



Academic Year : 2023/2024

**3D Generative Adversarial Network and TransUNet-based approach for  
Self-supervised region-aware segmentation of COVID-19 CT scans**

**Supervised by:** BANOUAR Oumayma

**Realized by:** DHIMEN Aymane, EL  
ANSARI Mostapha, EL FAHSI Aymane,  
KHALI M.Adam, KOUDIA Selma

# Outline

- Introduction
- Contextualization
- Problematic
- Initial approach
- Proposed approach for self supervised segmentation of Covid 19 CT scans
- Methods
- Results and discussion
- Conclusion and Future Work
- References

## Introduction

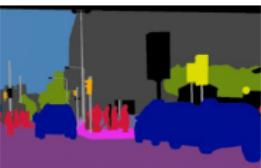
Image segmentation involves dividing an image into several regions or segments, often pixel by pixel, to identify and isolate different structures or objects present.

**Image segmentation can :**

- ▶ Ensure accuracy
- ▶ Reduce resource-intensive and time-consuming
- ▶ Quantify the detection



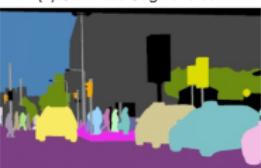
(a) Image



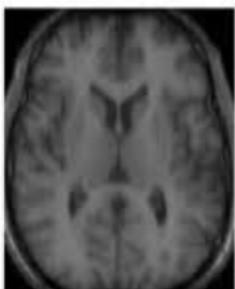
(b) Semantic Segmentation



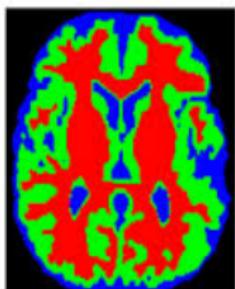
(c) Instance Segmentation



(d) Panoptic Segmentation



(a) coupe axial



(b) segmentation des tissus

Figure 1:Image segmentation

## Contextualization

**Medical image segmentation plays a crucial role in analysis and diagnosis**

- ▶ Locate and precisely delineate organs, tissues, or lesions.
- ▶ Quantification of features: size, shape, or volume

**The automatic medical image segmentation models include:**

- ▶ Fully supervised
- ▶ Semi-Supervised
- ▶ Weakly-Supervised
- ▶ Self-Supervised

**We chose to use self-supervised segmentation because it addresses challenges:**

- ▶ Not enough healthy CT scans ,
- ▶ Not Costly : No need to wait for the expert's feedback mainly in labelling data,

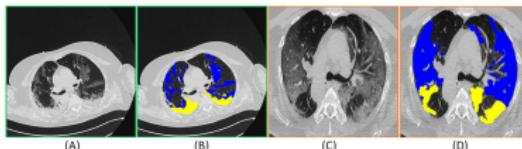


Figure 2:Example-segmentation of Covid-19 infection areas in the lungs

## Problematic

How can we develop a segmentation model capable of:

- ▶ Predicting lesions in COVID-19 CT scans without manual annotation at the pixel level.
- ▶ Achieving results that surpass the performance of existing state-of-the-art unsupervised and weakly-supervised segmentation techniques.

# Initial approach

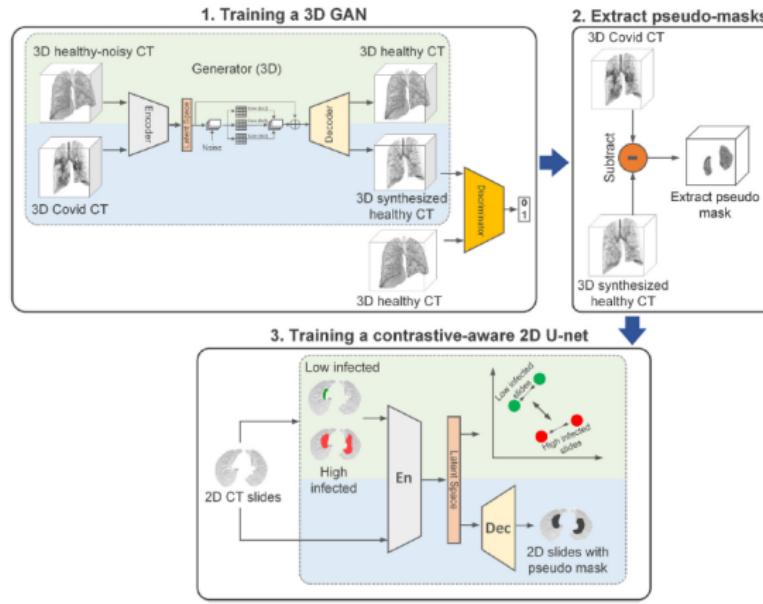
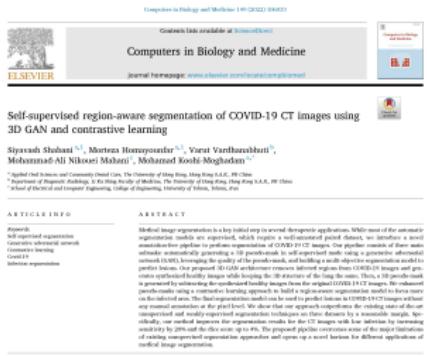


