Information Fusion Research (2024) Volume 2 Issue 1
doi: 10.59429/ifr.v2i1.6552

ARTS AND SCIENCE
ARTS AND SCIENCE PRESS PTE. LTD

RESEARCH ARTICLE

COVID-19 diagnosis based on CT scan image segmentation using multi-scale convolutional neural network

Riadh Ayachi[1*], Mouna Afif[1], Yahia Said[2], and Abdessalem Ben Abdelali[1]

[1] *Laboratory of Electronics and Microelectronics, Faculty of Sciences of Monastir, University of Monastir, Monastir, Tunisia*

[2] *Electrical Engineering Department, College of Engineering, Northern Border University, Arar, Saudi Arabia*

***Corresponding author:** Riadh Ayachi, riadh.ayachi@fsm.rnu.tn

ABSTRACT

The COVID-19 pandemic continues to pose significant health challenges worldwide. While traditional methods like Reverse Transcription Polymerase Chain Reaction (RT-PCR) are widely used for diagnosis, they often suffer from limitations in speed and accessibility. This chapter proposes an automatic COVID-19 diagnosis technique using CT scan image segmentation, offering a promising alternative for early detection. The proposed method leverages Convolutional Neural Networks (CNNs), enhanced with a multi-scale approach, to capture more detailed semantic features and achieve high-quality segmentation despite challenges such as low image resolution, varying infection characteristics, and limited training data. To address the data scarcity issue, various data augmentation techniques, including scale variation, translation, and rotation, were applied to enhance the model's performance. The proposed CNN was evaluated using the COVID-19 CT dataset, achieving an overall accuracy of 96.85%, with COVID-19 sensitivity at 93.31%, Common Pneumonia sensitivity at 92.93%, and a true negative rate of 97.82%. These results demonstrate the effectiveness and efficiency of the proposed method in detecting COVID-19 at an early stage, reducing the need for human intervention, and accelerating the diagnostic process. Compared to existing methods, this approach offers significant improvements in accuracy and reliability, making it a valuable tool in the ongoing fight against the pandemic.

Keywords: COVID-19; automatic diagnosis; multi-scale network; convolutional neural network; CT scan images segmentation

1. Introduction

The COVID-19 pandemic has created an urgent need for accurate and timely diagnostic methods to control the spread of the virus. Reverse Transcription Polymerase Chain Reaction (RT-PCR) has been the gold standard for diagnosing COVID-19 due to its high specificity in detecting viral RNA. However, RT-PCR tests are not without limitations; they often require specialized equipment, trained personnel, and can take several hours to days to yield results. Moreover, RT-PCR tests have been reported to have sensitivity rates ranging from 60% to 70%, leading to a significant number of false negatives, especially in the early stages of infection.

In contrast, chest Computed Tomography (CT) scans have emerged as a valuable tool for diagnosing

ARTICLE INFO
Received: 13 July 2024 | Accepted: 27 August 2024 | Available online: 9 September 2024

CITATION
Ayachi R, Afif1 M, Said Y, et al. COVID-19 Diagnosis Based on CT Scan Image segmentation using Multi-scale Convolutional Neural Network. *Information Fusion Research* 2024; 2(1): 6552.

COVID-19, particularly in cases where RT-PCR results are negative, but clinical symptoms suggest infection. CT scans can detect the characteristic lung abnormalities associated with COVID-19, such as ground-glass opacities and consolidations, even before symptoms fully develop. Several studies have demonstrated that CT scans have a higher sensitivity compared to RT-PCR, with some reports indicating sensitivity rates as high as 97%. However, the reliance on visual interpretation by radiologists introduces variability and potential delays, making automated image analysis critical.

CT scan image segmentation is crucial because it allows for the precise identification and quantification of infected regions within the lungs, enabling a more objective and consistent diagnosis. This is particularly important for monitoring disease progression and assessing the severity of the infection. However, the challenges of CT scan-based diagnosis include the low resolution and contrast of the images, which can obscure the visibility of early-stage abnormalities. Additionally, the variability in the appearance of COVID-19-related lung lesions, which can overlap with other pulmonary conditions such as common pneumonia, complicates accurate segmentation.

For more than a year, COVID-19 has been causing a wide health crisis in the world. Based on the statistics of the World Health Organization (WHO) [1], there are more than 126,359,540 cases with more than 2,769,473 deaths. Even after finding a vaccine, COVID-19 is still spreading and more cases have been confirmed. The gold standard for COVID-19 detection is RT-PCR. In clinical settings, nasopharyngeal and oropharyngeal swabs are routinely employed as nucleic acid screening samples. However, it has been discovered that detecting nucleic acid by swabbing the pharynx is prone to false negatives due to sample quality and the pharynx's low viral load [2]. Delaying treatment for individuals with clinical signs increases the risk of infection among medical workers. Misdiagnosis and increased risk of transmission are common with silent infections. CT is a crucial complement to RT-PCR because of its non-invasiveness, excellent resolution, and low noise [36]. Detecting infected cases was performed using RT-PCR. However, this method requires more than 24 hours to confirm the state of the suspected cases. Besides, the RT-PCR was not sensitive enough [2] to generate trusted results for early treatment of confirmed COVID-19 cases [3]. Since COVID-19 caused respiratory failure, it can be detected through CT scan images. this was a faster diagnosis solution to confirm suspected cases.

As illustrated in Fig. 1, it can clearly explain picture properties associated with COVID-19, such as ground-glass opacity (GGO) and lung consolidation. In the early stages of the illness, the extrapulmonary zone was mostly affected by several tiny patches. GGO emerged in numerous lung lobes once the illness developed, causing the interlobular septum to thicken and the appearance of the "crazy-paving sign." After progressing to the severe stage, both lungs displayed widespread lesions, mostly lung consolidation, along with lung structural distortion, bronchiectasis, pleural effusion, and other symptoms [37]. Accurate COVID-19 lesion segmentation is beneficial in assessing the degree of lung infection and disease development, and it serves as a basis for patient follow-up therapy. However, manual lesion labeling takes time, and accuracy is heavily influenced by clinicians' subjective knowledge and experience.

