and found popularity due to astounding breakthroughs in image and video processing (Loey et al., 2020).

Segmentation can be characterized as partitioning the image into homogeneous pixel groups. It can be performed in a highly sophisticated way, down to the level of a single pixel (Rogowska, 2009). Semantic pixel-level segmentation is the task of clustering image parts together that belong to the same object class (Wang et al., 2018). Semantic image segmentation is the method of converting an input image into a raster so that any pixel is labelled to a certain significance entity type from a predefined class nomenclature (Kang et al., 2020). The appearance of items is learned using the labelled pictures. Methods typically fuse two sub-tasks: an identification component that labels pixels with the class, and a regularization part that enforces other pixels in the neighbourhood to be in the same class. The recognizing aspect depends on the pixel or patch stage of the local appearance (Li et al., 2020). Various elements, including texture, colour statistics and SIFT, are used for defining the local appearance.

The main contributions of this research are: (1) proposing a new model for the detection of COVID-19 lesions in CT Lung images, (2) measuring the accuracy and the performance metrics of the proposed model, and (3) comparing the proposed model with other related works according to predefined metrics. One of the technical novelties for the presented work is the architecture of the proposed model. This work contains the following sections. The second section addresses past writings on relevant research. Section 3 discusses the dataset. Section 4 the proposed model for detecting COVID-19 infected regions. The article discusses the outcomes of the experiment and ends with potential work.

2 | RELATED WORK

Semantic Segmentation refers to assigning a class label to each pixel in the given image. This is done through Object identification and class labelling to each pixel in the given image (Garcia-Garcia et al., 2018). Learning and classifying models for vision and image processing have been applied in artificial intelligence, robotics, and computer vision, and have been used with effectiveness to identify objects and control segmentation in images. The semantic segmentation method is deeply used in most segmentation problems in general. It can be used in the autonomous car driving by segmenting the roads, signs, cars, and pedestrians. Also, it is heavily used in the medical domain such as tumour detection in CT and X-rays. To gain a better understanding of infection rates and to track and rapidly detect infections, the AI's capabilities are being leveraged to address the Covid-19 pandemic (Bhattacharya et al., 2021; González-Crespo et al., 2020), including the detection of COVID-19 in medical chest CT scan (Lassau et al., 2021). This section focuses on recent papers in terms of applying representative works related to semantic segmentation based on deep learning for CT image diagnosis.

In Oulefki et al. (2020), the writers demonstrated an algorithm for automatic COVID-19 Lung infection segmentation using chest CT images. Depending on 3D U-Net architecture, the proposed tool is capable of detecting irregular regions with low contrast between lesions and healthy tissues with an accuracy of 0.98 and a precision of 0.73. The authors in (Yang et al., 2020) developed and tested a 3D deep learning model which is called CODID-SegNet, for using the COVID-19 segmentation method on CT pictures of the chest. The planned network would consist of ASPs that contain features that are good at highlighting the COVID-19 boundaries and positions. Yang, He, et al. (2020)) introduced the MultiResUNet model to slice COVID-19 from CT pictures. They adjust certain convolutions to render activation functions simpler, incorporate regularization and dropout to manage to overfit. The average precision score for Right, left, disease, and background lung is 84.66, 87.04, 72.14 and 99.23 respectively. Outcomes of the segmentation study are analysed, and the required outcomes are achieved. It is critical for medical personnel to identify rapidly where the infection is located.

In Wang, Zhang, et al. (2020) introduced four transfer learning models based on U-Net. Three freely accessible non-COVID19 datasets were used in their research. The suggested hybrid-encoder learning approach includes a dedicating encoder for the COVID 19 infection function and

general lung lesion features. Results show that a Hybrid-encoder state-of-the-art neural network model achieves accuracy, sensitivity, F1-score of 0.994, 0.682, 0.707, respectively, when trained with only minimal data. In Ma et al. (2020)), the authors create a benchmark dataset for COVID-19 segmentation. Authors gave three datasets to segment lung and infection from COVID-19. Also, the authors include the guidance of standard preparation, measurement measures, and the corresponding code. Through training on 40 baselines, researchers can save resources by preventing needing to train their techniques and can concentrate on designing their approaches. In Voulodimos et al. (2020)) introduced a segmentation model of pneumonia based on CT images. The proposed model based on U-Nets as convolutional neural architecture based on a real dataset. The analysis revealed that Completely Convolutional Neural Networks correctly segment the entity groups in the specified dataset.

In Yang, Li, et al. (2020), the authors create a new open benchmark dataset of COVID-19. The research data collection includes 349 CT photographs from 216 patients and 463 non-CT pictures. They show the reliability of the dataset for the creation of an AI-based diagnostic model of COVID-19. They create a machine learning algorithm that achieves an accuracy score of .89, on a test collection of initial CT photos from hospitals. In Fan et al. (2020)) proposed COVID-19 segmentation model for CT pictures. The proposed model classifies infected regions from chest CT images. A parallel partial decoder that includes high-level features as features is used to generate a global map. They present a semi-supervised segmentation system that is focused on a random propagation technique, which only uses a few labelled training pictures and utilizes heavily uncategorized data. There is tremendous scope for the method to be incorporated in the diagnosis of COVID-19.

3 | DATASET

Data were collected in this study from (Jun et al., n.d.). These CT scans are numbered with numbers from one to twenty. Left lung, right lung, and multiple-sided disease were assessed by two radiologists who collaborated, with the help of an expert, a chest radiologist, to be present in this case. The dataset is considered a heterogeneous one as it includes both healthy and COVID-19 and CT scans from two sources. The first source includes 10 patients from a single source of CT scans, while the second source includes 10 patients' CT scans from the radiopaedia website (Radiopaedia.Org, the Wiki-Based Collaborative Radiology Resource, n.d.). The first source consists of 2581 images of CT scans with its labelled images. The second source (radiopaedia) consists of 939 images of CT scans with its labelled images. The total number of images from the two sources is 3520 CT scans images with its labelled images. Figure 1 illustrates samples from the two sources for the original dataset.