Figure 3 : Original Pipeline



## 1. Introduction

Medical image segmentation is the process of dividing an image into several sectors to identify diagnostic and clinical characteristics of the images. This is a crucial initial step in several therapeutic applications, such as cancer detection [1–3]. Over the past few years, many effective approaches for medical image segmentation have been reported [4–10]. These approaches can be categorized into two main groups: supervised and unsupervised (UNS) methods. These methods train the model on paired data sets consisting of the original medical images and their corresponding ground truth segmentation maps. They often require a large amount of labeled data to train the model, which is often unaffordable for experts [1–3]. However, one significant limitation of these approaches is the requirement of using a well-annotated dataset. This is not always possible, particularly in the medical image domain,

where creating medical image datasets is expensive and time-consuming due to the sensitive nature of the data and the need for highly specialized medical knowledge. Therefore, the demand for a well-annotated dataset for unsupervised medical image segmentation is high. On the other hand, the necessity for a (well-annotated) dataset is a great bottleneck for the success of the proposed methods [11–13]. Recently, some researchers have proposed a novel pipeline to segment COVID-19 CT images without any annotation of the images [14–16]. The basic idea behind this pipeline is to generate COVID-19 corrupted tomography (CT) image without any annotation of the images. This image is then used as an input to a 3D Generative Adversarial Network (GAN) to generate healthy images. The healthy images produced by GAN would then be subtracted from the original CT images to extract the ROI structure of the lung regions. Finally, a 2D U-Net is trained to segment the healthy images while keeping the ROI structure of the lung regions without the need for human annotation. We train a multi-objective segmentation network to both pseudo mask and contrastive losses. Estimated below is a summary of the main contributions of our work:

\* Corresponding author.  
† Mortes Hosseynzadeh (M. Hosseynzadeh).

✉ These authors contributed equally.

<https://doi.org/10.1016/j.combiomed.2022.105683>  
Received 27 August 2021; Received 27 November 2021; Accepted 29 August 2022  
Available online 27 September 2022  
© 2022 Elsevier Ltd. All rights reserved.

