Foundation University Journal of Engineering and Applied Sciences

FUJEAS
Vol. 4, Issue 2, 2023.
DOI:10.33897/fujeas.v4i2.782

Research Article

COVID-19 Lungs CT Scan Lesion Segmentation

Muhammad Shariq Shoaib *, Zobia Suhail

Department of Computer Science, Faculty of Computing and Information Technology, University of the Punjab, Lahore, Pakistan.

* **Corresponding author**: mscsf21m507@pucit.edu.pk

Article Citation:
Shoaib et al. (2023). "COVID-19 Lungs CT Scan Lesion Segmentation". *Foundation University Journal of Engineering and Applied Sciences*
DOI:10.33897/fujeas.v4i2.782

Abstract:

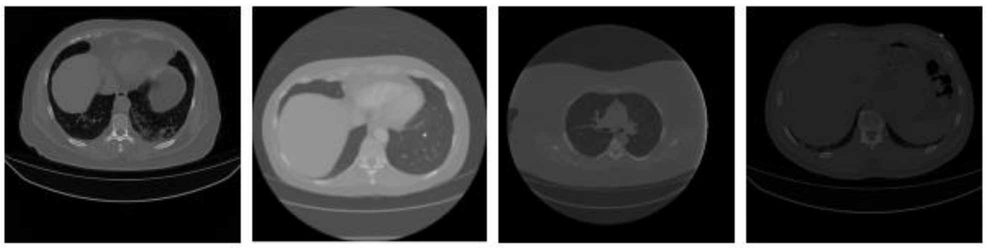
The outburst of the novel coronavirus 2019 has caused a multinational pandemic that has impacted a huge number of individuals around the globe. One of the primary indications of COVID-19 is the formation of lesions in the lungs, which can cause severe harm to the respiratory system and lead to death. In the following study, we submitted a novel strategy for making lung window CT scans and mediastinal window CT scans similar, to input it into a customized U-Net based model to achieve a decent degree of accuracy in segmenting these lung lesions. The method suggested in this research study is based on specialized image processing algorithms to normalize the CT scans' pixel intensity level and uniform the mediastinal and lung window CT scans. This allows us to accurately segment the lung lesions using a UNet model with a single channel input. We were able to achieve an IOU score of 82.4%, which is a significant addition to the existing Medical World. Additionally, the suggested approach is on par with cutting-edge methods.

Keywords: Pandemic; COVID-19; CT Scan; Image Processing; Lesion; Mediastinal Window; U-Net; Segmentation.

Published by
Foundation University
Islamabad.
Web: https://fui.edu.pk/

1. Introduction

Novel Corona virus 2019 a.k.a COVID - 19 is an illness of the respiratory system caused by a Novel Corona virus called SARS-CoV-2. In December 2019, a new coronavirus was identified in Wuhan, China. This coronavirus, now known as COVID-19, quickly spread throughout China and the rest of the world. As of March 2023, there have been over 400 million cases of COVID-19 and over 6 million deaths worldwide. The virus can be transmitted from person to person. In the early stages of infection with this virus, which closely resembled pneumonia, some patients developed severe symptoms of acute respiratory tract infection (ARI). Like many other lung diseases, it was observed that often COVID-19 patients' lungs had lesions formed inside them. A lesion is a localized area of tissue damage or disease. Pulmonary lesions can be benign (noncancerous) or malignant (cancerous). 3 cm or larger lesions inside the lungs are deemed to be severe malfunctioning masses and are usually ministered as malignant until confirmed otherwise. Early detection of lung lesions can lead to earlier treatment, improving patient outcomes and potentially saving lives. Till now according to WHO [1], Coronavirus cases have been found in 642,898,302 people out of which 6,625,569 deaths and 622,129,473 people recovered which shows the severity of this disease [Stats collected on 20/11/2022]. To analyze the lesions and detect them, a lung CT scan is done on suspected patients and medical professionals have to manually identify areas having lesions, which is very much time taking and unpleasant for them especially when there are numerous cases every day. In order to automate this process, data of many Covid patients having lesions in their lungs were prepared and publicly published, after which computer scientists took part in automating this and assisting medical professionals in this and assisting medical professionals in this pandemic. The dataset used in this research consists of CT scans of COVID patients containing two sorts of windows, that are: The Lung Window CT scan and the Mediastinal Window CT scan. Fundamentally to discern lungs from the whole Lungs CT scan images, a pristine strategy integrates technique based on histogram equalization and Optimum thresholding along Binary holes fill-based preprocessing in order to achieve similar data to be fed to our Neural network. So that it handles the Lung CT scan of different window types (Mediastinal and Lung Window) efficiently. The lung window and mediastinal window comprise different histograms because they represent distinct grayscale classes. However, we likewise noticed a minute transition in the histogram among different kinds of 2D CT scans of the mediastinal window, as shown in Figures 1 and 2. This suggests that there may be some variability in the grayscale distribution of mediastinal structures, which could be due to factors such as patient anatomy, scanner settings, or image acquisition techniques. The histograms of the Lung and Mediastinal windows in CT scans show that they have different local neighboring conditions around areas of interest. The slight difference between the histograms of the Mediastinal windows suggests that an efficient pre-processing method is needed to convert both windows into similar ranged 2D slices with two classes of grayscale. Additionally, an automated way to standardize both windows and extract the lungs from them for further processing is needed.



(a) Lung CT scan having Mediastinal Window (b) Lung CT scan having Lung Window (c) Lung CT scan having Lung Window (d) Lung CT scan having Mediastinal Window

Figure 1: Intensity variation in same window CT scans

The contribution of the following study is to present an austerer approach for the segmentation of lungs with lesions in CT scans of Covid patients using a five-layered encoder and decoder U-Net shown in Figure 3, based architecture. Specifically, this thesis focused on preprocessing lung CT scans with different window types (i.e., lung window and mediastinal window) and segmenting out the area of lungs