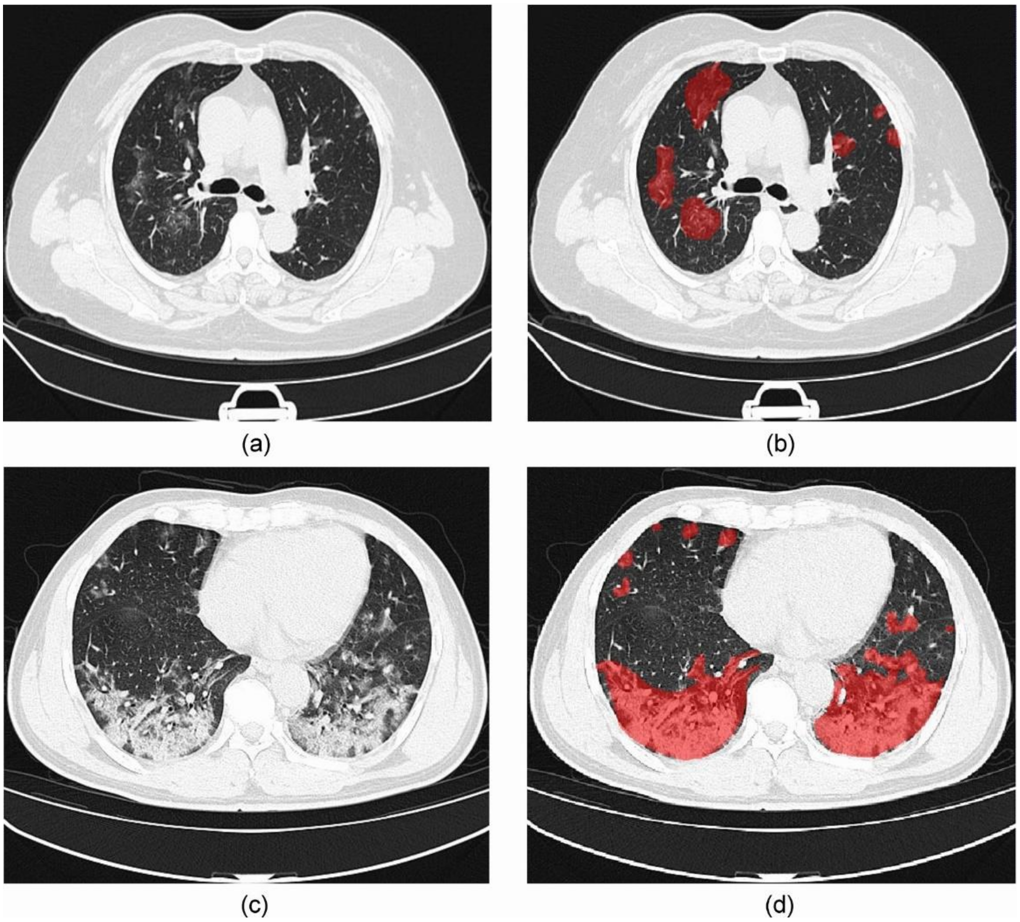

Figure 1: illustration of GGO and consolidation. a present a lung with multi GGO, b GGO lesions are covered with red, c represents a lung with consolidation, and d the consolidations and GGO lesions are covered with red

Considering CT scan images, a solution for fast COVID-19 diagnosis is still challenging due to the need for experts in medical imaging interpretation. For a case or two, it is feasible but for hundreds of cases, it is impossible to do the job perfectly. Besides, COVID-19 infection presents features similar to pneumonia which makes it harder to differentiate them. So, image segmentation was a promising idea to trait CT scan images faster with more accuracy and reduce human efforts. The CT scan image segmentation faces many challenges such as variation in position, texture, and size, low contrast of the Ground Glass Opacity (GGO) area, and noisy images. Besides, data availability is a major concern in building high-performance image segmentation algorithms. Two main factors are behind the lack of data. First, the CT scan images for suspected cases are not collected and stored for research purposes. Second, the annotation of the CT scan images is time-consuming and requires the presence of a medical expert to perform a perfect annotation.

The technique of separating one or more items from a given image based on their class is known as image segmentation. The fundamental concept behind this procedure is to create an image with different values for pixels within selected objects and pixels that are not selected. We will discuss both traditional and neural network-based segmentation approaches in this section. Furthermore, we will go through the benefits and drawbacks of each group. The traditional technique is region-based segmentation, which consists of fundamental image processing processes. Thresholding is the most often utilized process; it establishes one or two thresholds that serve as limits for the pixel values. The procedure then neutralizes pixels with values that are outside of the interval defined by the two thresholds. Regional growth segmentation is another well-known strategy. This algorithm begins by choosing a collection of pixels to serve as seeds. Iteration refers to the process of adding neighboring pixels to previously generated areas. The Region expanding algorithm iterates until it converges, or until all seeded regions cease growing. Edge-detection segmentation techniques might also be mentioned. The theory behind these strategies is based on the notion that every two neighboring segments have an edge. On a given input image, the edge segmentation approach employs a convolution process. The resulting image exhibits a strong contrast between the edges and the background. The major aspect of this procedure is the filter type. The Sobel [34]and Laplacian [35] operators are the most commonly utilized.

In the few past years, deep learning techniques have been considered a great solution for computer vision applications such as scene recognition [4], object detection [5], and image segmentation [6]. Particularly, deep learning techniques have been deployed successfully for medical image segmentation [7]. The main idea behind the high performance of deep learning techniques was the use of deep neural networks with the ability to self-learning by automatically extracting features from the input data. Moreover, the convolutional neural network (CNN) was the most used for image processing tasks due to its composition of convolution layers that enable the detection of complex patterns perfectly and the use of non-linear activation layers that allow to process of a huge amount of data without been overfitted and increase the learning capability of the network.

Despite the demonstrated potential of CT scans for COVID-19 diagnosis, existing methods for CT image analysis, particularly those relying on manual interpretation or traditional machine learning techniques, face significant limitations. Current approaches often struggle with the variability in the appearance of COVID-19-related lung lesions, which can overlap with other conditions such as common pneumonia, leading to misdiagnosis. Furthermore, many traditional methods are not robust against the inherent challenges of CT imaging, such as low resolution, poor contrast, and noise, which can obscure early-stage infection signs.

Another major limitation of existing methods is the reliance on large, well-annotated datasets for training. In the context of COVID-19, the availability of such datasets is limited due to the novelty of the virus and the rapid progression of the pandemic. This scarcity of training data hinders the performance of conventional models, making them less effective in clinical practice.

The specific gap in the literature that our proposed multi-scale CNN approach aims to address is the need for a more robust and accurate method for segmenting CT images that can overcome these challenges. Traditional CNN models, while effective in many image analysis tasks, often fail to capture the multi-scale nature of the features present in CT images of COVID-19 patients. These models may miss subtle but crucial details that are essential for early and accurate diagnosis.

The proposed multi-scale CNN approach addresses this gap by incorporating multiple scales into the analysis, allowing the model to capture both fine and coarse features simultaneously. This enhances the model's ability to distinguish between COVID-19 and other lung conditions, even in low-quality images, and reduces the dependence on large datasets by leveraging data augmentation techniques. As a result, our method provides a more reliable and accurate tool for COVID-19 diagnosis, offering significant improvements over existing methods in terms of robustness, accuracy, and generalizability.

Since CT scan images require more powerful techniques to handle the presented challenges, we proposed a multi-resolution CNN model that can detect fine-grained semantic features. The multi-resolution has been obtained through special pyramid pooling (SPP) [8]. Different from the original SPP, in the proposed network there is no need to fix the size of the features map. So, the generated feature maps were concatenated with a shortcut connection with its input and then used as input to the next layer. The proposed SPP was designed to build a multi-scale network able to collect features at different scales. This enables the network to achieve better segmentation performances.

CT scan images for COVID-19 cases are not available with a big amount which causes many problems especially to train CNN models. To enhance the performance of the proposed CNN, we proposed many data augmentation techniques to elevate the amount of training data. The data augmentation techniques enhance the generalization power of the model by presenting more challenging training examples. The proposed data augmentation techniques consist of flipping, translation, rotation, and contrast variation.

Since getting perfect annotated images is a hard challenge and most of the available annotations are noisy, we propose to use a custom loss function. The dice loss function [9] was designed to balance background and foreground classes.