Figure 4: First page of original research article

## Proposed approach for self supervised segmentation of Covid 19 CT scans

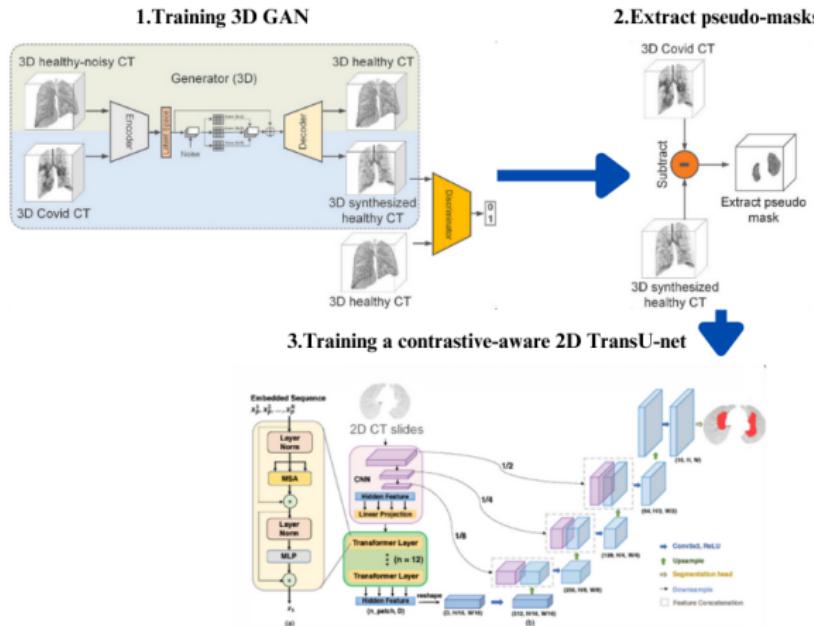


Figure 5 : Our Pipeline: GAN (with Perlin Noise) + contrastive-aware TransUNET Model

## Proposed approach for self supervised segmentation of Covid 19 CT scans

- ▶ Phase 1: 3D GAN for Synthesizing Pseudo-Masks
- ▶ Phase 2: Synthesizing Pseudo Masks
- ▶ Phase 3: Segmenting the CT scans using TransUNET

## Methods

### Phase 1: 3D GAN for Synthesizing Pseudo-Masks

In our study, we used a 3D Generative Adversarial Network (GAN) to transform CT scans of lungs affected by COVID-19 into their healthy counterparts.

- ▶ Train a multi-objective 3D GAN to transform COVID-19 CT sections into healthy CT sections.
- ▶ Employ two losses simultaneously: one for mapping healthy-to-healthy sections and another for generating semi-healthy sections from infected ones.
- ▶ Distinguish real healthy CT scans from synthesized healthy scans.

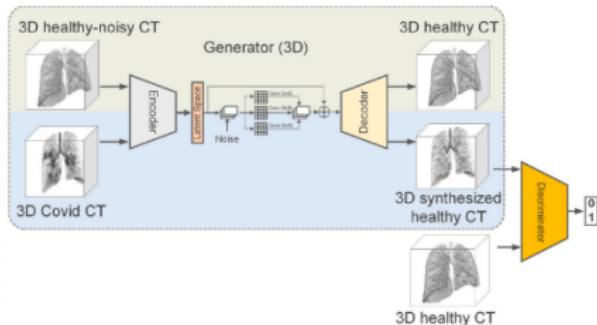


Figure 6 : 3D GAN

## Methods

### Phase 1: 3D GAN for Synthesizing Pseudo-Masks

To build an effective 3D GAN, two main challenges were addressed:

- ▶ **Ensuring the model's generality for various infections:**

The model's architecture was modified to include a noise adder (Perlin Noise) operator in the encoder-decoder's latent space, which is represented as follows:

$$\hat{I}_H = G(I_C) = De(En(I_C) + U_N), \quad (1)$$

where  $En : I_C \in \mathbb{R}^{W \times H \times L} \rightarrow L_N$  symbolizes the encoder,  
 $De : L_N \rightarrow \hat{I}_H \in \mathbb{R}^{W \times H \times L}$  is the decoder, and  $U_N$  is the noise operator. The noise adder perturbs the latent space to train the generator to remove a wide variety of infections.

## Methods

### Phase 1: 3D GAN for Synthesizing Pseudo-Masks

#### ► Maintaining the original lung structure in the generated sections

we updated the model's weights by incorporating healthy-noisy sections into the training phase. These images were created by adding Perlin noise to healthy 3D CT scans.

The loss function for the proposed 3D GAN is defined as:

$$\begin{aligned} \max_D V(D, G) = & \mathbb{E}_{I_H \sim p_{\text{healthy}}} [\log D(I_H)] \\ & + \mathbb{E}_{I_C \sim p_{\text{covid}}} [\log(1 - D(G(I_C)))] \\ & + \mathbb{E}_{I_H^{Pe} \sim p_{\text{noisy\_Healthy}}} [G(I_H^{Pe}) - I_{H_2}], \end{aligned} \tag{2}$$

where  $I_H$  denotes the real healthy 3D volume, and  $I_H^{Pe}$  is the healthy-noisy image with Perlin noise. By training with both losses, the 3D GAN learns to generate a variety of healthy outputs corresponding to the COVID-19 input images.