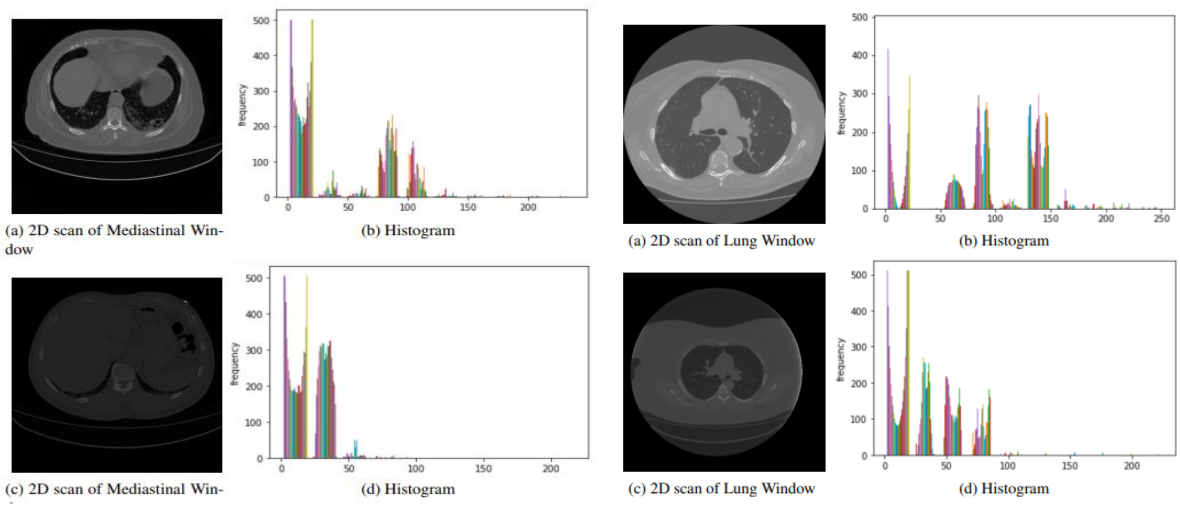


Figure 2: Histograms of different windows of lung CT scan

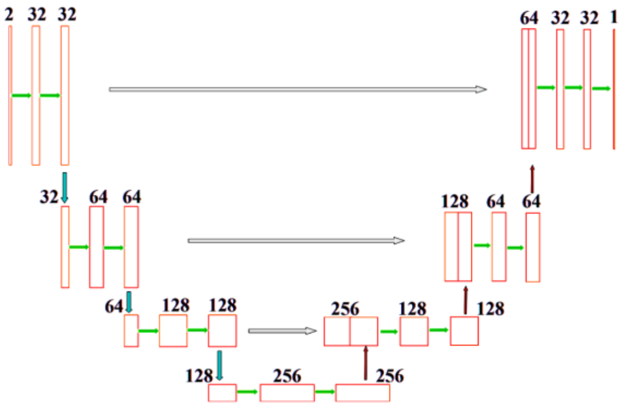


Figure 3: U-Net architecture

having lesions. To our knowledge, this approach is novel in the literature. It uses a simpler architecture than existing methods, requiring fewer layers and parameters, while still achieving comparable accuracy. In conclusion, this study's anticipated contribution is a novel method for the lungs with lesion segmentation in CT scan images employing more straightforward U-Net based architecture dealing with various window kinds of Lung CT scans.

1.1. Research Objectives

This research contributes a novel method leveraging specialized image processing algorithms to normalize pixel intensity levels of CT scans and align mediastinal and lung window CT scans. This approach enables accurate segmentation of lung lesions utilizing a UNet model with a single-channel input. With an achieved IOU score of 82.4%, this contribution represents a substantial advancement in medical imaging, offering promising implications for improved diagnosis and treatment of respiratory diseases.

1.1.1. Research Objective 1

To propose a novel strategy for aligning lung window CT scans and mediastinal window CT scans, aiming to enhance the accuracy of lesion segmentation in COVID-19 patients' lungs.

1.1.2. Research Objective 2

To develop and evaluate a customized U-Net based model capable of accurately segmenting lung lesions from CT scans normalized through the proposed alignment strategy, with the goal of achieving a high Intersection over Union (IOU) score as a significant contribution to the medical field.

1.2. Research Contributions

This study makes a significant contribution to improving healthcare outcomes by addressing a pressing problem in medical imaging. By introducing a novel method based on specialized image processing algorithms, the research effectively normalizes CT scans' pixel intensity levels and aligns mediastinal and lung window CT scans. This innovative approach enables accurate segmentation of lung lesions using a UNet model with a single-channel input. The achieved IOU score of 82.4% represents a noteworthy advancement in the field of medical imaging, providing clinicians with enhanced diagnostic tools for respiratory diseases. This empirical contribution not only adds to the existing body of knowledge but also offers practical solutions that can directly impact patient care, ultimately contributing to better health outcomes and quality of life.

2. Related Work

A number of publications give an overview of the subject of deep learning in medical image analysis, focusing on its practical applications, such as those by JG Lee et al. [4], Lakhani et al. [5], Suzuki [6] [7], Suhail et al. [54], Razak et al. [8], and Jiangg et al. [9] [10]. For instance, Razzak et al. [8] predict that by 2021, investment in medical image analysis will be equal to that in the entire analysis market in 2016, excluding only imaging. Utmost, exploration in medical vision and imaging is presently done with two-dimensional images and data, but some recent studies, identical to the study by Kayalibay et al. [11] and Roth et al. [12], have explored, the operation of three-dimension based filters(kernels) to use the entire details of three-dimensional medical visuals and images.

The strategy presented by Dimililer et al. in their paper [13] combines two image processing techniques, the Gaussian filter and the Discrete Wavelet Transform (DWT), to segment lung lesions in medical images. The authors reported an accuracy of 96% with the proposed method, which is a significant improvement over other methods. Adel Oulefki, et al. [14], mentioned the main hurdle in the segmentation of lesions is similar grayscale values to the values of soft tissues, Intensity in-homogeneity, and presence of the artifact. They proposed an enhancement technique and compared it with recent and advanced state-of-the-art segmentation techniques on the same data, achieving an Accuracy of 0.989 using the metric of masking and multi-level thresholding of images for image segmentation by simply optimizing K.E (KapurEntropy). Active contour models are a form of image segmentation technique that entails growing a contour to suit the limits of an object of interest in an image, as described by Younes Akbari et al. in [15].

Medical imaging is one of the areas in which deep learning has shown remarkable performance, especially in tasks such as segmentation, classification, and detection [16]. It provides various neural networks and attention models that are well-suited for processing medical images. These models comprise famous network backbones like AlexNet [17], GoogleNet [18], ResNet [19], and others. Gao et al. [20] proposed a network called a Dual-branch combination network a.k.a DCN, which exhibits the accurate detection and segmentation of lesions formed due to Novel Coronavirus-19 using CT scans. DCN was developed to achieve classification and lesion segmentation which resulted in high accuracy. CT scan images are preprocessed first using U-net and then DCN is applied which uses FC Net to get lesion segmentation. Dual-branch combination network reached a classification accuracy: of 96.74% on its internal data and 92.87% on its external validating data. M Lavanya et al. [21], present a method for lung lesion capturing in the Lung CT scans using a combination of the "Fuzzy Local Information Cluster Means" a.k.a FLICM algorithm and a backpropagation network for classification. Yu at el. [22], proposed a technique in which novel Dense-G.A.Ns was designed to get clear images and U-NET fused with a multilayer attention mechanism used to get accurate lesion segmentation. Further related work resulted in a Dense Generative Adversarial Network(D-GAN). Precision by this proposed method was as high as 0.94. Yao et al. [23], used a normalcy-recognizing network (NormNet) and came up with a label-free solution by making the model learn strong patterns of lesions. They synthesized lesions from which a normalcy-recognizing network was learnt which separates normal tissue from lesions. This work outperformed most unsupervised anomaly detection methods. Grady et al. [24], highlighted that due to the diversity in the types, shapes, and textures of lesions, it is really hard to come up with a solution working for every type of lesion in a CT. The Walker algorithm has been used to produce multiple 2D segmentation outputs and a resulting final 3D segmentation. The quantitative analysis stated it to be used for clinical purposes. Cai at el. [25], used a technique to segment lesions from anywhere in the whole body as universal lesion segmentation (ULS) for CT scans. Manually annotating such huge data requirements was never easy. So weakly supervised model was used here by a combination of a High-Resolution Network (HRNet) and Regional Level set (RLS) which resulted in the finest performance stats on the publicly available large-scale Deep-Lesion data set, concluded this work by RLS as it enhanced performance significantly. Many of these methods have shown high accuracy and robustness in the segmentation of COVID-19 lung lesions but still, most of the architectures used were based on either a single type of CT input i.e. Mediastinal Window or Lung Window. And also requires high computational units for training a large number of epochs.