The proposed CNN model was evaluated using the COVID-19 CT dataset [10]. the dataset was collected from different patients with and without COVID-19 infection. The dataset contains 6000 CT scan images divided into 3 sets, a training set, a validation set, and a testing set. The proposed CNN model achieved an overall accuracy of 96.85%, 93.31% of COVID-19 sensitivity, 92.93% of Common Pneumonia sensitivity, and 97.82% of the true negative rate.

The rest of the chapter is organized as follows: section 2 presents related works on CT scan images segmentation for COVID-19 detection. the proposed approach was described and detailed in section 3. Section 4 was reserved for pressing and discussing experimental results. In section 5, we present the conclusion.

2. Related works

The widespread of COVID-19 has made a need for an urgent collaboration between computer science engineers and medical staff to invent automatic diagnosis tools that reduce human effort and interaction in addition to producing more trusted results concerning suspected cases. CT scan image segmentation was a good solution for COVID-19 detection. So, many works were proposed to build a robust segmentation system. For a complete review of proposed methods for COVID-19 diagnoses, readers may refer to the review in [11].

Ter-Sarkisov [12] proposed CT scan image segmentation based on the mask region-based convolutional neural network (R-CNN) [13]. The Mask R-CNN was designed for object detection with the generation of a segmentation mask over the localized object. It is composed of two networks. The first is the backbone used for feature extraction and the second is the region proposal network (RPN) [14] used to generate region proposals. The output of the RPN was concatenated with the output of the backbone through a region of interest (RoI) pooling layer that fixes the size of the feature maps connected to the fully connected layers. to use the Mask R-CNN for a full segmentation, the bounding boxes prediction was eliminated and only the classification part was conserved. The proposed Mask R-CNN was trained and evaluated on the COVIDx-CT dataset and an overall accuracy of 91.66% was achieved.

A CT scan image segmentation based on a multi-task neural network was proposed in [15]. The proposed network was composed of inception recurrent residual layers. Those layers were a modified version of the inception proposed in GoogleNet [16] where residual connections were added in addition to memory units from recurrent neural networks. The proposed multi-task neural network was composed of an input layer followed by five inception recurrent residual layers, an average pooling layer, and a Softmax layer for output. The rectified linear units (ReLU) were used for activation layers. The proposed network was evaluated on a custom-made dataset with 300 images. The proposed network has achieved an accuracy of 84.67%.

A dual-branch combination network (DCN) [17] was proposed for computer-aided COVID-19 diagnosis through CT scan image segmentation. The proposed network was capable of segmenting CT scan images and classifying detected lesions simultaneously. To accomplish both tasks, the DCN was designed based on three subnetworks. First, the U-Net [18] was used for lung segmentation to detect lesions. Second, a classification network based on ResNet 50 [19] was used to classify lesions. Finally, an encoder-decoder convolutional neural network was used to segment lesions. To make the network focus on lesion classification a lesion attention module was proposed. The proposed network was evaluated using a collected dataset. CT scan images were collected from patients in a local hospital in Yueyang City, China, and other hospitals in the neighboring cities. The collected data were automatically preprocessed to extract lung area then a set of experienced radiologists manually annotated the data. A total of 6130 slices were used for training and 204265 slices from other hospitals were used for testing. An accuracy of 96.74% was achieved using the local testing data and an accuracy of 92.87% was achieved on the testing data from other hospitals.

Yang et al. [20] proposed a semi-supervised learning method for the segmentation of CT scan images to detect COVID-19. The proposed method consists of a 3D fully convolutional neural network (FCN) with skip connections between different levels of the feature extraction stage. The proposed method was based on the design of the U-Net [18]. To eliminate the problem of data variability across the world due to the difference in medical imaging equipment, a dataset was collected from China, Japan, and Italy to train and evaluate the proposed method. The best accuracy achieved at evaluation was 60.7%.

The U-net model [18] was enhanced with an attention mechanism and a boundary loss function for the segmentation CT scan imaged to detect COVID-19 infection [22]. The proposed method was used to develop an automatic segmentation system that detects infected regions in CT scan lung images. The developed system takes advantage of the U-Net model with additional improvements. An attention mechanism was proposed to increase the segmentation ability by extracting more relevant features. Besides, a boundary loss function was proposed to improve the accuracy of detecting small and imbalanced infected regions. The proposed system was evaluated on three public datasets [23-25]. A sensitivity of 0.73 was achieved.

Suri et al. [26] evaluated state-of-the-art segmentation models for CT scan image segmentation to detect COVID-19 infection. Many models such as ResNet-SegNet, VGG-SegNet, PSPNet [27], VGG-UNet, and ResNet-UNet. A dataset composed of 3000 images was collected from 40 positive cases. Experimental results proved that ResNet-UNet has the best accuracy compared to other models. A dice value of 0.77 was achieved as the best performance of the ResNet-UNet model.

For detecting COVID-19 infections, a COVID-19 hierarchical segmentation network was proposed in [28]. The proposed network follows the U-Net style. The network is composed of cascaded residual attention inception modules that work in encoder-decoder methodology. Also, spectral-spatial and depth attention network was integrated to enhance the overall performance of the segmentation model. This network combines depthwise separable convolution with expansion and a mix of spectral and max pooling layers. this technique was developed to reduce information loss in the encoding-decoding process. An average binary cross entropy loss function with additional dice loss was deployed to reduce the false negative and false positive rates. An accuracy of 0.965 was achieved on a custom-made dataset.

Amara et al. [29] proposed a virtual diagnosis platform for detecting early COVID-19 infections from CT scan images. The proposed model is based on the combination of convolutional autoencoders with channels upsampling and downsampling. The proposed network was designed to require less storage memory and computation effort. The input image was processed with two convolutional autoencoders in a parallel manner and the outputs were combined to generate the final prediction. Good results were achieved with a dice score of 0.764.

The COVID-Rate framework [30] was proposed to automatically segment CT scan images and to detect COVID-19 infection. The proposed framework was based on the U-Net model [18] with unsupervised learning. The main architecture was improved through dilated residual blocks and variant kernel sizes in the encoder. This improvement enables feature extraction from variable receptive fields. Besides, a squeeze and expansion module was deployed to enhance model generalization and recalibrate channel-wise feature maps. All max pooling layers were replaced with convolution layers with a stride of 2 to reduce information loss at the decoding stage [31]. The proposed model was trained using a hybrid loss function that combines focal Tversky loss [32] and weighted binary cross entropy loss [33]. A dice score of 0.8069 was achieved which proved the efficiency of the proposed model.

The study in [38] suggested creating a new application for COVID-19 segmentation and analysis using deep learning. A neural network that uses context aggregation as its foundation was used to build the suggested system. The context fuse model (CFM), attention mix module (AMM), and residual convolutional module (RCM) are the three primary components of this network. Ground glass opacity and consolidation area are two primary COVID-19-related areas that the created system can identify in CT scans. Common pneumonia and instances of COVID-19 are typically associated with these lesions. The COVID-x-CT dataset has been used in both training and testing trials. In comparison to state-of-the-art performances, the created system showed superior and more competitive outcomes, according to the acquired results. The numerical results show that the suggested work is effective since it achieves a level of accuracy that is more than 96.23%.