4 | THE PROPOSED MODEL ARCHITECTURE

In this part, a detailed overview of the introduced architecture for semantic segmentation will be presented. The proposed architecture flowchart is presented in Figure 2. The model passes through three phases. The first phase is the pre-processing, while the second phase is the training. The third phase is the testing phase. In the pre-processing phase, all images in the dataset were resized to 128 $\times$ 128 pixels to reduce the processing time in the training phase. Moreover, all the pictures were normalized to remove any anomalies in the dataset images. In the training phase, the

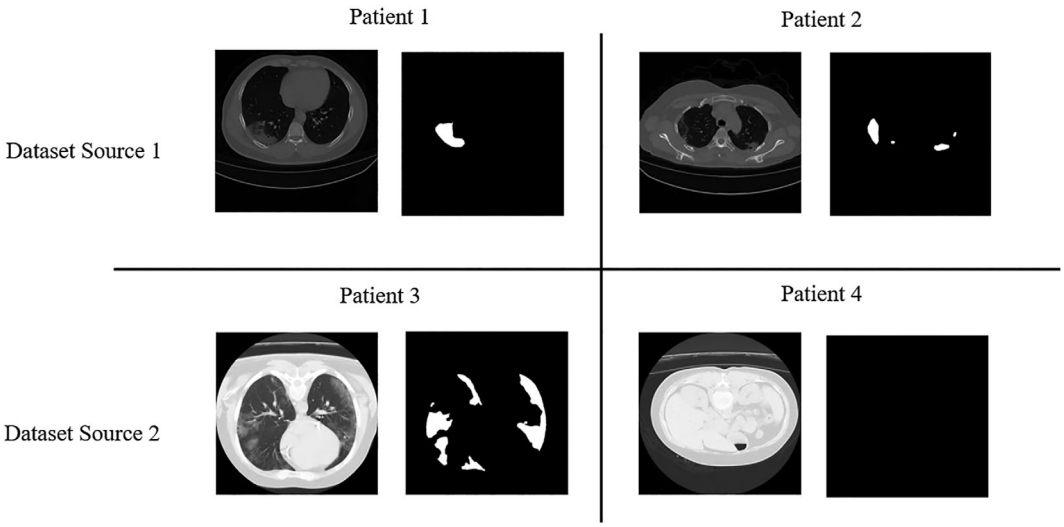

FIGURE 1 Dataset sample images from classified to it is the original source

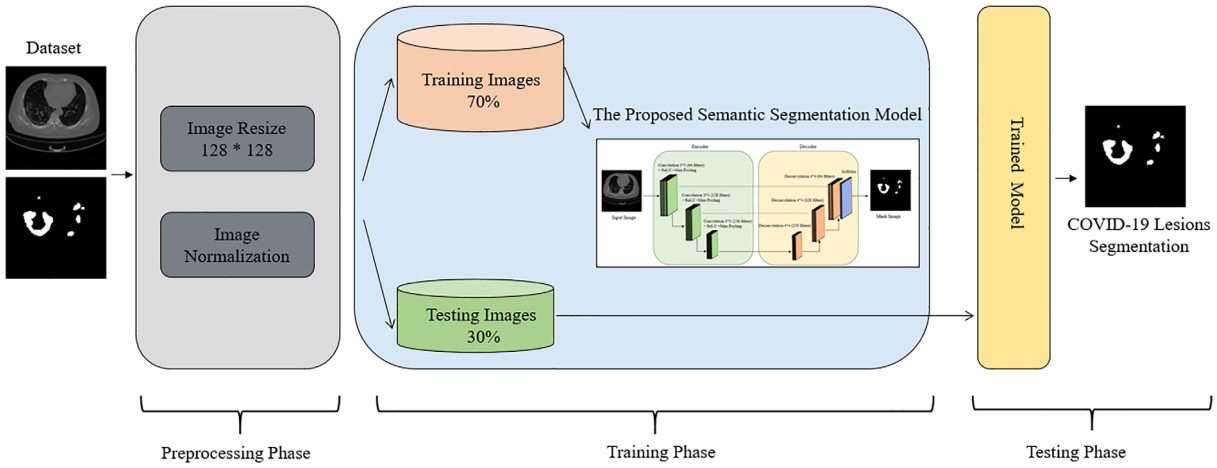

FIGURE 2 The proposed model flowchart

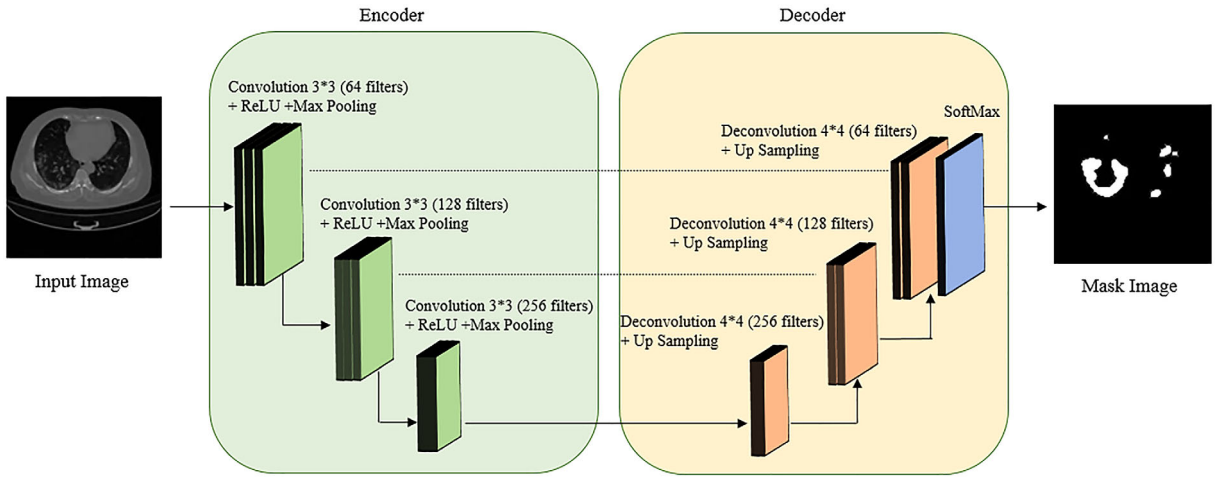

FIGURE 3 The architecture of the proposed semantic segmentation model

images were split into two parts. The first part for training purposes with 70% of the original data, While the second part was 30% for testing purposes. The training part will be fed into the proposed model for semantic segmentation which is presented in Figure 3 and will be discussed next section. The testing phase is used for testing the accuracy of the introduced architecture over the test data.

In this semantic segmentation, the architecture has two separate components as presented in Figure 3. The first component is the encoder with three stages (s_{e1},s_{e2},s_{e3}). Every stage consist of one convolutional layer, one ReLU layer, and one max-pooling layer with total nine layers. Encoder which is a down sampling operation is aimed at capturing semantic or context information. The input is fed up with the encoder part. The input image is an image of CT for the chest with size 128 $\times$ 128 pixels and normalized. The encoder's first convolutional layer has 64 filters followed by ReLU and max-pooling layer. The output of the first max-pooling layer is then fed into the second convolutional layer (contains 128 filters) and then it is followed by ReLU and the second max-pooling layer. The output of the previous max pooling is fed into the third convolutional layer which contains 256 filters then followed by ReLU and max-pooling layer.

After extracting the compressed features, the decoder returns the features to their original scale, thereby reducing the sum of knowledge loss due to compression. Decoder which is an up-sampling operation is aimed at recovering spatial information. Common decoders include bilinear interpolation, deconvolution and dense up sampling convolution. In the proposed model, the deconvolution option has been used. The decoder function is the opposite function of the encoder. The second component is the encoder within 3 stages (s_{d1},s_{d2},s_{d3}). The layer of the encoder component consists of the deconvolutional layers and the SoftMax layer. The first deconvolutional layer in the encoder consists of 256 filters, while the second deconvolutional layer consists of 128 filters and the third deconvolutional layer consists of 64 filters. The SoftMax layer is the last in the decoder component followed by the pixel classification layer. The output of the decoder layer is a mask image of the COVID-19 lesions in the lung as presented in Figure 2.