3. The Research Problem

The advent of the COVID-19 pandemic has presented a pressing challenge to global healthcare systems, particularly in the realm of diagnosis and treatment. With over 400 million reported cases and more than 6 million deaths worldwide as of March 2023, the urgency to streamline diagnostic processes and improve patient outcomes has never been more apparent. COVID-19, caused by the novel coronavirus SARS-CoV-2, often manifests in severe respiratory symptoms akin to pneumonia, leading to the formation of pulmonary lesions in affected individuals. These lesions, indicative of tissue damage or disease, necessitate early detection and intervention for improved prognosis and patient survival.

Despite the advancements in medical imaging techniques, such as computed tomography (CT) scans, the manual identification of lesions within lung scans remains a time-consuming and labor-intensive task for healthcare professionals, particularly in the midst of a pandemic where the volume of cases overwhelms medical facilities. The need for automated tools to assist in the detection and segmentation of pulmonary lesions in COVID-19 patients is evident.

Existing approaches to automate lesion detection often rely on complex algorithms and deep learning architectures, which may require extensive computational resources and expertise to implement effectively. Moreover, these methods may not adequately address the variability in CT scan window types, such as lung window and mediastinal window, which exhibit distinct grayscale distributions and local neighboring conditions.

Thus, the primary research problem addressed in this study is the development of a novel and efficient method for the segmentation of lungs with lesions in CT scans of COVID-19 patients. Specifically, the challenge lies in preprocessing lung CT scans with varying window types and accurately delineating the regions of interest containing pulmonary lesions. This research aims to bridge the gap between the need for rapid, automated lesion detection and the practical constraints faced by healthcare professionals during the COVID-19 pandemic.

4. Proposed Solution

We have already discussed many lung lesion segmentation techniques developed that might have some pre-processing procedures consisting of standardizing images and extracting features. In these research works [13], [14] and [22] the authors marked lung lesion regions in two to three procedures, i.e., Image pre-processing, feature extraction or enhancement, and feeding to network. Some of them just handled segmentation with the raw images. Just like the foremost approach, our proposed method consists of three degrees i-e, Image Pre-processing for standardization, Encoding Image, and Decoding i.e., Generating Mask, with the use of a Neural Network. Our proposed segmentation approach is shown in Figure 4.

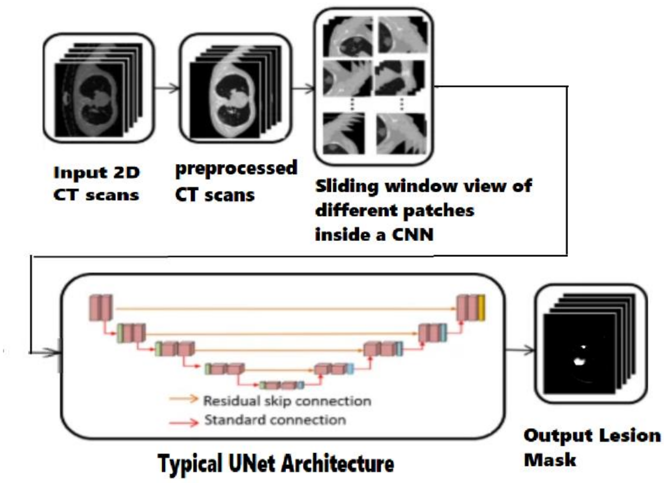
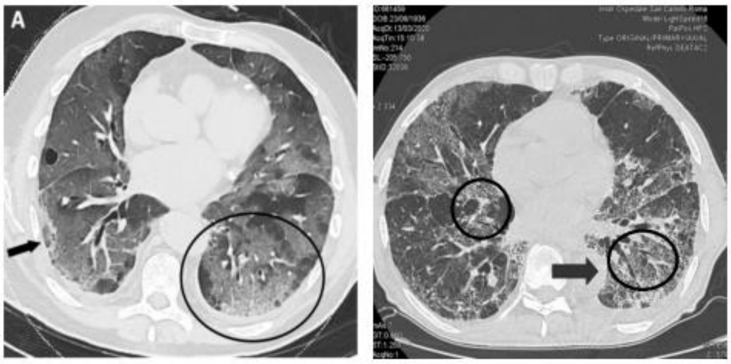


Figure 4: Overview of the proposed segmentation algorithm

4.1. Dataset and Data Selection

The dataset used for the problem stated in this study is a publicly available & published dataset [26], consisting of Lung CT scans of Covid patients, which has been collected from different published COVID datasets in a single place. The author claimed that they have built this immense lung CT-scan data set [27] for Novel CoronaVirus19 by curating data from seven publicly available data sets listed in the [28], [29], [30], [31], [32], [33] and [34]. As we only needed data for lesion segmentation from the data collection, we worked on only three of these datasets i.e. [31], [33], and [34], which had shared COVID-19 lesion masks. The dataset we used was also published separately at Kaggle [26], these three publicly available datasets, have 2729 CT scan images and ground truth binary mask pairs. All diverse types of lesions found in lungs in this dataset are mapped to white shade for consistency across datasets. Out of these 2729 images, 2171 were chosen on two bases, as follows Scans containing text in scans or some other annotation pointers like arrows or circular boundaries. Patient CT Scans having < 30 slices with high noise or CT having any anomaly/unwanted annotations. An example of rejected scans is below in Figure 5. As we proposed an approach to preprocessing images using conventional image processing, in order to make a multi-window image similar to the DL model, we selected 2171 images having no textual or irrelevant annotations on the input image itself. The selected data collection comprises labeled data from 68 patients (10, 9, and 49 images from 3 public datasets having masks). Data was further split into train, val, and test on the basis of unique patient CT scans with a distribution of 50, 5, and 13 patients respectively.



(a) Sample 1 (b) Sample 2

Figure 5: Sample images of rejected scans

4.2. Preprocessing of Images

4.2.1. Smoothing and Normalization

CT scan images can be noisy, especially if they are captured from old equipment. This noise can affect the results of subsequent analyses, so it is important to apply smoothing techniques to the images. Almost all segmentation techniques use some form of pre-processing before segmenting the images. This is to ensure that the results are consistent and accurate. In the following study, we used normalization and smoothing/blurring as preprocessing steps. Gaussian smoothing [35] is the most typically utilized technique for de-noising images, and it was found to be effective for our dataset. The kernel size of (3x3 and $\sigma = 0.5$) was used, which performed better than average smoothing. The samples in our data set contain a range of intensities of pixels from 0-255. To restrict the pixel intensity range to 0–1, we used image normalization [36]. This also helps a network to reach minima and converge faster.

4.2.2. Histogram Equalization

This is a typical procedure for enhancing the details of a low-featured or low-quality image. It works in a similar way to histogram-based stretching, but it often produces more aesthetically satisfactory outcomes with a variety of input visuals. H.E (Histogram Equalization) [37] involves making the resulting image's histogram as balanced as possible (where the across-the-board shape of the histogram stays the same). In essence, this causes a histogram with a mountain of tightly spaced pixels that appears to spread out or flatten, making lighted pixels appear lighter and unlit pixels appear darker. The H.E technique includes 4 principal steps:

a) Compute the running sum of the weights of the histogram.
b) Standardize the weights derived from the step above [step one] by just dividing by the total count of intensities.
c) Multiply the above-derived values from [step two] by the highest grey intensity and round off it.
d) Now, map the origin grey intensities to the outcomes from [step three], using 1-to-1 correspondence.