SCTV-UNet [39] was proposed as a solution to address these issues. More visual layer information is gained to identify the normal pixels between nearby sick regions by merging spatial and channel attentions on the encoder. The issue of poor contrast and hazy boundaries produced by BCE in conventional U-shaped networks is addressed by employing the composite function DTVLoss, which zeroes in on the pixels inside the affected region. Based on the results of the experiment, the suggested SCTV-UNet can be a useful tool for clinical COVID-19 detection and study, since its segmentation impact is much enhanced for COVID-19 segmentation.

Newson et al. [40] proposed to autonomously segmenting COVID-19 CT data using an Encoder-Decoder convolutional neural network (ED-CNN) model. This way, a model that is different from what's already out there was created; it's simple to understand and replicate, which means it will be easier to apply in the real world and won't need as much training. It was shown that thoracic CT automated segmentation prediction can accurately identify infected lung areas. When peer review is not an option, this segmentation automation can speed up the contouring process by verifying manual contouring instead. It can also quickly indicate infection so patients can be sent for additional treatment, saving time and money. On average, competing models employ roughly 1,000 times more parameters than the suggested model's 49 k. Shorter training durations are achieved by our method due to its reliance on a relatively compact model. This allows for easy retraining of the model using additional data and might lead to the affordable implementation of "personalized medicine" procedures.

3. Proposed Approach

In this section, we present and detail the proposed CNN model used for CT scan image segmentation. The proposed CNN was designed to collect fine-grained features from multi-scale layers to achieve better performance. To build the multi-scale network, we proposed to use an SPP block with residual connections instead of pooling layers. the SPP is composed of many max-pooling layers at different scales. A part of each max-pooling is used to generate the final output. Then, the outputs were concatenated with input through residual connection and used as input to the next layer. The proposed SPP is presented in figure 2. The SPP conserves the spatial information while pooling layers lose the spatial information and its implantation for segmentation tasks is not desirable. The proposed SPP has a skip connection between the input and output to transfer high-resolution semantic information to the top of the network. The use of the SPP for the multi-scale network is more stable in the training process and results in better accuracy. The proposed SPP allows to use of input images with different resolutions without any modification to the network and getting good results.

Lung anomalies caused by COVID-19 can range from very modest to quite variable from patient to patient. By effectively identifying both tiny and big lesions, the multi-scale method improves the CNN's capability to detect these alterations. This is vital for diagnosing the disease at an early stage, when it may only cause subtle, imperceptible alterations in the lung tissue.

It can be challenging to extract significant features using single-scale approaches from CT scans due to difficulties such as low resolution, noise, and poor contrast. To get around this, the multi-scale method lets the CNN make up for low-quality images at one scale with higher-quality, more informative features at another. This results in segmentation that is more stable, even when the imaging circumstances are not ideal. Because COVID-19 symptoms can manifest in a wide variety of ways, it might be difficult for a model trained on a specific dataset to apply its findings to novel instances. By training the CNN to identify and segment lesions of varying sizes and shapes, the multi-scale method improves generalizability and makes the network more responsive to a wide range of patient demographics and disease presentations. Accurately segmenting contaminated areas in CT scans is made easier using the multi-scale CNN, which extracts and integrates data at several scales. This results in more accurate disease burden estimation, which is critical for determining the seriousness of infection, informing treatment choices, and tracking the development of illness.

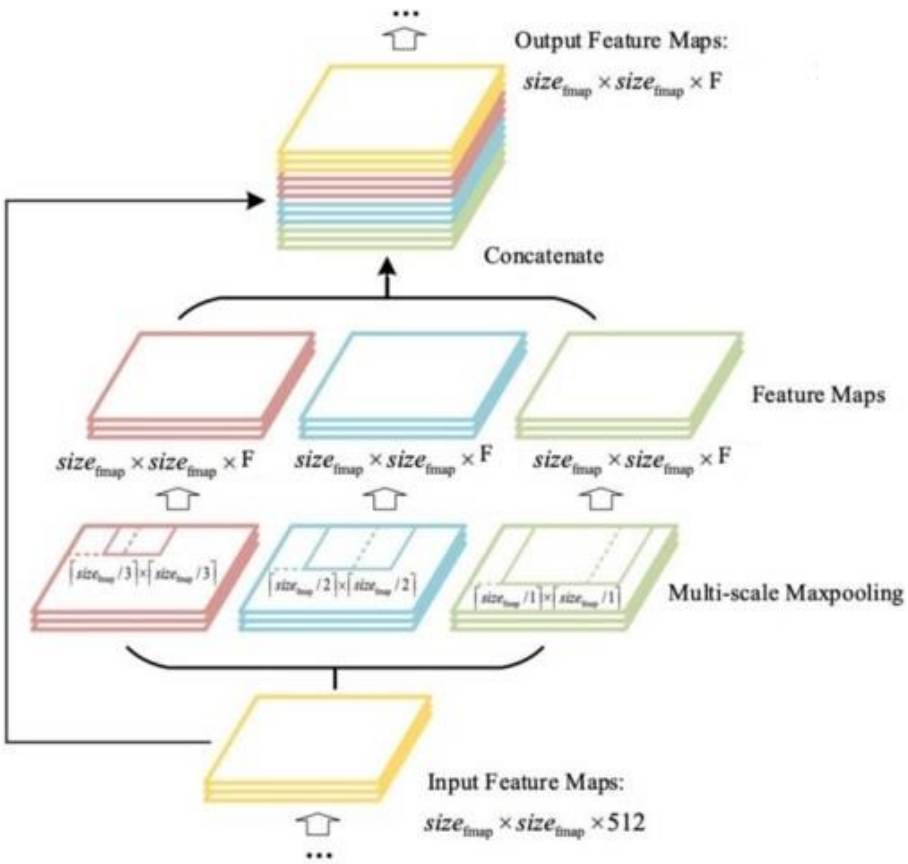

Figure 2: Proposed spatial pyramid pooling

For the segmentation of CT scan images, we proposed a multi-scale CNN model following the U-Net [18] fashion style due to its efficiency for medical image segmentation. The proposed CNN is composed of four main blocks. First, the SPP block was used to generate the different resolutions in the network Second, the very resolution stage starts and ends with a multi-scale block. Third, for the generated resolution stages, an upsampling block is used to enable the concatenation of the stage output with higher resolution stages. Finally, residual blocks were used to extract features at different resolution stages. Figure 3 illustrates the proposed multi-scale CNN with different blocks. For all blocks, we proposed to use the Scaled Exponential Linear Units (SELU) [21] as an activation function that helps to stabilize the training process and enhances the overall performance. The SELU has proved its efficiency by allowing the training of neural networks with few layers perfectly and achieving high performance through self-regularization without the need for regularization layers. besides, it eliminates the problem of vanishing and exploding gradients.