5 | EXPERIMENTAL RESULTS

A commercial software program (MATLAB) was used to develop the proposed semantic segmentation model. The execution was particular to run in a central processing unit (CPU). All tests were conducted on a computer server with an Intel Xeon CPU, about 100 GB of RAM with Titan X GPU.

The experimental trial setups were designed to capture the accuracy of the quantification along with the planned model's performance specifications. The following setups were adopted during the experimental trials:

- Two sources of data as illustrated in the dataset section and will be noted as DS1 for dataset source 1 and D_{S2} for dataset source 2. Moreover, D_{S3} is a merged source of D_{S1} and D_{S2}.
- The training phase will be experimented with 70% of the data, while 30% will be used in the testing phase.
- The number of test images for D_{S1} is 774, for D_{S2} is 281 and for D_{S3} is 1055 images.
- Five trials were conducted with different images randomly selected for every trial for the training/testing phases.

Along with the following parameters set up for the proposed semantic segmentation deep learning model:

- Batch size: 64
- Momentum: 0.9
- Number of epochs: 300
- Learning rate: 0.001
- Optimizer: Adam

The following accuracy and performance metrics will be measured during the experimental trials and they are (Gaspari et al., 2020) (Goutte & Gaussier, 2005):

- Accuracy: the ratio between the total pixels correctly identified of a class over the number of pixels of that label as presented in Equation (1).

$$\text{Accuracy} = \frac{C_{\text{tp}} + C_{\text{tn}}}{(C_{\text{tp}} + C_{\text{fp}}) + (C_{tn} + C_{fn})} \tag{1}$$

where C_{tp} is the counter of correctly identified pixels in the class, C_{fp} is the counter of miss-labelled of rest of pixels of the classes, C_{tn} is the counter of correctly identified pixels of the rest of the classes, and C_{fn} is the counter of miss-labelled pixels in the class. Sensitivity, Specificity, Precision, Accuracy and F_1 Score (Goutte & Gaussier, 2005) are introduced from Equation (2) to Equation (5).

$$\text{Sensitivity} = \frac{C_{\text{tp}}}{(C_{\text{tp}} + C_{fn})} \tag{2}$$

$$\text{Specificity} = \frac{C_{\text{tp}}}{(C_{\text{fp}} + C_{tn})} \tag{3}$$

$$\text{Precision} = \frac{C_{\text{tp}}}{(C_{\text{tp}} + C_{\text{fp}})} \tag{4}$$

$$\text{F}_1\ \text{Score} = \frac{2 * C_{\text{tp}}}{(2 * C_{\text{tp}} + C_{\text{fp}} + C_{fn})} \tag{5}$$

- Global Accuracy: the ratio between the total pixels of one class overall number of pixels.
- Mean Accuracy: the average of the Accuracy values of all the classes
- IoU: Intersection over-union is an estimation metric that is used to determine the object detection precision on a given dataset as presented in Equation (6).

$$IoU = \frac{area(C \bigcap G)}{area(C \bigcup G)} \quad (6)$$

where C is the predicted mask and G is the ground truth mask.