4.2.3. Thresholding via Multi-Otsu Method

Multi-Otsu thresholding [38] is a variant of Otsu's method, which is a popular thresholding technique for the segmentation of the desired images. Similar to Otsu's method, Multi Otsu thresholding also identifies the optimal values of the threshold by maximizing the variance between two classes of pixels (those above and below the threshold). However, unlike other thresholding methods like Adaptive, Global, or region-based thresholding techniques and also Otsu's method, which uses only one value as a threshold for the whole input image or regions, Multi-Otsu thresholding uses multiple values of threshold to segment the input image into more than 2-classes and allows for more complex ranges to be thresholded, due to which it was best fit for us to use it for this complex problem where we have variation of intensities due to different windowed Lung CT-scans. To implement Multi Otsu thresholding, first calculate the histogram of the image, which is a distribution of pixel intensities in the image, before implementing Multi Otsu thresholding. The best threshold values are then determined by computing the histogram's cumulative sum. For each threshold value, the algorithm separates the image into two groups successively. It then analyses the deviation between the groups and chooses the threshold value that minimizes the variance. Up till the necessary number of courses is reached, the process is repeated. As shown in Eq. 1.

$$\max_k \sum_{i=1}^{k} p_i\,(\mu_k - \mu_w)^2 + \sum_{i=k+1}^{L} p_i\,(\mu_L - \mu_k)^2 \tag{1}$$

where: L is the number of intensity levels in the image

- *k* is the number of thresholds to be found
- *pi* is the normalized histogram of the image at intensity level i
- *µk* is the mean intensity value of the pixels in the kth class
- *µw* is the global mean intensity value of the image

4.2.4. Binary Hole Filling

The binary fill holes algorithm fills holes in a binary image by starting from the boundary pixels and propagating inwards, filling in any white pixels (with a value of 1) that are surrounded by black pixels (with a value of 0). The binary fill holes operation is used to fill in the gaps of a binary image. It accepts a binary input array with gaps that need to be filled. The code then extracts the logical not from the input array to produce a binary mask. It produces a transient binary array with the same dimensions as the mask. If an output array is specified as an argument, the function uses the structuring element to execute a binary dilation operation on the temporary array and inserts the results into the specified array. The output array's logical not is then used. If not, it carries out the same dilation procedure and then, after taking into account its logical not, returns the array. On the complement of the input image, beginning at the outer edge of the image, the dilation operation is carried out. The holes are not filled in during the dilation procedure since they are not connected to the boundary. The filled-in version of the input binary array is the outcome after taking its complement.

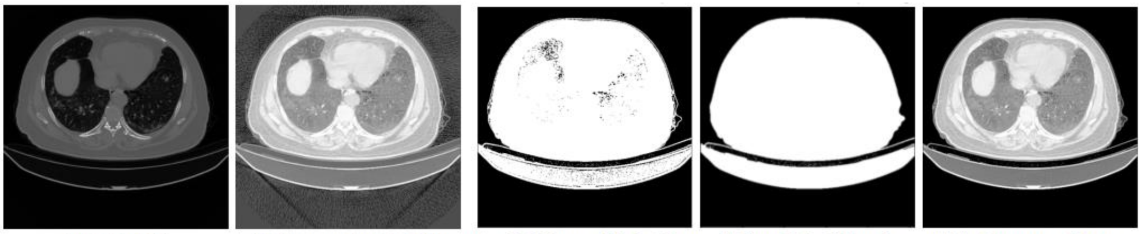
```
function BINARY FILL HOLES(input,structure = None,output = None,origin = 0)
  mask ← logical not(input)
  temp ← zeros(mask.shape,bool)
  inplace ← is instance?(output,ndarray)
  if inplace then
        binary dilation(temp,structure,−1,mask,out put,1,origin)
        logical not(output,output)
  else
        output ← binary dilation(temp,structure,−1,mask,None,1,origin)
        logical not(output,output)
        return output
  end if
end function
```

4.3. UNet

Convolutional neural networks [39], and CNNs significantly enhanced image identification and detection tasks, particularly between 2011 and 2015 when the top-5 mistake rate in the imagenet challenge dropped from over 25% to about 4% [40] [41]. The best classification accuracy is achieved by new designs like ResNet [19] and Inception V3 [42]. UNet is a widespread deep-learning architecture that is commonly employed for image segmentation tasks in various domains including medical imaging, remote sensing, and computer vision. The architecture was introduced by Ronneberger et al. in 2015, and since then, it has evolved into one of the most widespread models for visual segmentation-relevant tasks [43]. Conventional UNet architecture is shown above in Figure 3. UNet is a highly customizable architecture, and its design can be adjusted to suit various applications.

4.3.1. Proposed UNet Architecture

So for the problem stated, we developed a customized U-net architecture having a lower number of learnable layers as compared to heavy U-net architectures, which leads to less training time and epoch and also less computational cost. Figure 6 shows the top-view modeling of the presented U-net architecture in 2D. And in Figure 7, shows the 3D modelling of the presented U-net architecture. The



(a) Original CT (b) Histogram Equalized CT (c) Multi-Otsu Global Mean Thresholded mask (d) Binary Holes filled mask (e) Final Preprocessed Output

Figure 6: Outcome after pre-processing applied

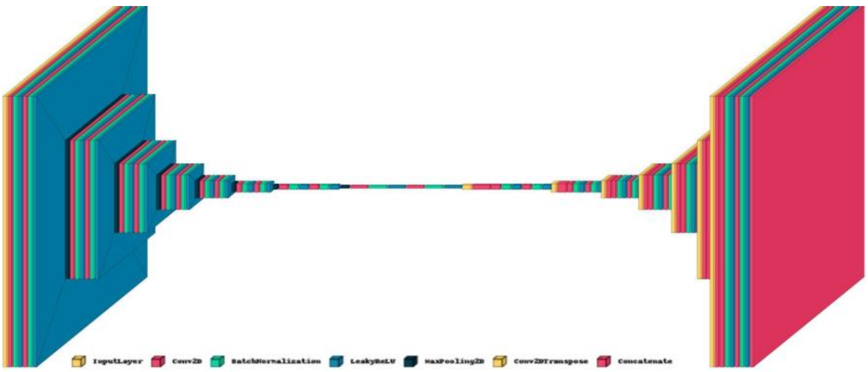


Figure 7: 3D view of proposed U-net architecture

model architecture starts with a series of encoder blocks, each consisting of 2 CL (convolutional layers) pursued by BN(batch-normalization) and leaky ReLU function as activation. Each encoder block is then followed by max-pooling to downsize the spatial dimensions of the feature maps. The outcome of each encoding block is then handed on to the decoding grid, where it is upsampled and concatenated with the corresponding encoder block's output. Each decoder block comprises a spatial increasing or up-sampling layer followed by a concatenation layer and two convolutional layers. The final output layer uses the sigmoid activation function.