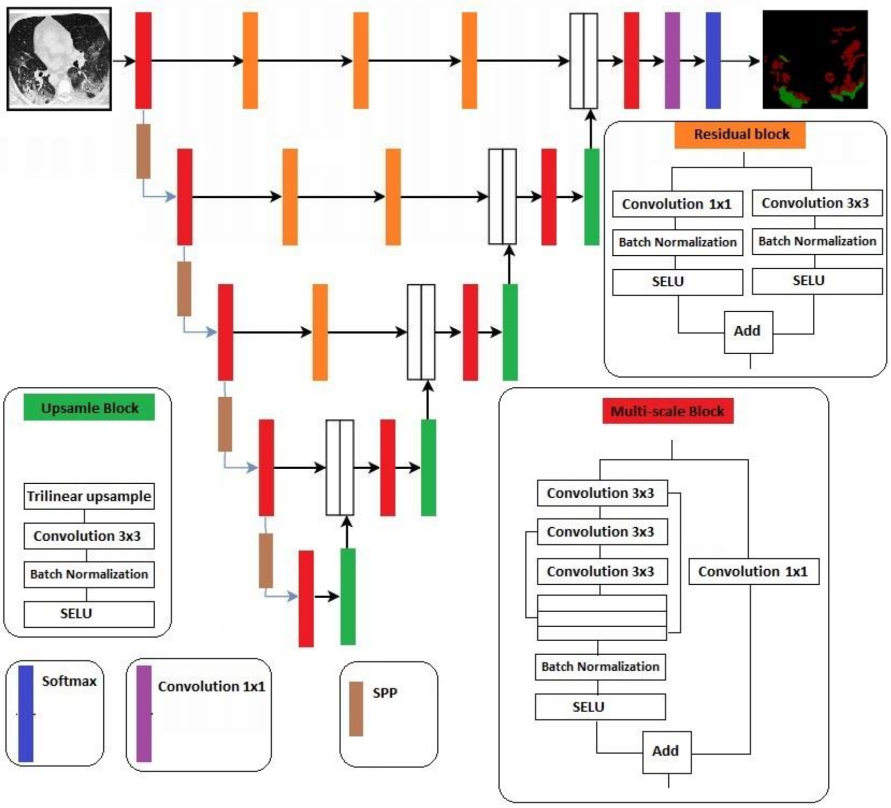

Figure 3: Proposed CNN model for CT scan image segmentation

The residual block is composed of convolution layers with 3x3 and 1x1 kernel sizes followed by a match normalization layer and an activation layer. The path with a 1x1 convolution kernel was used to reduce the computation complexity. The upsampling block is composed of a bilinear upsample followed by a 3x3 convolution layer, a batch normalization layer, and an activation layer. The multi-scale block is composed of 3 convolution layers with a kernel size of 3x3 concatenated in inverse order and followed by a batch normalization layer and an activation layer. A residual connection through a 1x1 convolution layer was used to reduce computation complexity and avoid vanishing gradient problems.

After collecting features from different resolutions, a 1x1 convolution layer was applied followed by a softmax layer for output generation. The output represents the density of each class in the input image. To compute the loss, we proposed to use a dice loss [9] instead of a cross-entropy loss. The use of the dice loss for segmentation problems is preferred due to its ability to balance the background/foreground predictions. For medical image segmentation especially CT scan images, the region of interest occupies a tiny size in the image. So, the model training will be stuck in a local minimum and most of its prediction will be mapped to the background. Thus, will decrease the performance y missing of partially detecting the region of interest. The main focus was on designing a loss function that gives the foreground more importance compared to the background. The dice loss was proposed as a solution to balance the predictions and increase the foreground sensitivity. The dice loss computes the intersection and union over the foreground. If the model detects all pixels as background, the intersection will be zero and the error will be raised to its maximum which is one, and if the model detects all pixels as foreground, then the union will be too large and the loss goes to one. The dice loss tries to minimize the error and the loss simultaneously by maximizing the intersection and minimizing the union over the foreground. The dice loss can be computed as 1

$$L_{dice} = 1 - \frac{1}{C-1}\sum_{c=2} \frac{\sum_{n=1}^{N} 2 \cdot \hat{y}_{c,n} \cdot y_{c,n} + \epsilon}{\sum_{n=1}^{N} \hat{y}_{c,n} + y_{c,n} + \epsilon} \tag{1}$$

Where C is the number of classes which are two in this case. $\hat{y}_{c,n}$ is the predicted label of the pixel n and $y_{c,n}$ is the ground-truth label of pixel n. The background class was not considered in the dice loss since it has no importance in the studied task and makes the model focus only on the foreground to gain more performance.

4. Experiments and results

For all training and evaluation of the Proposed approach, a desktop running a Linux operating system equipped with an Intel i7 CPU, 32 GB of RAM, and Nvidia GTX960 GPU was used. The proposed CNN model was developed based on the TensorFlow deep learning framework with GPU support through the CUDA and cuDNN libraries. The open cv library was used for the load and display of the images.

4.1 Dataset

To train the proposed CNN model, we proposed the use of a CT scan dataset [10] collected in China. The dataset consists of 617775 CT scan images from 3777 patients. The collected images contain different pneumonia cases such as viral pneumonia, bacterial pneumonia, mycoplasma pneumonia, and COVID-19 pneumonia. The main challenge was the ability to differentiate COVID-19 pneumonia from other common pneumonia. Only 4695 CT scan images were manually labeled for the segmentation task. The rest of the images were used for a classification task which was not investigated in this work. The main focus of this work is on the CT scan image segmentation task. The dataset was divided into a training set and a testing set where 70% was used for training and 30% was used for testing.

4.2 Training

To increase the amount of training data and to enhance the generalization power of the model, data augmentation techniques were applied. Data augmentation has many benefits including avoiding the overfitting problem by presenting many variations of the input data. First, a data augmentation technique based on random scale variation was applied. The scaling factor was varied between 0.8 and 1.2. this technique was considered as aspect ratio data augmentation due to the independent sampling for both the x- and y-axis. Second, a data augmentation technique based on random flipping on the x-axis was applied. The images were flipped to present more challenging conditions and train the model to detect harder samples. Third, the image cropping technique was used as data augmentation by using 70% and 50% random crop of the image as new data. Fourth, rotation with different angles and translations were applied. Finally, a contrast variation was performed to make the model robust against the difference in contrast caused by imaging tools.

Techniques like scale variation and translation expose the model to features at different sizes and positions, helping it learn to recognize lesions even when they appear small or blurry due to low resolution. Rotation and flipping further ensure that the model is not dependent on any specific orientation, making it more resilient to variations in image quality. Brightness and contrast adjustment directly simulate the conditions of low contrast in CT images, teaching the model to identify features even when they are not distinguishable from the surrounding tissues. This is crucial for accurately segmenting COVID-19-related lung abnormalities, which might be faint or diffuse in some scans.

The Adam optimizer was used for loss optimization with an initial learning rate value of 0.0001, a weight decay of 0.0001, epsilon was fixed to 0.0000001, beta_1 = 0.9, and beta_2 =0.999. The learning rate was optimized through multiplication by a factor of 0.95 every 5 epochs and the weight decay was regularized using a decoupled weight decay regularization [21]. Optimizing the learning rate has allowed for stabilizing the training in the final epochs. The proposed model was trained for 50 epochs each with 1000 iterations.