- Mean IoU: IOU is a commonly used measure for semantic segmentation, which first counts the IOUs and then computes their average.
- Weighted IoU: The individual Class IOU averages were calculated using the total number of pixels as a weight.
- BF Score: acronym for Boundary F_1 contour matching score, it indicates how close the predicted boundary of a class matches the ground truth boundary.
- Mean BF Score: the average of the BF_{Score} of all the classes.

5.1 | Testing performance outcomes

According to the previous section metrics, testing performance is considered the most common metric in deep learning evaluation metrics. Those metrics will be measured for the different dataset sources D_{S1}, D_{S2} and D_{S3}. Table 1 presents the global accuracy, mean accuracy, sensitivity, specificity, precision, and F_1 score for the different dataset sources.

Table 1 illustrates that for D_{S1}, the results for all metrics super pass the achieved results for D_{S2} and D_{S3}. One of the reasons for better accuracy for D_{S1} due to the homogeneity of the sources according to the image's quality and its labelled masks. The D_{S2} is captured from the radiopaedia website, which by its definition contains images of different qualities. The achieved results are not the highest, as it achieves 0.9842 in global accuracy and 0.9775 in mean accuracy but still a promising result. For the combined dataset source (D_{S3}). The proposed model achieved 0.9930 in global accuracy and close to D_{S1}, and 0.9717 in mean accuracy.

Table 1 showed that the achieved results for D_{S1} were the highest ones with 0.9932, 0.9872, 0.9937, 0.9812, 0.9817 and 0.9874 for global accuracy, mean accuracy, precision, sensitivity, specificity, and F1 score accordingly. The achieved results due to the homogeneity and the quality of the images from dataset Source 1.

The achieved result in Table 1 was calculated from the confusion matrix for D_{S1}, D_{S2} and D_{S3}. Figure 4 presents the confusion matrices for the different dataset sources.

One of the most common accuracies in deep learning is the accuracy for every class. Table 2 presents the accuracy for every class for the proposed model.

Table 2 illustrates that in the background class accuracy, D_{S3} achieves the highest accuracy with 0.9935. For COVID-19 class accuracy, D_{S1} achieves the highest accuracy possible with 0.9812 due to the quality and homogeneity of the images in D_{S1}. For background pixels, it is all black in all dataset sources and more data means higher accuracy in deep learning, and D_{S3} contains the highest number of images among the other dataset sources.

5.2 | Weighted IoU, mean IoU and BF score outcomes

Weighted IoU, Mean IoU and Mean BF Score is the most commonly used parameters in object and region of interest detection. Table 3 presented the IoU metrics and Mean BF score for the proposed model for different dataset sources.

Table 3 shows that for D_{S1} and D_{S3} the weighted IoU is very close to each other with 0.9892 for D_{S1} and 0.9886 for D_{S3}. In the mean IoU metric and the mean BF score, D_{S3} achieved the highest accuracy possible with 0.8746. The achieved results were due to the large number of images in D_{S3}.

Figure 5 illustrates the mean IoU metric for the dataset sources. The figure shows the number of images for mean IoU achieved accuracy. For instance, in D_{S1}, there are 300 images from the test images achieved mean IoU $= 1$. In D_{S2}, there are 80 images from the test images achieved mean IoU $= 1$. In D_{S3}, there are 160 images from the test images achieved mean IoU $= 0.8$.

TABLE 1 Accuracy results with its performance metrics

Dataset source	Global accuracy	Mean accuracy	Precision	Sensitivity	Specificity	F1 score
DS1	0.9932	0.9872	0.9937	0.9812	0.9817	0.9874
DS2	0.9842	0.9775	0.9842	0.9706	0.9706	0.9774
DS3	0.9930	0.9717	0.9932	0.9499	0.9499	0.9711