4.3.2. Encoder Details

Here in the encoder part of the architecture, the height & width of feature maps decreases due to operations like convolution and pooling, which downsample the input. This downsampling helps extract hierarchical features efficiently and sequentially block by block by capturing coarse-details at lower resolutions in encoded form, and form highly informative features enriched flatten embedding just before the decoder part. Summary of an encoder block of the proposed UNet architecture. The input is a grayscale image, of size 512x512 with only one channel. The block starts with a convolutional layer with 8 filters, pursued by BN (batch-normalization) and the LeakyReLU activation function. Then, another convolutional layer with 8 filters, batch normalization, and LeakyReLU activation is added. A max-pooling layer with a 2x2 kernel and stride of two is used to downsample the feature maps. Next, two convolutional layers with 16 filters each are added with batch normalization and LeakyReLU activation. Another max-pooling layer is applied to decrease spatial information or down-sample the obtained feature maps. Then, two convolutional layers with 32 filters each are added with batch normalization and LeakyReLU activation. Another max-pooling layer is applied to downsample the feature maps. Finally, a convolutional layer with 64 filters, batch normalization, and LeakyReLU activation is added to obtain the output feature maps. A total of 7 encoder blocks are there having Height and Width decreasing by a factor of 1/2, resulting initial height of 512 decreasing down to 8 before the bridge.

4.3.3. Decoder Details

The decoder block of the U-Net architecture takes in the feature maps from the corresponding encoder block and then upsamples them. It concatenates the upsampled feature maps with the feature maps from the corresponding encoder block to form the input for the next layer of the decoder block. The first layer in the decoder block is a transposed convolutional layer that upsamples the feature maps by a factor of 2. The resultant of this layer is connected with the feature maps from the respectively interconnected encoder block, and the resulting tensor is provided into a sequence of convolutional layers with a smaller count of kernels compared to the corresponding encoder block. Each convolutional layer is pursued by a batch normalization layer and a leaky ReLU activation function. The decoder block consists of three stages. Each stage comprises a transposed convolutional layer, followed by concatenation, two convolutional layers, batch normalization, and leaky ReLU activation functions. The number of filters in each stage is halved, resulting in feature maps with a reduced spatial resolution. The output of the decoder block is the final segmentation mask. The number of decoder blocks is also equal to encoders i.e. 7.

Table 1: Stats of params of architecture proposed

Total params	Learn-able params	Non-Learn-able params
31,138,089	31,125,865	12,224

A Sigmoid-Loss function is among the few most popular losses or loss functions that have been used with UNet architecture. It is based on a binary cross-entropy function of loss that maps the predicted values into the range of 0 to 1 using the sigmoid function. See Eq. 2.

$$L_{sigmoid} = -\frac{1}{N}\sum_{i=1}^{N}(y_i \log(p_i) + (1 - y_i)\log(1 - p_i)) \qquad (2)$$

5. Experiment and Results

5.1. Training and Validation Stats

The network was trained for 50 epochs using Adam optimizer [44] with the default values of 0.001 for the learning rate and 0.9 for the momentum. and early stopping was applied, here are the stats we achieved on training: Training IoU: 84.17% and Validation IoU: 82.89%, as shown in Figure 8.

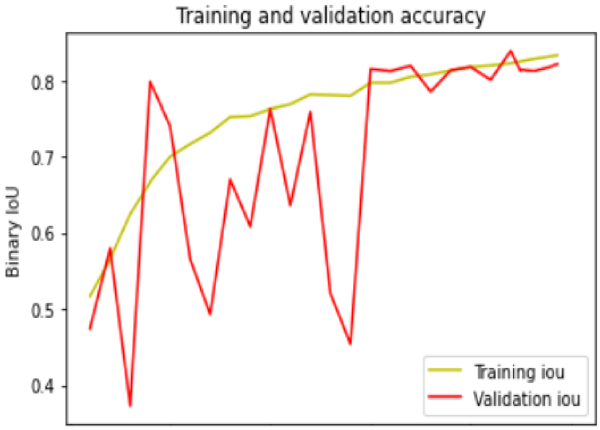
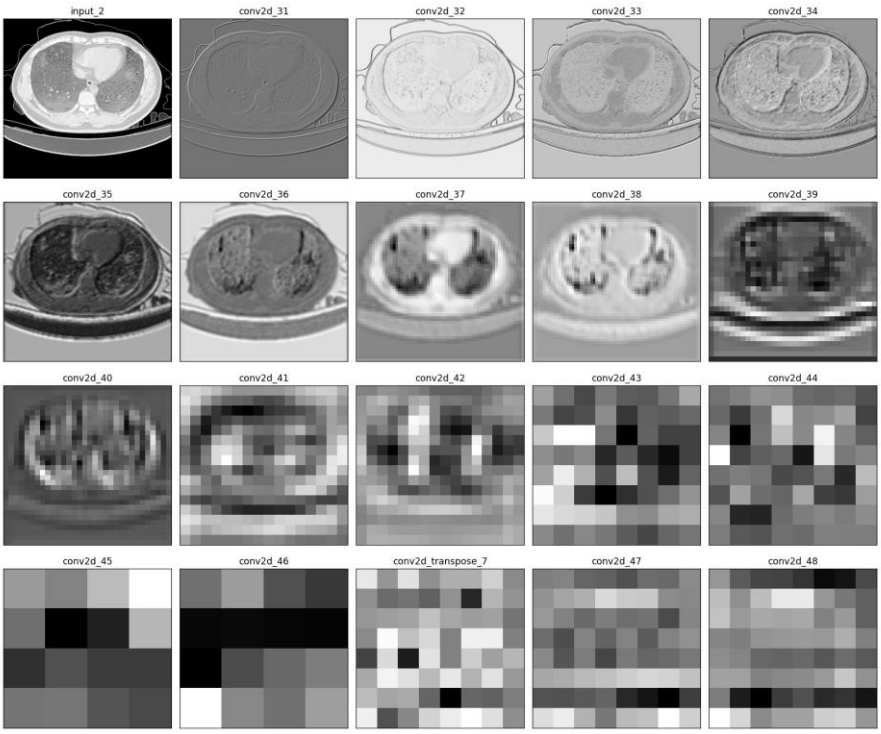


Figure 8: Train validation IoU

5.2. Extracted Features Visualization

We presented the stats obtained from training our custom U-Net network and here's an example showcasing the visualization of the features extracted. These features are derived from a testing data sample, shedding light on the inner workings of our U-Net architecture in Figure 9.