## Methods

### Phase 2: Synthesizing Pseudo Masks

- ▶ Synthesize pseudo masks by subtracting the generated healthy CT scans from the original COVID-19 CT scans.
- ▶ Pseudo masks emphasize the infected regions within the CT scans.
- To achieve this, we employ the pseudo mask, denoted as  $\hat{M}$ , which is computed as follows:

$$\hat{M} = I_C - \hat{I}_H \quad (3)$$

Where:

- $I_C$  represents the COVID-positive CT scan, which is an entry of the trained 3D GAN model generator.
- $\hat{I}_H$  represents the output of the 3D GAN, consisting of synthesized pseudo-healthy images.

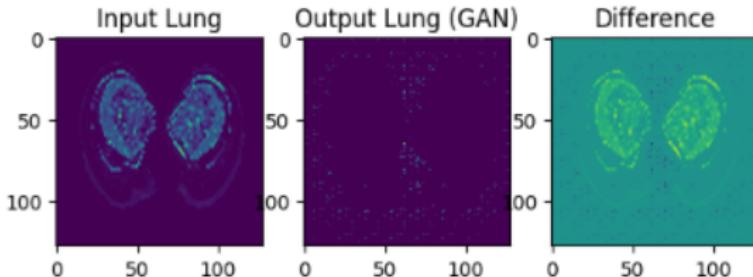


Figure 7 : Example of Mask extraction

## Methods

### Phase 3: Segmenting the CT scans using TransUNET

- ▶ Adapt the segmentation phase by incorporating the TransUNET architecture for precise COVID-19 CT image segmentation.
- ▶ Utilize the generated pseudo masks from the previous step as input data for training TransUNET segmentation model.
- ▶ Highlight TransUNET's capabilities in effectively segmenting COVID-19-infected regions with exceptional accuracy.