4.3 Evaluation

The proposed CNN model was evaluated using many evaluation metrics for a fair comparison against state-of-the-art models. The overall accuracy was used to determine the performance of the model in the training. The COVID-19 sensitivity is used for evaluating the model performance when detecting CIVID-19 cases. The Common Pneumonia sensitivity is the rate of detecting pneumonia with no COVID-19 infection. The true negative rate refers to the percentage of true negative predictions. The dice score was used to measure the difference between the predicted segmentation and the ground truth. The proposed CNN model achieved an overall accuracy of 96.85%, 93.31% of COVID-19 sensitivity, 92.93% of Common Pneumonia sensitivity, 97.82% of the true negative rate, and a dice score of 0.95. The achieved results proved the efficiency of the proposed multi-scale CNN model for the segmentation of CT scan images. to further improve the performances of the proposed model, we consider a comparison against state-of-the-art models for CT scan image segmentation for COVID-19 detection. **Table 1** presents a comparison against the most recent models for the segmentation of CT scan images.

Table 1: comparison against the most recent model for the segmentation of CT scan images for COVID-19 detection

Model	**Overall accuracy (%)**	**COVID-19 sensitivity (%)**	**Pneumonia sensitivity (%)**	**True negative rate (%)**
COVID-CT-Mask-Net [12]	91.66	90.80	91.62	-
COVID_MTNet [15]	94.52	94.66	-	-
DCN [17]	95.99	93.59	89.14	97.55
Multi-scale CNN (ours)	96.85	93.31	92.93	97.82

The proposed model has outperformed the most recent model for the segmentation of CT scan images for COVID-19 detection in terms of accuracy and sensitivity. Based on the reported results, the proposed model is very effective for the studied task. The visual output of the proposed multi-scale CNN is presented in **figure 4.**

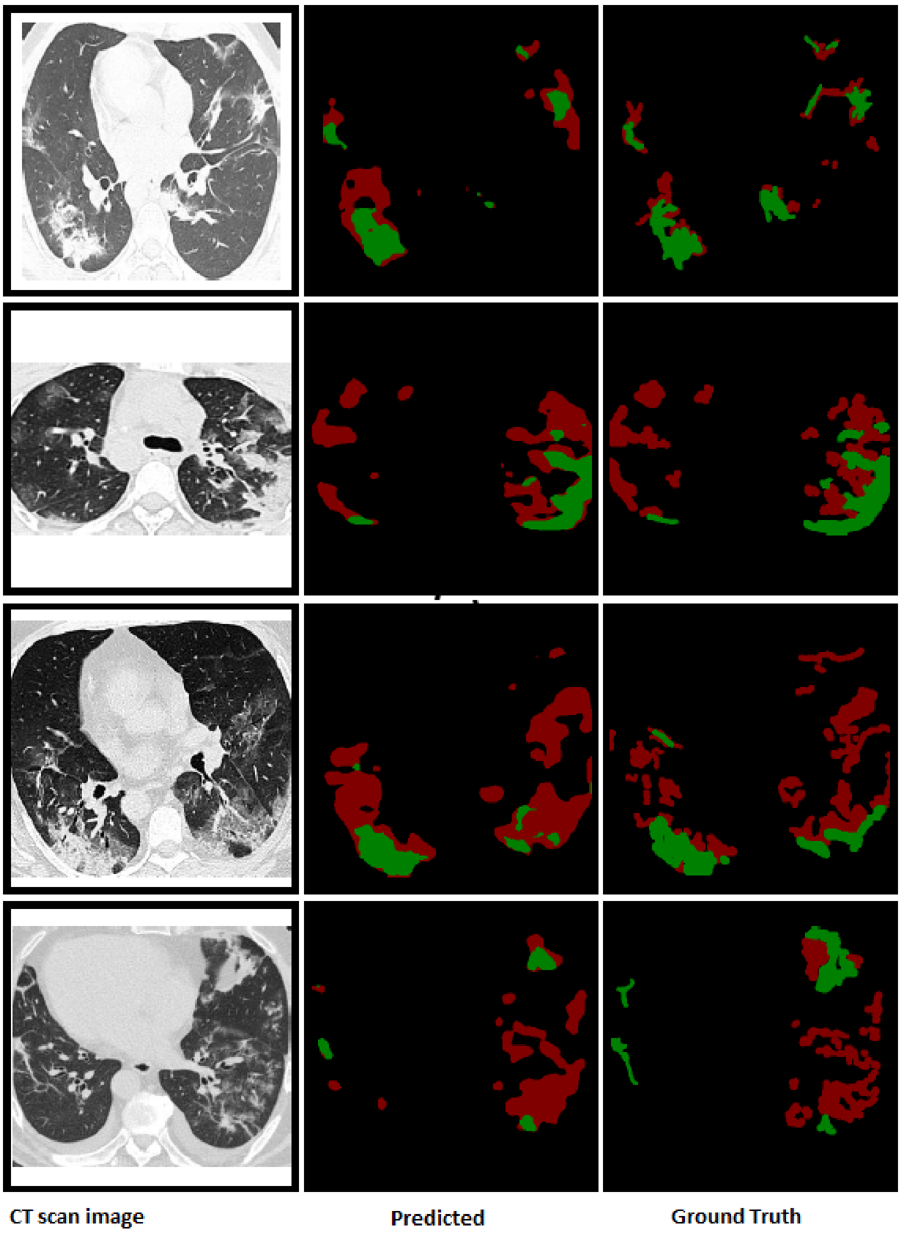

Figure 4: Visual output of the segmentation model, where the red color refers to the GGO class and the green color refers to the consolidation

4.4 Discussion

The main goal of this work was to build a powerful approach for the segmentation of CT scan images to detect COVID-19 infections. The use of a multi-scale CNN model was very effective in achieving the desired goal. As reported earlier high performance was achieved which outperforms the performance of most recent approaches for the same task.

The use of the SPP layers has improved the accuracy by enabling the reconstruction of semantic features without losing spatial information. compared to the use of pooling layers, the SPP is much more efficient spatial for the segmentation task. The proposed multi-scale block has helped to extract more relevant features while keeping the low computation complexity. The upsample module was very important to manipulate the feature map sizes without damaging the spatial information. The use of the SELU activation function has improved the accuracy compared to the rectified linear unit (ReLU) function.

In order to better segment CT images of lung anomalies due to COVID-19, this work presents a new multi-scale Convolutional Neural Network (CNN). In contrast to conventional CNNs, which only analyze pictures at one scale, the suggested multi-scale method enables the network to concurrently collect coarse and fine characteristics. The wide variety of COVID-19 symptoms, from faint ground-glass opacities to more noticeable consolidations, makes this skill crucial for detection. The approach improves diagnostic accuracy by achieving more accurate and thorough segmentation by combining information from several scales.

Since CT images can have wildly varying quality, the multi-scale CNN architecture aims to fix issues like low contrast and resolution. With the capability to handle pictures of various sizes, the model becomes more resistant to these changes, guaranteeing accurate identification of COVID-19 in various imaging scenarios. Since the quality of CT scans might vary in real-world clinical situations, this improvement greatly enhances the model's practical usefulness.

Scale variation, translation, rotation, flipping, and brightness/contrast alteration are only a few of the data augmentation strategies utilized in this study, which acknowledges the scarcity of well-annotated CT datasets for COVID-19. In order to improve the CNN's ability to generalize to new, unknown situations, these strategies artificially enhance the variety of the training data. In the early phases of a pandemic, when data gathering is still occurring, models generally suffer owing to insufficient training data. This contribution addresses a significant gap in the current literature by addressing this issue.