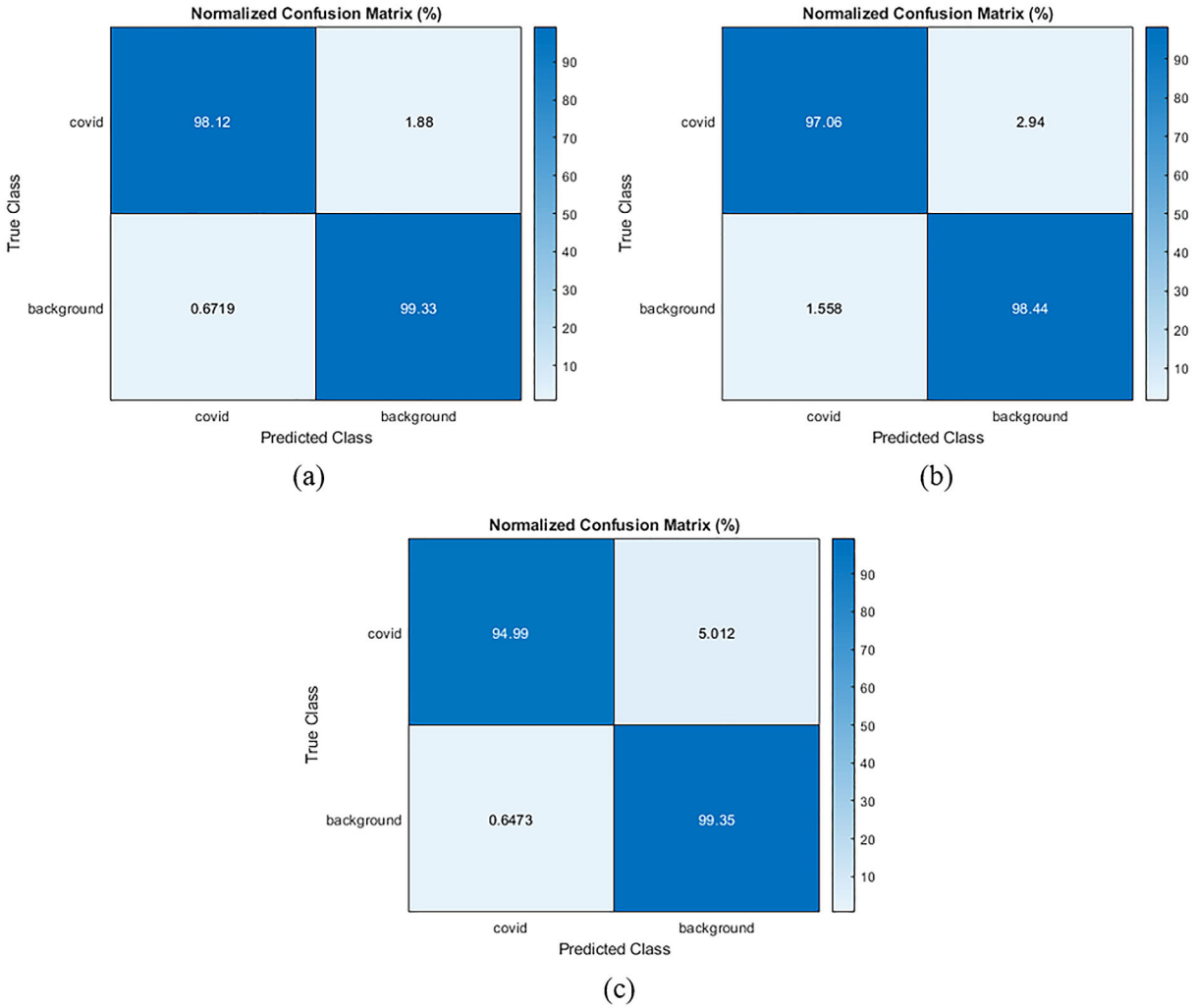

FIGURE 4 Confusion matrix for (a) D_{S1}, (b) D_{S2}, (c) D_{S3}

TABLE 2 Testing accuracy for the different class for dataset sources

Testing accuracy	D_{S1}	D_{S2}	D_{S3}
COVID-19 Class	0.9812	0.9706	0.9499
Background Class	0.9933	0.9844	0.9935

TABLE 3 IoU metrics and mean BF score for different dataset sources

Dataset source	Weighted IoU	Mean IoU	Mean BF score
D_{S1}	0.9892	0.7807	0.8591
D_{S2}	0.9759	0.7468	0.8042
D_{S3}	0.9886	0.7990	0.8746

5.3 | Results discussion and comparison with related works

From the previous subsection, we can deduce that the homogeneity of the images and their number affects the obtained results. The proposed architecture over D_{S3} achieved the highest metrics in Section 5.2 Weighted IoU, Mean IoU and Mean BF Score with 0.9886, 0.7990, 0.874, accordingly. Those metrics concerns with the model accuracy in detecting and localizing the COVID-19 Lesions in CT lungs images. While in testing accuracy for classification of the pixels whether is to belong to COVID-19 or background class the proposed model achieves the highest accuracy possible over D_{S1} and close result to D_{S3} according to global accuracy metric with 0.9932 in D_{S1} and 0.9930 in D_{S3}. So, the research decision