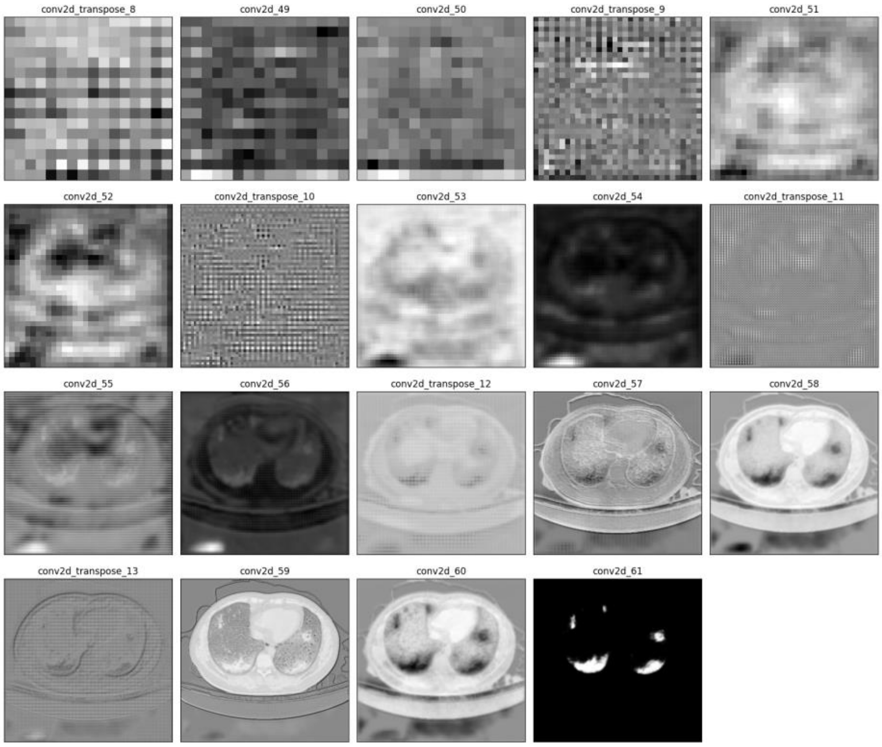


Figure 9: Insights from UNet Conv2D layer outputs: Exploring intermediate representations

Along with Encoder, Decoder side also clearly showed how features were being highlighted from the network by assigning positive and negative weights, which are more evident if we look at the very last layers. In the final layer, the Network eradicated unimportant areas and just outcomes with the lesion-highlighted area.

5.3. Results

We used 388 unique CT images for testing which comprised of different windows and also different orientations in order to generalize the testing dataset. The proposed algorithm enabled us to achieve IoU scores of up to 82.4% and Dice scores of up to 80.48%. A few examples of the outcomes of the current presented solution are presented below in Figure 10, where;

a) Original CT Scan (Mediastinal Window and Lung Window)
b) Preprocessed Image for standardizing
c) Lung Lesion Ground Truth Mask
d) Predicted Lung Lesion Mask of our Proposed Technique

5.4. Comparative Analysis with some Existing Techniques

Here we explored a few already existing Covid-19 Lung Lesion Segmentation strategies and then analyzed their performance with our suggested segmentation strategy. As we worked on segmentation and to assess the performance of our presented algorithm we chose 3 main metrics (IoU, DSC, and Accuracy) which are already described above. Secondly, our main focus was to design a method that works decently on multi-window lung CT scans. We observed results on the unseen combined dataset with the Mediastinal window and the Lung Window data. However, it is challenging to evaluate our method with other Novel CoronaVirus-19 lesion segmentation techniques due to the unavailability and unacquainted details of most datasets, differences in the splitting of datasets, and annotation quality. So, we compared our stats with other different existing algorithms and segmentation techniques with their claimed results in lung lesion segmentation experimentation on the same metric, mainly dice score.

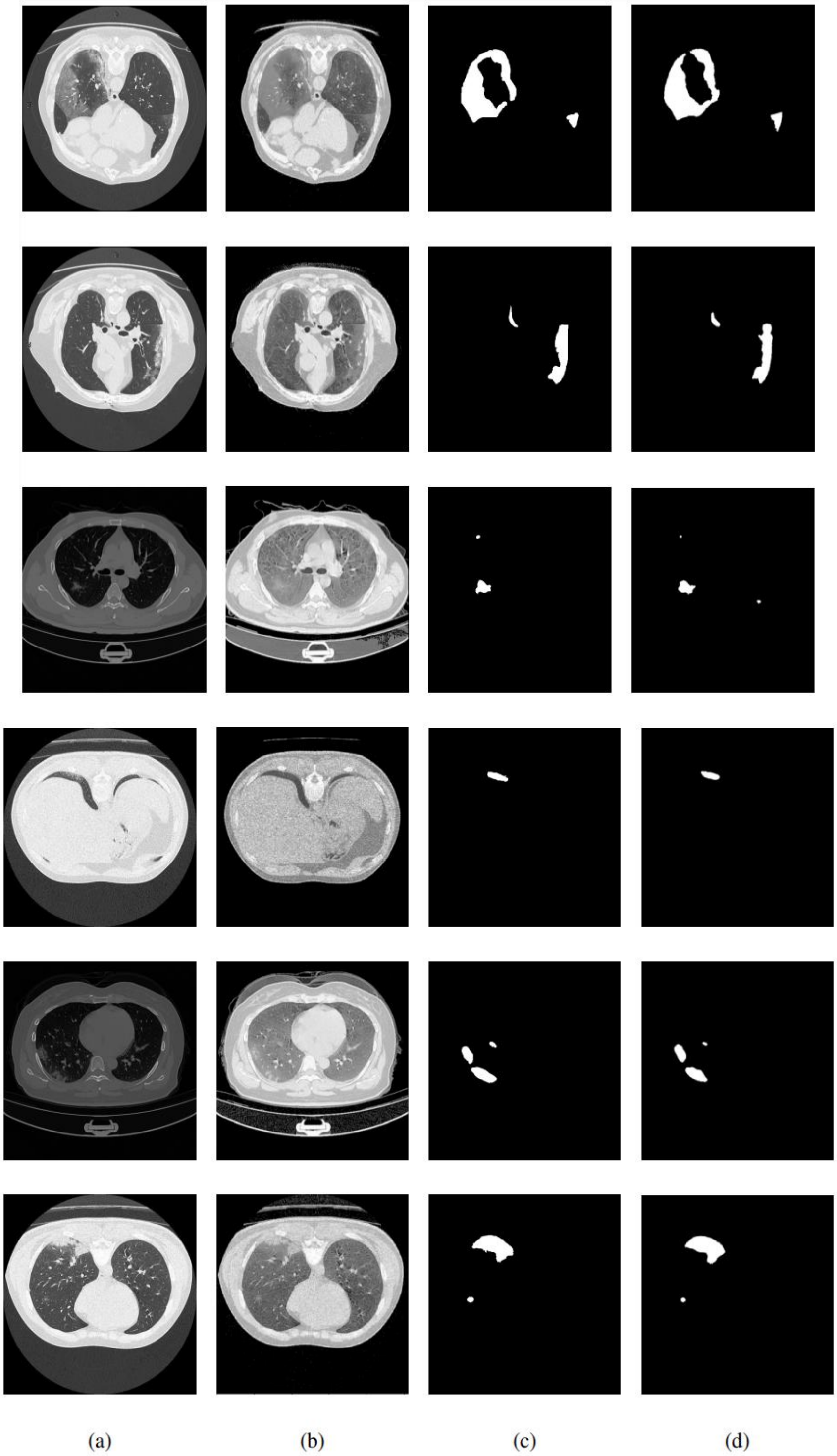


Figure 10: Lung CT scans along with results of several covid-19 patients (shown in sequence)

This is shown in Table 2.

6. Conclusions and Future Direction

In conclusion, the present study addressed the challenging task of Covid-19 lung CT scan lesion segmentation by introducing a novel image processing technique and a custom-designed U-Net architecture. Our proposed approach standardizes multi-window lung CT scans and keeps the simple U-Net architecture is robust for high-quality 512 × 512 CT images while providing competitive results.