With a high sensitivity (93.31%), true negative rate (97.82%), and accuracy (96.85%) on the COVID-19 CT dataset, the suggested method significantly improves diagnostic performance. When compared to previous approaches, the multi-scale CNN with considerable data augmentation provides superior accuracy and dependability. This is especially true when it comes to differentiating between COVID-19, common pneumonia, and healthy patients.

Radiologists still need to manually interpret CT scans, which may be a time-consuming and inaccurate process; a new study automates that process. Clinical choices may be made more quickly and with more confidence using the suggested strategy since it both shortens the diagnostic procedure and makes COVID-19 detection more consistent.

Although this study primarily focuses on COVID-19, the data augmentation methodologies and multi-scale CNN architecture offered here have wider applications in medical image analysis. Contributing to the improvement of automated diagnosis tools in general, the methodologies described here can be extended to different illnesses and imaging modalities.

5. Conclusions

The COVID-19 pandemic has exacerbated global health crises due to its rapid transmission and infection rates. While CT scan imaging offers a quicker and more reliable diagnostic alternative to RT-PCR tests, manual segmentation of the vast number of CT images is impractical and labor-intensive. This chapter introduces an automated approach for CT scan image segmentation using a multi-scale Convolutional Neural Network (CNN) model. The model features a CNN architecture with multiple resolution stages to capture detailed features across different scales, and incorporates a Spatial Pyramid Pooling (SPP) layer to enhance the multi-resolution capabilities of the network. Evaluated on a publicly available dataset, the proposed multi-scale CNN model demonstrated exceptional performance, achieving an overall accuracy of 96.85%, a COVID-19 sensitivity of 93.31%, a Common Pneumonia sensitivity of 92.93%, and a true negative rate of 97.82%, along with a dice score of 0.95. These results affirm the model's effectiveness in accurately segmenting CT images for early COVID-19 detection. The proposed method significantly streamlines the segmentation process, reducing manual labor and expediting diagnosis, which is crucial for timely and accurate COVID-19 detection in clinical settings. Despite its strong performance, the model's evaluation was limited to a single publicly available dataset, which may restrict its applicability to other contexts or patient populations. Additionally, while the data augmentation techniques used were effective, they do not fully address all potential variations encountered in real-world scenarios. Future work should focus on expanding and diversifying datasets to enhance the model's generalizability, integrating this segmentation approach with other diagnostic tools for improved accuracy, and validating the model in diverse clinical settings. Additionally, exploring advanced data augmentation methods or alternative learning techniques could further improve performance and address data limitations.

Acknowledgment: Acknowledgments: The authors extend their appreciation to the Deanship of Scientific Research at King Khalid University for funding this work through Small Groups. (Project under grant number 367/43).

Conflict of interest

The authors declare no conflict of interest.

References

1. WHO Coronavirus (COVID-19) Dashboard. Available at: https://covid19.who.int/ last accessed 01/04/2021.
2. Ai, Tao, Zhenlu Yang, Hongyan Hou, Chenao Zhan, Chong Chen, Wenzhi Lv, Qian Tao, Ziyong Sun, and Liming Xia. "Correlation of chest CT and RT-PCR testing for coronavirus disease 2019 (COVID-19) in China: a report of 1014 cases." Radiology 296, no. 2 (2020): E32-E40.
3. Fang, Yicheng, Huangqi Zhang, Jicheng Xie, Minjie Lin, Lingjun Ying, Peipei Pang, and Wenbin Ji. "Sensitivity of chest CT for COVID-19: comparison to RT-PCR." Radiology 296, no. 2 (2020): E115-E117.
4. Afif, Mouna, Riadh Ayachi, Yahia Said, and Mohamed Atri. "Deep learning based application for indoor scene recognition." Neural Processing Letters (2020): 1-11.
5. Ayachi, Riadh, Yahia Said, and Mohamed Atri. "A Convolutional Neural Network to Perform Object Detection and Identification in Visual Large-Scale Data." Big Data 9, no. 1 (2021): 41-52.
6. Minaee, Shervin, Yuri Y. Boykov, Fatih Porikli, Antonio J. Plaza, Nasser Kehtarnavaz, and Demetri Terzopoulos. "Image segmentation using deep learning: A survey." IEEE Transactions on Pattern Analysis and Machine Intelligence (2021).
7. Renard, Félix, Soulaimane Guedria, Noel De Palma, and Nicolas Vuillerme. "Variability and reproducibility in deep learning for medical image segmentation." Scientific Reports 10, no. 1 (2020): 1-16.
8. He, Kaiming, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. "Spatial pyramid pooling in deep convolutional networks for visual recognition." IEEE transactions on pattern analysis and machine intelligence 37, no. 9 (2015): 1904-1916.
9. Milletari, Fausto, Nassir Navab, and Seyed-Ahmad Ahmadi. "V-net: Fully convolutional neural networks for volumetric medical image segmentation." In 2016 fourth international conference on 3D vision (3DV), pp. 565-571. IEEE, 2016.
10. Kang Zhang, Xiaohong Liu, Jun Shen, Zhihuan Li, Ye Sang, Xingwang Wu, Yunfei Zha, Wenhua Liang, Chengdi Wang, Ke Wang, Linsen Ye, Ming Gao, Zhongguo Zhou, Liang Li, Jin Wang, Zehong Yang, Huimin Cai, Jie Xu, Lei Yang, Wenjia Cai, Wenqin Xu, Shaoxu Wu, Wei Zhang, Shanping Jiang, Lianghong Zheng, Xuan Zhang, Li Wang, Liu Lu, Jiaming Li, Haiping Yin, Winston Wang, Oulan Li, Charlotte Zhang, Liang Liang, Tao Wu, Ruiyun Deng, Kang Wei, Yong Zhou, Ting Chen, Johnson Yiu-Nam Lau, Manson Fok, Jianxing He, Tianxin Lin, Weimin Li, Guangyu Wang,. Clinically applicable ai system for accurate diagnosis, quantitative measurements, and prognosis of covid-19 pneumonia using computed tomography. Cell, 2020.
11. Shi, Feng, Jun Wang, Jun Shi, Ziyan Wu, Qian Wang, Zhenyu Tang, Kelei He, Yinghuan Shi, and Dinggang Shen. "Review of artificial intelligence techniques in imaging data acquisition, segmentation and diagnosis for covid-19." IEEE reviews in biomedical engineering (2020).
12. Ter-Sarkisov, Aram. "Covid-ct-mask-net: Prediction of covid-19 from ct scans using regional features." medRxiv (2020).
13. He, Kaiming, Georgia Gkioxari, Piotr Dollár, and Ross Girshick. "Mask r-cnn." In Proceedings of the IEEE international conference on computer vision, pp. 2961-2969. 2017.
14. Ren, Shaoqing, Kaiming He, Ross Girshick, and Jian Sun. "Faster r-cnn: Towards real-time object detection with region proposal networks." arXiv preprint arXiv:1506.01497 (2015).
15. Alom, Md Zahangir, M. M. Rahman, Mst Shamima Nasrin, Tarek M. Taha, and Vijayan K. Asari. "Covid_mtnet: Covid-19 detection with multi-task deep learning approaches." arXiv preprint arXiv:2004.03747 (2020).
16. Szegedy, Christian, Vincent Vanhoucke, Sergey Ioffe, Jon Shlens, and Zbigniew Wojna. "Rethinking the inception architecture for computer vision." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 2818-2826. 2016.
17. Kai Gao, Jianpo Su, Zhongbiao Jiang, Ling-Li Zeng, Zhichao Feng, Hui Shen, Pengfei Rong, Xin Xu, Jian Qin, Yuexiang Yang, Wei Wang, Dewen Hu,. "Dual-branch combination network (DCN): Towards accurate diagnosis and lesion segmentation of COVID-19 using CT images." Medical image analysis 67 (2021): 101836.
18. Ronneberger, Olaf, Philipp Fischer, and Thomas Brox. "U-net: Convolutional networks for biomedical image segmentation." In International Conference on Medical image computing and computer-assisted intervention, pp. 234-241. Springer, Cham, 2015.
19. He, Kaiming, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. "Deep residual learning for image recognition." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 770-778. 2016.
20. Dong Yang, Ziyue Xu, Wenqi Li, Andriy Myronenko, Holger R. Roth, Stephanie Harmon, Sheng Xu, Baris Turkbey, Evrim Turkbey, Xiaosong Wang, Wentao Zhu, Gianpaolo Carrafiello, Francesca Patella, Maurizio Cariati, Hirofumi Obinata, Hitoshi Mori, Kaku Tamura, Peng An, Bradford J. Wood, Daguang Xu,. "Federated semi-supervised learning for COVID region segmentation in chest CT using multi-national data from China, Italy, Japan." Medical image analysis 70 (2021): 101992.
21. Loshchilov I, Hutter F Decoupled weight decay regularization. In: seventh international conference on learning representations (ICLR). Ernest N. Morial Convention Center, New Orleans, USA, (2019).