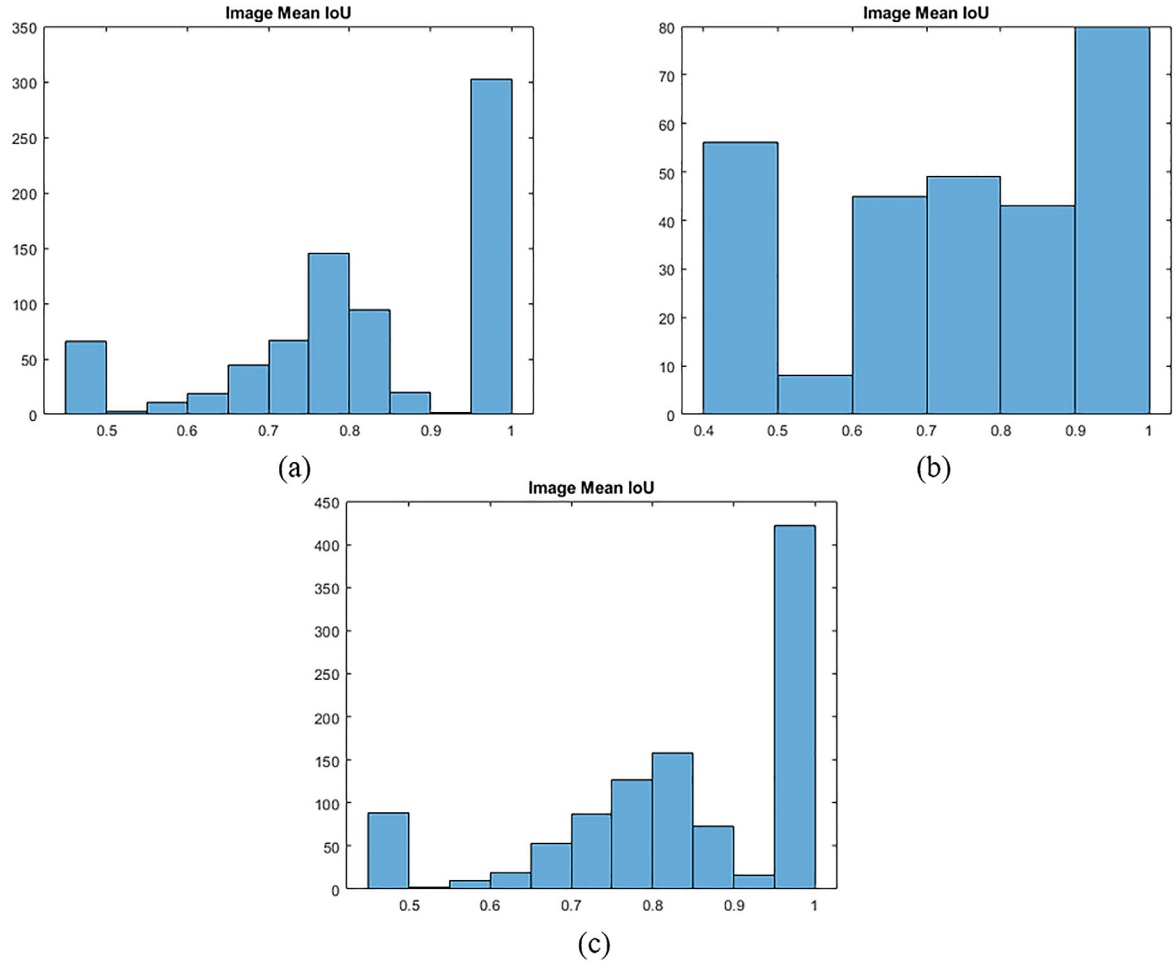

FIGURE 5 Mean IoU for (a) D_{S1}, (b) D_{S2}, and (c) D_{S3}

was to select the proposed model for D_{S3} as the main results of this research with 0.9930 global accuracies, 0.9886, 0.7990, 0.874 for Weighted IoU, Mean IoU and Mean BF Score.

Figure 6 illustrated the final image segmentation result for the proposed model. The images in the (d) label present the predicted segmentation of COVID-19 lesions in CT lungs. The difference between the predicted mask and the original mask is considered small as it achieved according to the previous section around 0.9930 accuracies with weighted IoU equal to 09886.

Table 4 presents a comparison of results with related works that used the same dataset. Table 4 shows clearly that the proposed model outperformed the related works in terms of accuracy, f1 score, precision, sensitivity and IoU. None of the related works reported the IoU instead they (Wang, Zhang, et al., 2020) (Ma et al., 2020) report Dice Similarity Coefficient which is very similar to IoU. The Dice Similarity Coefficient tends to measure the average performance, while the IoU scores the worst-case performance.

6 | CONCLUSION

COVID-19 pandemic is one of the devastating pandemics throughout modern history. Its effects did not reflect only on human lives but also on world economics, global trading and consumption patterns. Advances in the computer science field and in particular in deep learning helped in improving the medical field's precision and speed. In this article, a deep learning semantic segmentation model for COVID-19 lesions detection in a limited chest CT dataset was presented. The proposed learning approach integrated functionality from non-COVID19 databases successfully and achieves substantial progress over other learning solutions. Studies showed that trained with a minimal dataset, proposed model strategy focused on pre-trained model achieves a mean Sensitivity, F1-score, and Accuracy of 0.95, 0.97 and 0.99, respectively, with improved generalization and lower overfitting threats for segmenting COVID-19 infection. The research offers information regarding the effects of transmitting expertise from non-COVID19 lung lesions in the light of learning from several lung lesion datasets, contributing to successful models. These

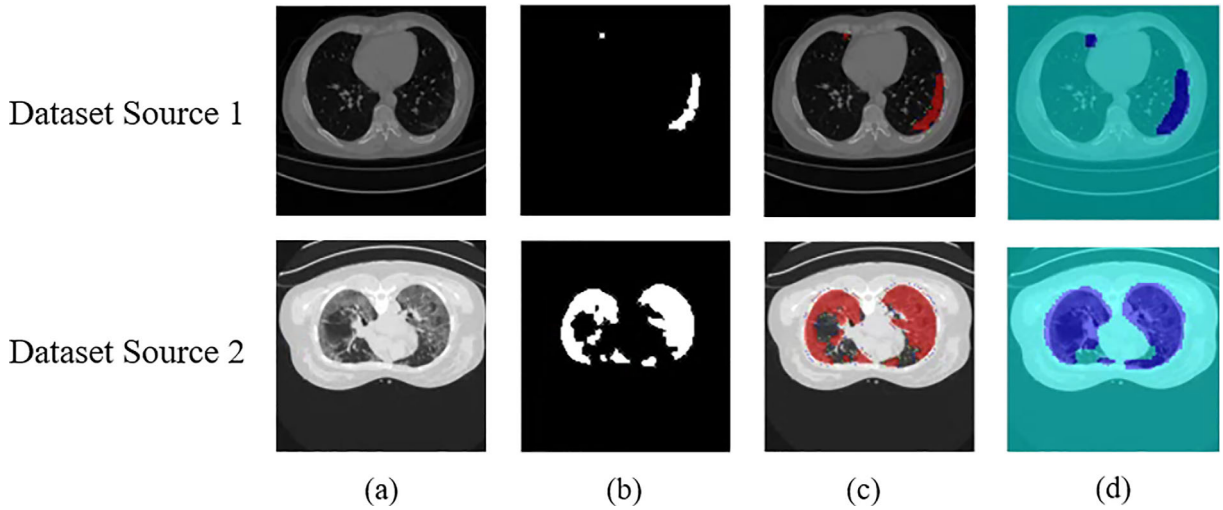

FIGURE 6 The final image segmentation result for the proposed model was (a) the original CT image, (b) the original mask image, (c) the original mask image over the original CT image, and (d) produced mask of the proposed model over the original CT image