Table 2: Generic lung lesion segmentation performance comparison

Methodology	Window Type	Dice	IoU	Accuracy
U-Net(VGG-16 backbone) without preprocessing [43]	Multi window	0.439	-	-
U-Net++ (VGG-16 backbone) [45]	Multi window	0.581	-	-
Dense U-Net [46]	Multi window	0.515	-	-
Gated U-Net [47]	Multi window	0.623	-	-
Inf-Net [48]	Multi window	0.682	-	-
COPLE-Net [49]	Multi window	0.682	-	-
Dense GAN and multi-layer attention [22]	Multi window	0.683	-	-
LABEL-FREE SEGMENTATION [23]	Multi window	0.698	-	-
Seg-Net [50]	Multi window	0.705	-	-
BiSe-Net [51]	Multi window	0.706	-	-
ESP-Net [52]	Multi window	0.706	-	-
Automatic COVID-19 LIR segmentation [14]	Multi window	0.714	-	0.98
nnU-Net [53]	Multi window	0.781	-	-
Proposed Method	Multi window	0.8048	0.8248	0.99

The utilization of the proposed algorithm enabled us to achieve IoU scores of up to 82.4% and Dice scores of up to 80.48%, utilizing a mere 31M total parameters. Our research contributes to the field of medical image analysis by providing an efficient and accurate approach for Covid-19 lung CT scan lesion segmentation. These promising results suggest that our method has great potential to be applied to other medical image analysis tasks and can perform as a standard for forthcoming investigations in this domain.

7. References

[1] "World Health Organization - death count toll," https://covid19.who.int/.

[2] "American Council on Science and Health - Coronavirus: COVID Deaths In U.S. By Age, Race," https://www.acsh.org/news/2020/06/23/coronavirus-covid-deaths-us-age-race-14863

[3] Z. T. Al-sharify, T. A. Alsharify and N. Al-sharify, "A critical review on medical imaging techniques (CT and PET scans) in the medical field, " In *Proceedings of IOP Conference Series Materials Science and Engineering*, IOP Publishing, 2020, vol. 870(1).

[4] J. G. Lee, S. Jun, Y. W. Cho, H. Lee, G. B. Kim, J. B. Seo and N. Kim, "Deep learning in medical imaging: general overview, " *Korean Journal of Radiology*, vol. 18(4), p.570, 2017.

[5] P. Lakhani, D. L. Gray, C. R. Pett, P. Nagy and G. Shih, "Hello world deep learning in medical imaging," *Journal of Digital Imaging*, vol. 31, pp. 283–289, 2018.

[6] K. Suzuki, "Overview of deep learning in medical imaging," *Radiological Physics and Technology*, vol. 10(3), pp. 257–273, 2017.

[7] K. Suzuki, "Survey of deep learning applications to medical image analysis," *Medical Imaging Technology*, vol. 35(4), pp. 212–226, 2017.

[8] M. I. Razzak, S. Naz and A. Zaib, "Deep learning for medical image processing: Overview, challenges and the future," *Classification in BioApps: Automation of Decision Making*, pp. 323–350, 2018.

[9] F. Jiang, A. Grigorev, S. Rho, Z. Tian, Y. Fu, W. Jifara, K. Adil and S. Liu, "Medical image semantic segmentation based on deep learning," *Neural Computing and Applications*, vol. 29, pp. 1257–1265, 2018.

[10] J. Jiang, P. Trundle and J. Ren, "Medical image analysis with artificial neural networks," *Computerized Medical Imaging and Graphics*, vol. 34(8), pp. 617–631, 2010.

[11] B. Kayalibay, G. Jensen and P. van der Smagt, "CNN-based segmentation of medical imaging data," *arXiv preprint arXiv:1701.03056*, 2017.

[12] H. R. Roth, C. Shen, H. Oda, M. Oda, Y. Hayashi, K. Misawa and K. Mori, "Deep learning and its application to medical image segmentation," *Medical Imaging Technology*, vol. 36(2), pp. 63–71, 2018.

[13] K. Dimililer, A. Hesri and Y. Kirsal Ever, "Lung lesion segmentation using gaussian filter and discrete wavelet transform," In *ITM Web of Conferences*, 2017, vol. 11, p.01018.

[14] A. Oulefki, S. Agaian and T. Trongtirakul, "Automatic COVID-19 lung infected region segmentation and measurement using CT-scans images," *Pattern Recognition*, vol. 114, p.107747, 2020.

[15] Y. Akbari, H. Hassen, S. Al-Maadeed and S. M. Zughaier, "COVID-19 lesion segmentation using lung CT scan images: Comparative study based on active contour models," *Applied Sciences*, vol. 11(17), p.8039, 2021.

[16] J. Ker, L. Wang, J. Rao and T. Lim, "Deep learning applications in medical image analysis," *IEEE Access*, vol. 6, pp. 9375–9389, 2017.

[17] A. Krizhevsky, I. Sutskever and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Communications of the ACM*, vol. 60(6), pp. 84–90, 2017.

[18] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke and A. Rabinovich, "Going deeper with convolutions," In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, IEEE, 2015, pp. 1–9.

[19] K. He, X. Zhang, S. Ren and J. Sun, "Deep residual learning for image recognition," In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, IEEE, 2016, pp. 770–778.

[20] K. Gao, J. Su, Z. Jiang, L. L. Zeng, Z. Feng, H. Shen, P. Rong, X. Xu, J. Qin, Y. Yang, W. Wang and D. Hu, "Dual-branch combination network (DCN): Towards accurate diagnosis and lesion segmentation of COVID-19 using CT images," *Medical Image Analysis*, vol. 67, p.101836, 2021.

[21] M. Lavanya and P. Kannan, "Lung lesion detection in CT scan images using the fuzzy local information cluster means (FLICM) automatic segmentation algorithm and back propagation network classification," *Asian Pacific Journal of Cancer Prevention: APJCP*, vol. 18(12), p.3395, 2017.

[22] J. Zhang, L. Yu, D. Chen, W. Pan, C. Shi, Y. Niu, X. Yao, X. Xu and Y. Cheng, "Dense GAN and multi-layer attention based lesion segmentation method for COVID-19 CT images," *Biomedical Signal Processing and Control*, vol. 69, p.102901, 2021.

[23] Q. Yao, L. Xiao, P. Liu and S. K. Zhou, "Label-free segmentation of covid-19 lesions in lung CT, " *IEEE Transactions on Medical Imaging*, vol. 40(10), pp. 2808–2819, 2021.

[24] M. P. Jolly and L. Grady, "3D general lesion segmentation in CT," In *5th IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, IEEE, 2008, pp. 796–799.

[25] Y. Tang, J. Cai, K. Yan, L. Huang, G. Xie, J. Xiao, J. Lu, G. Lin and L. Lu, "Weakly- supervised universal lesion segmentation with regional level set loss, " In *Proceedings of 24th International Conference on Medical Image Computing and Computer Assisted Intervention–MICCAI 2021*, Springer International Publishing, 2021, Part II 24, pp. 515–525.