22. Ahmed, Imran, Abdellah Chehri, and Gwanggil Jeon. "A Sustainable Deep Learning-Based Framework for Automated Segmentation of COVID-19 Infected Regions: Using U-Net with an Attention Mechanism and Boundary Loss Function." Electronics 11, no. 15 (2022): 2296.
23. COVID-19 CT scans from Italian Society of Medical and Interventional Radiology. Available at: https://sirm.org/category/senza-categoria/covid-19/. Last accessed: 02/13/2023.
24. COVID-19 pneumonia. Available at: https://radiopaedia.org/playlists/25887/. Last accessed: 02/13/2023.
25. COVID-19 CT segmentation dataset. Available at: http://medicalsegmentation.com/covid19/. Last accessed: 02/13/2023.
26. Suri, Jasjit S., Sushant Agarwal, Gian Luca Chabert, Alessandro Carriero, Alessio Paschè, Pietro SC Danna, Luca Saba et al. "COVLIAS 1.0 Lesion vs. MedSeg: An Artificial Intelligence Framework for Automated Lesion Segmentation in COVID-19 Lung Computed Tomography Scans." Diagnostics 12, no. 5 (2022): 1283.
27. Zhao, Hengshuang, Jianping Shi, Xiaojuan Qi, Xiaogang Wang, and Jiaya Jia. "Pyramid scene parsing network." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 2881-2890. 2017.
28. Punn, Narinder Singh, and Sonali Agarwal. "Chs-net: A deep learning approach for hierarchical segmentation of covid-19 via ct images." Neural Processing Letters 54, no. 5 (2022): 3771-3792.
29. Amara, Kahina, Ali Aouf, Hoceine Kennouche, A. Oualid Djekoune, Nadia Zenati, Oussama Kerdjidj, and Farid Ferguene. "COVIR: A virtual rendering of a novel NN architecture O-Net for COVID-19 Ct-scan automatic lung lesions segmentation." Computers & Graphics 104 (2022): 11-23.
30. Enshaei, Nastaran, Anastasia Oikonomou, Moezedin Javad Rafiee, Parnian Afshar, Shahin Heidarian, Arash Mohammadi, Konstantinos N. Plataniotis, and Farnoosh Naderkhani. "COVID-rate: an automated framework for segmentation of COVID-19 lesions from chest CT images." Scientific Reports 12, no. 1 (2022): 1-18.
31. Ayachi, Riadh, Mouna Afif, Yahia Said, and Mohamed Atri. "Strided convolution instead of max pooling for memory efficiency of convolutional neural networks." In Proceedings of the 8th International Conference on Sciences of Electronics, Technologies of Information and Telecommunications (SETIT'18), Vol. 1, pp. 234-243. Springer International Publishing, 2020.
32. Abraham, Nabila, and Naimul Mefraz Khan. "A novel focal tversky loss function with improved attention u-net for lesion segmentation." In 2019 IEEE 16th international symposium on biomedical imaging (ISBI 2019), pp. 683-687. IEEE, 2019.
33. Afshar, Parnian, Anastasia Oikonomou, Farnoosh Naderkhani, Pascal N. Tyrrell, Konstantinos N. Plataniotis, Keyvan Farahani, and Arash Mohammadi. "3D-MCN: a 3D multi-scale capsule network for lung nodule malignancy prediction." Scientific reports 10, no. 1 (2020): 1-11.
34. Yao, Yuqin. "Image segmentation based on Sobel edge detection." In 2016 5th International Conference on Advanced Materials and Computer Science (ICAMCS 2016), pp. 141-144. Atlantis Press, 2016.
35. Casaca, Wallace, Luis Gustavo Nonato, and Gabriel Taubin. "Laplacian coordinates for seeded image segmentation." In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 384-391. 2014.
36. Fang, Yicheng, Huangqi Zhang, Jicheng Xie, Minjie Lin, Lingjun Ying, Peipei Pang, and Wenbin Ji. "Sensitivity of chest CT for COVID-19: comparison to RT-PCR." Radiology 296, no. 2 (2020): E115-E117.
37. Zheng, Bingbing, Yaoqi Liu, Yu Zhu, Fuli Yu, Tianjiao Jiang, Dawei Yang, and Tao Xu. "MSD-Net: Multi-scale discriminative network for COVID-19 lung infection segmentation on CT." Ieee Access 8 (2020): 185786-185795.
38. Afif, Mouna, Riadh Ayachi, Yahia Said, and Mohamed Atri. "Deep learning-based technique for lesions segmentation in CT scan images for COVID-19 prediction." Multimedia Tools and Applications 82, no. 17 (2023): 26885-26899.
39. Liu, Xiangbin, Ying Liu, Weina Fu, and Shuai Liu. "SCTV-UNet: a COVID-19 CT segmentation network based on attention mechanism." Soft Computing (2023): 1-11.
40. Newson, Kiri S., David M. Benoit, and Andrew W. Beavis. "Encoder-decoder convolutional neural network for simple CT segmentation of COVID-19 infected lungs." PeerJ Computer Science 10 (2024): e2178.
41. Moosavi, Abdoulreza S., Ashraf Mahboobi, Farzin Arabzadeh, Nazanin Ramezani, Helia S. Moosavi, and Golbarg Mehrpoor. "Segmentation and classification of lungs CT-scan for detecting COVID-19 abnormalities by deep learning technique: U-Net model." Journal of Family Medicine and Primary Care 13, no. 2 (2024): 691-698.