TABLE 4 Comparison results with related works

	Description	ACC	F1 score	Precision	Sensitivity	Dice similarity coefficient/IoU
(Voulodimos et al., 2020)	FCN and Unet	0.97	0.65	0.95	0.55	–
(Wang, Wang, et al., 2020; Wang, Zhang, et al., 2020)	Pretrained Multi-Lesions	0.99	0.71	–	0.68	0.707
(Ma et al., 2020)	40 pre-trained baseline models	–	–	–	–	0.700
(Yang, He, et al., 2020; Yang, Li, et al., 2020)	contrastive self-supervised learning with transfer learning	0.89	0.89	–	–	–
(Fan et al., 2020)	Deep Network (Inf-Net)	–	–	0.50	0.87	0.59
(Fan et al., 2020)	Deep Network (Semi Inf-Net)	–	–	0.51	0.86	0.59
Proposed Model		0.99	0.97	0.99	0.95	0.799

outcomes showed that transfer learning from other medical activities for CT image segmentation is feasible, and this can be applied to other tasks. One of the possible future works is the apply the proposed model in others segmentation application domains such as brain CT scans to measure its performance. Also, building a deeper semantic model with more layers and test its performance against the proposed model.

CONFLICT OF INTEREST

On behalf of all authors, the corresponding author states that there is no conflict of interest.

DATA AVAILABILITY STATEMENT

Research data are not shared.

ORCID

Mohamed Hamed N. Taha https://orcid.org/0000-0003-0200-2918
Mohamed Loey https://orcid.org/0000-0002-3849-4566

REFERENCES

Bhattacharya, S., Reddy Maddikunta, P. K., Pham, Q.-V., Gadekallu, T. R., Krishnan, S., Chowdhary, S. R., Alazab, C. L., & Jalil Piran, Md., M. (2021). Deep learning and medical image processing for coronavirus (COVID-19) pandemic: A survey. *Sustainable Cities and Society*, *65*, 102589. https://doi.org/10.1016/j.scs.2020.102589

Fan, D.-P., Zhou, T., Ji, G.-P., Zhou, Y., Chen, G., Fu, H., Shen, J., & Shao, L. (2020). Inf-net: Automatic COVID-19 lung infection segmentation from CT images. *IEEE Transactions on Medical Imaging*, *39*(8), 2626–2637. https://doi.org/10.1109/TMI.2020.2996645

Garcia-Garcia, A., Orts-Escolano, S., Oprea, S., Villena-Martinez, V., Martinez-Gonzalez, P., & Garcia-Rodriguez, J. (2018). A survey on deep learning techniques for image and video semantic segmentation. *Applied Soft Computing*, *70*, 41–65. https://doi.org/10.1016/j.asoc.2018.05.018

Gaspari, A., Natale, E., De Silvestri, A., & D'Emilia, G. (2020). Effect of measurement uncertainty on artificial vision methods, for quality control on composite components.

Ghazy, R. M., Almaghraby, A., Shaaban, R., Kamal, A., Beshir, H., Mourси, A., Ramadan, A., & Taha, S. H. N. (2020). A systematic review and meta-analysis on chloroquine and hydroxychloroquine as monotherapy or combined with azithromycin in COVID-19 treatment. *Scientific Reports*, *10*(1), 22139. https://doi.org/10.1038/s41598-020-77748-x

González-Crespo, R., Herrera-Viedma, E., Dey, N., Fong, S. J., & Li, G. (2020). Finding an accurate early forecasting model from small dataset: A case of 2019-nCoV novel coronavirus outbreak. *International Journal of Interactive Multimedia and Artificial Intelligence*, *6(special issue on soft computing)*. *6*, 132–140. https://www.ijimai.org/journal/bibcite/reference/2752

Goutte, C., & Gaussier, E. (2005). A probabilistic interpretation of precision, recall and F-score, with implication for evaluation. *European Conference on Information Retrieval*, *3408*, 345–359.

Jun, M., Cheng, G., & Yixin, W.. (n.d.). COVID-19 CT lung and infection segmentation dataset. https://zenodo.org/record/3757476

Kang, G., Wei, Y., Yang, Y., Zhuang, Y., & Hauptmann, A. (2020). Pixel-level cycle association: A new perspective for domain adaptive semantic segmentation. *Advances in Neural Information Processing Systems*, *33*, 3569–3580.

Khalifa, N. E. M., Loey, M., Mawgoud, A. A., & Taha, M. H. N. (2020). Empirical study and enhancement on deep transfer learning for skin lesions detection. *Journal of Theoretical and Applied Information Technology*, *98*(20), 1351–1361.

Khalifa, N. E. M., Smarandache, F., Manogaran, G., & Loey, M. (2021). A study of the Neutrosophic set significance on deep transfer learning models: An experimental case on a limited COVID-19 chest X-ray dataset. *Cognitive Computation*. *5*, 1–13. https://doi.org/10.1007/s12559-020-09802-9

Khalifa, N. E. M., Taha, M. H. N., Ezzat Ali, D., Slowik, A., & Hassanien, A. E. (2020). Artificial intelligence technique for gene expression by tumor RNA-Seq data: A novel optimized deep learning approach. *IEEE Access*, *8*, 22874–22883. https://doi.org/10.1109/ACCESS.2020.2970210

Kwee, T. C., & Kwee, R. M. (2020). Chest CT in COVID-19: What the radiologist needs to know. *Radiographics*, *40*(7), 1848–1865. https://doi.org/10.1148/rg.2020200159

Lassau, N., Ammari, S., Chouzenoux, E., Gortais, H., Herent, P., Devilder, M., Soliman, S., Meyrignac, O., Talabard, M.-P., Lamarque, J.-P., Dubois, R., Loiseau, N., Trichelair, P., Bendjebbar, E., Garcia, G., Balleyguier, C., Merad, M., Stoclin, A., Jegou, S., … Blum, M. G. B. (2021). Integrating deep learning CT-scan model, biological and clinical variables to predict severity of COVID-19 patients. *Nature Communications*, *12*(1), 634. https://doi.org/10.1038/s41467-020-20657-4