[26] "Large COVID-19 CT scan slice dataset — kaggle," https://www.kaggle.com/datasets/maedemaftouni/large-covid19-ct-slice-dataset

[27] M. Maftouni, A. C. C. Law, B. Shen, Y. Zhou, N. Ayoobi Yazdi and Z. Kong, "A robust ensemble-deep learning model for COVID-19 diagnosis based on an integrated CT scan images database," In *Proceedings of IISE Annual Conference*, 2021, pp. 632–637.

[28] J. Zhao, Y. Zhang, X. He and P. Xie, "COVID-CT-Dataset: a CT scan dataset about COVID-19," arXiv preprint *arXiv:2003.13865*, 2020.

[29] P. Afshar, S. Heidarian, N. Enshaei, F. Naderkhani, M. J. Rafiee, A. Oikonomou, F. B. Fard, K. Samimi, K. N. Plataniotis and A. Mohammadi, "COVID-CT-MD, COVID-19 computed tomography scan dataset applicable in machine learning and deep learning," *Scientific Data*, vol. 8(1), p.121, pp. 1–8, 2021.

[30] J. P. Cohen, P. Morrison, L. Dao, K. Roth, T. Q. Duong and M. Ghassemi, "Covid-19 image data collection: Prospective predictions are the future, " *arXiv preprint arXiv:2006.11988*, 2020.

[31] S. P. Morozov, A. E. Andreychenko, N. A. Pavlov, A. V. Vladzymyrskyy, N. V. Ledikhova, V. A. Gombolevskiy, I.A. Blokhin, P. B. Gelezhe, A. V. Gonchar and V. Y. Chernina, "MosMedData: Chest CT Scans With COVID-19 Related Findings Dataset," *arXiv preprint arXiv:2005.06465*, 2020.

[32] M. Rahimzadeh, A. Attar and S. M. Sakhaei, "A fully automated deep learning-based network for detecting COVID-19 from a new and large lung CT scan dataset," *Biomedical Signal Processing and Control*, vol. 68, p.102588, 2020.

[33] M. Jun, G. Cheng, W. Yixin, A. Xingle, G. Jiantao, Y. Ziqi, Z. Minqing, L. Xin, D. Xueyuan, C. Shucheng and W. Hao, "COVID-19 CT Lung and Infection Segmentation Dataset," In: Zenodo, 2020, https://doi.org/10.5281/zenodo.3757476.

[34] "COVID-19" 2020. [Online] http://medicalsegmentation.com/covid19/ [Online: Accessed 23 December 2020].

[35] Wikipedia, "Gaussian blur — Wikipedia, the free encyclopedia," http://en.wikipedia.org/w/index.php?title=Gaussian%20bluroldid=1154879725, 2023, [Online: Accessed 19 May 2023].

[36] Wikipedia, "Normalization (image processing) — Wikipedia, the free encyclopedia," http://en.wikipedia.org/w/index.php?title=Normalization%20(image%20processing)oldid=1080487172, 2023, [Online: Accessed 19 May 2023].

[37] W. Zhihong and X. Xiaohong, "Study on histogram equalization," In *2011 2nd International Symposium on Intelligence Information Processing and Trusted Computing*, IEEE, 2011. 2011, pp. 177–179, doi: 10.1109/IPTC.2011.52.

[38] C. Wang, J. Yang and H. Lv, "Otsu multi-threshold image segmentation algorithm based on improved particle swarm optimization," In *2019 IEEE 2nd International Conference on Information Communication and Signal Processing (ICICSP)*, IEEE, 2019, pp. 440–443.

[39] Y. LeCun and Y. Bengio, "Convolutional networks for images, speech, and time series," *The Handbook of Brain Theory and Neural Networks*, vol. 3361(10), 1995.

[40] P. Agrawal, R. Girshick and J. Malik, "Analyzing the performance of multilayer neural networks for object recognition," In *Proceedings of 13th European Conference on Computer Vision–ECCV 2014,* Springer International Publishing, 2014, Part VII 13, pp. 329–344.

[41] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein and A. C. Berg, "Imagenet large scale visual recognition challenge," *International Journal of Computer Vision*, vol. 115, 2015.

[42] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens and Z. Wojna, "Rethinking the inception architecture for computer vision," In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition,* IEEE, 2015, pp. 2818–2826.

[43] O. Ronneberger, P. Fischer and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," In *Proceedings of 18th International Conference on Medical Image Computing and Computer-assisted Intervention–MICCAI*, Springer International Publishing, 2015, Part III 18, pp. 234–241.

[44] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.

[45] Z. Zhou, M. M. Rahman Siddiquee, N. Tajbakhsh and J. Liang, "Unet++: A nested u-net architecture for medical image segmentation," In *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: 4th International Workshop, DLMIA 2018, and 8th International Workshop, ML-CDS 2018, Held in Conjunction with MICCAI 2018,* Springer International Publishing, 2018, pp. 3–11.

[46] X. Li, H. Chen, X. Qi, Q. Dou, C. W. Fu and P. A. Heng, "H-DenseUNet: hybrid densely connected UNet for liver and tumor segmentation from CT volumes," *IEEE Transactions on Medical Imaging*, vol. 37(12), pp. 2663–2674, 2018.

[47] J. Schlemper, O. Oktay, M. Schaap, M. Heinrich, B. Kainz, B. Glocker and D. Rueckert, "Attention gated networks: Learning to leverage salient regions in medical images," *Medical Image Analysis*, vol. 53, pp. 197–207, 2019.

[48] L. Zhou, Z. Li, J. Zhou, H. Li, Y. Chen, Y. Huang, D. Xie, L. Zhao, M. Fan, S. Hashmi and F. Abdelkareem, "A rapid, accurate and machine-agnostic segmentation and quantification method for CT-based COVID-19 diagnosis," *IEEE Transactions on Medical Imaging*, vol. 39(8), pp. 2638–2652, 2020.

[49] G. Wang, X. Liu, C. Li, Z. Xu, J. Ruan, H. Zhu, T. Meng, K. Li, N. Huang and S. Zhang, "A noise-robust framework for automatic segmentation of COVID-19 pneumonia lesions from CT images," *IEEE Transactions on Medical Imaging*, vol. 39(8), pp. 2653–2663, 2020.

[50] V. Badrinarayanan, A. Kendall and R. Cipolla, "Segnet: A deep convolutional encoder-decoder architecture for image segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39(12), pp. 2481–2495, 2017.

[51] X. Wang, K. Yu, S. Wu, J. Gu, Y. Liu, C. Dong, Y. Qiao and C. C. Loy, "Esrgan: Enhanced super-resolution generative adversarial networks," In *Proceedings of the European Conference on Computer Vision (ECCV) Workshops*, 2018.

[52] S. Woo, J. Park, J. Lee and I. S. Kweon, "CBAM: Convolutional block attention module," In *Proceedings of the European Conference on Computer Vision (ECCV)*, Springer, 2018, pp. 3–19.

[53] F. Isensee, P. F. Jaeger, S. A. Kohl, J. Petersen and K. H. Maier-Hein, "nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation," *Nature methods*, vol. 18(2), pp. 203–211, 2021.

[54] M. Nasir, M. S. Farid, Z. Suhail and M. H. Khan, "Optimal Thresholding for Multi-Window Computed Tomography (CT) to Predict Lung Cancer," *Applied Sciences*, vol. 13(12), p.7256, 2023.