Li, Y., Liu, Y., Liu, G., & Guo, M. (2020). Weakly supervised semantic segmentation by iterative superpixel-CRF refinement with initial clues guiding. *Neurocomputing*, *391*, 25–41. https://doi.org/10.1016/j.neucom.2020.01.054

Loey, M., Manogaran, G., & Khalifa, N. E. M. (2020). A deep transfer learning model with classical data augmentation and CGAN to detect COVID-19 from chest CT radiography digital images. *Neural Computing and Applications*. *32*, 1–13. https://doi.org/10.1007/s00521-020-05437-x

Loey, M., Manogaran, G., Taha, M. H. N., & Khalifa, N. E. M. (2021a). A hybrid deep transfer learning model with machine learning methods for face mask detection in the era of the COVID-19 pandemic. *Measurement*, *167*, 108288. https://doi.org/10.1016/j.measurement.2020.108288

Loey, M., Manogaran, G., Taha, M. H. N., & Khalifa, N. E. M. (2021b). Fighting against COVID-19: A novel deep learning model based on YOLO-v2 with ResNet-50 for medical face mask detection. *Sustainable Cities and Society*, *65*, 102600. https://doi.org/10.1016/j.scs.2020.102600

Ma, J., Wang, Y., An, X., Ge, C., Yu, Z., Chen, J., Zhu, Q., Dong, G., He, J., He, Z., & others. (2020). Towards data-efficient learning: A benchmark for COVID-19 CT lung and infection segmentation. *Medical Physics*, *48*(3), 1197–1210.

Oulefki, A., Agaian, S., Trongtirakul, T., & Kassah Laouar, A. (2020). Automatic COVID-19 lung infected region segmentation and measurement using CT-scans images. *Pattern Recognition*, 107747. *114*, https://doi.org/10.1016/j.patcog.2020.107747

Perumal, V., Narayanan, V., & Rajasekar, S. J. S. (2021). Detection of COVID-19 using CXR and CT images using transfer learning and Haralick features. *Applied Intelligence*, *51*(1), 341–358. https://doi.org/10.1007/s10489-020-01831-z

Radiopaedia.org, the wiki-based collaborative Radiology resource. (n.d.). https://radiopaedia.org/

Rogowska, J. (2009). Chapter 5-Overview and fundamentals of medical image segmentation. In I. N. Bankman (Ed.), *Handbook of medical image processing and analysis* (2nd ed., pp. 73–90). Academic Press. https://doi.org/10.1016/B978-012373904-9.50013-1

Vandenberg, O., Martiny, D., Rochas, O., van Belkum, A., & Kozlakidis, Z. (2020). Considerations for diagnostic COVID-19 tests. *Nature Reviews Microbiology*, *19*, 1–13. https://doi.org/10.1038/s41579-020-00461-z

Voulodimos, A., Protopapadakis, E., Katsamenis, I., Doulamis, A., & Doulamis, N. (2020). Deep learning models for COVID-19 infected area segmentation in CT images. *MedRxiv*.

Wang, L., Wang, Y., Ye, D., & Liu, Q. (2020). Review of the 2019 novel coronavirus (SARS-CoV-2) based on current evidence. *International Journal of Antimicrobial Agents*, *55*(6), 105948. https://doi.org/10.1016/j.ijantimicag.2020.105948

Wang, P., Chen, P., Yuan, Y., Liu, D., Huang, Z., Hou, X., & Cottrell, G. (2018). Understanding convolution for semantic segmentation. *IEEE Winter Conference on Applications of Computer Vision (WACV)*, *2018*, 1451–1460. https://doi.org/10.1109/WACV.2018.00163

Wang, Y., Zhang, Y., Liu, Y., Tian, J., Zhong, C., Shi, Z., Zhang, Y., & He, Z. (2020). Does non-COVID19 lung lesion help? Investigating transferability in COVID-19 CT image segmentation. *ArXiv Preprint ArXiv*, *2006*, 13877.

Yan, Q., Wang, B., Gong, D., Luo, C., Zhao, W., Shen, J., Shi, Q., Jin, S., Zhang, L., & You, Z. (2004). COVID-19 chest CT image segmentation—A deep convolutional neural network solution. *ArXiv*, *2020*, 10987. http://arxiv.org/abs/2004.10987

Yang, Q., Li, Y., Zhang, M., Wang, T., Yan, F., & Xie, C. (2020). Automatic segmentation of COVID-19 CT images using improved MultiResUNet. *Chinese Automation Congress (CAC)*, *2020*, 1614–1618. https://doi.org/10.1109/CAC51589.2020.9327668

Yang, X., He, X., Zhao, J., Zhang, Y., Zhang, S., & Xie, P. (2020). COVID-CT-dataset: A CT scan dataset about COVID-19. ArXiv:2003.13865 [Cs, Eess, Stat]. http://arxiv.org/abs/2003.13865

AUTHOR BIOGRAPHIES

Nour Eldeen M. Khalifa, Department of Information Technology, Faculty of Computers & Artificial Intelligence, Cairo University, Cairo 12613, Egypt

Gunasekaran Manogaran, University of California, Davis, California, USA; College of Information and Electrical Engineering, Asia University, Taichung, Taiwan

Mohamed Hamed N. Taha, Department of Information Technology, Faculty of Computers & Artificial Intelligence, Cairo University, Cairo 12613, Egypt

Mohamed Loey, Department of Computer Science, Faculty of Computers and Artificial Intelligence, Benha University, Benha 13518, Egypt

How to cite this article: Khalifa, N. E. M., Manogaran, G., Taha, M. H. N., & Loey, M. (2021). A deep learning semantic segmentation architecture for COVID-19 lesions discovery in limited chest CT datasets. *Expert Systems*, e12742. https://doi.org/10.1111/exsy.12742