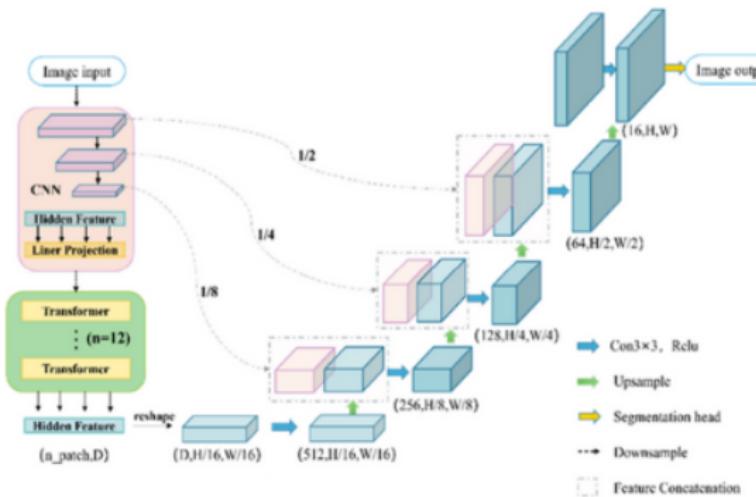


Figure 8: Framework Overview - TransUNet Architecture

# Enhanced TransUNET Segmentation Pipeline Incorporating Transformers

---

**Algorithm 1** Enhanced TransUNET Segmentation Pipeline Incorporating Transformers
 

---

- 1: **Inputs:**  $I_c$ : COVID-19 3D volume,  $I_h$ : Healthy 3D volume
  - 2: **Input:**  $I_p$ : High infected COVID-19 2D slices,  $I_n$ : Low infected COVID-19 2D slices
  - 3: **Initialize:**  $G$ : 3D Generator,  $D$ : 3D Discriminator
  - 4: **Initialize:**  $U_{ed}$ : TransUNET (UNet with Transformer modules),  $U_e$  referring to the encoder part only
  - 5: **Initialize:** *SEL*: Sensitivity Enhanced Loss, *ECL*: Enhanced Contrastive Loss  
for each epoch in the first training phase (3D GAN G-D Model training) do
  - 6: Freeze  $G$ , Unfreeze  $D$
  - 7: Calculate generated healthy volumes  $\hat{I}_H$  using  $G(I_c)$
  - 8: Update  $D$  using  $I_h$  and  $\hat{I}_H$ , by calculating adversarial loss for  $D$  as  
$$\nabla_{\theta_D} \mathbb{E}[(\log D(I_h) + \log(1 - D(\hat{I}_h)))]$$
  - 9: Freeze  $D$ , Unfreeze  $G$
  - 10: Update  $G$  using  $I_c$ , by calculating adversarial loss for  $G$  as  
$$\nabla_{\theta_G} (\log(1 - D(\hat{I}_h)))$$
  - 11: Add Perlin Noise to the healthy CT volumes to obtain  $I_H^{Pe}$
  - 12: Update  $G$  using MSE loss of pairs of  $\hat{I}_h^{Pe}$  and  $I_h$  as input and output, respectively  
$$\nabla_{\theta_G} \mathbb{E}[\|G(\hat{I}_h^{Pe}) - I_h\|_2]$$
  - 13:  
for each epoch in the second training phase (2D Segmentation Model training)  
do
  - 14: Perform a forward pass of the batch through TransUNet
  - 15: Calculate the contrastive loss
  - 16: Calculate the mean squared error (MSE) loss
  - 17: Calculate the sensitivity-enhanced loss
  - 18: Combine losses and perform backpropagation
  - 19: Update model weights with the optimizer
  - 20: Periodically plot loss graphs and sample segmentations
  - 21: Save model checkpoints
  - 22:
  - 23: **Output:** Segmented infection volumes
  - 24: **Output:** Updated  $G$  and  $U_{ed}$  parameters
-

# Methods

## Loss functions

Our advanced segmentation model employs two novel loss functions, each meticulously crafted to enhance the model's performance in segmenting COVID-19 infections in lung CT scans. Here, we present the mathematical formulations of these loss functions.

### Enhanced Contrastive Loss

$$L_{\text{contrastive}} = - \log \frac{\sum_{i=1}^N \exp\left(\frac{\mathbf{f}_i \cdot \mathbf{f}_p}{\tau}\right)}{\sum_{i=1}^N \sum_{j=1}^N \exp\left(\frac{\mathbf{f}_i \cdot \mathbf{f}_j}{\tau}\right)} \quad (4)$$

where  $\mathbf{f}_i$  and  $\mathbf{f}_j$  are the flattened and normalized feature vectors of the  $i^{th}$  and  $j^{th}$  samples, respectively,  $\mathbf{f}_p$  represents the feature vector of a positive sample,  $\tau$  is the temperature parameter, and  $N$  is the number of samples.

# Methods

## Loss functions

When determining the optimal value for  $\tau$ , stochastic gradient descent can be employed:

### 1. Initialize parameters:

Initialize the model parameters randomly or with specific values.

### 2. Initialize initial learning rate:

Choose an initial learning rate ( $\eta_0$ ).

### 3. Repeat until convergence or maximum iterations reached:

#### 3.1 Randomly choose a pair of examples ( $\mathbf{f}_i, \mathbf{f}_p$ ):

#### 3.2 Compute the partial gradient:

$$\nabla J(\tau) = \frac{\mathbf{f}_i \cdot \mathbf{f}_p}{\tau^2} \left( \frac{\exp(\mathbf{f}_i \cdot \mathbf{f}_p / \tau)}{\sum_{k=1}^N \exp(\mathbf{f}_k \cdot \mathbf{f}_p / \tau)} - \frac{\sum_{j=1}^N \exp(\mathbf{f}_i \cdot \mathbf{f}_j / \tau)}{(\sum_{k=1}^N \sum_{l=1}^N \exp(\mathbf{f}_k \cdot \mathbf{f}_l / \tau))} \right)$$

#### 3.3 Update temperature ( $\tau$ ) using a variable learning rate ( $\eta_t$ ):

$$\tau \leftarrow \tau - \eta_t \nabla J(\tau)$$

### 4. Stopping condition:

Stop if the maximum number of iterations is reached.

# Methods

## Sensitivity Enhanced Loss

### ► Sensitivity Enhanced Loss

- Aims to reduce false negatives in medical diagnostics.
- Formula:

$$L_{\text{sensitivity}} = \frac{1}{N} \sum_{i=1}^N [BCE(\mathbf{p}_i, \mathbf{t}_i) \times (1 + \beta \times FN(\mathbf{p}_i, \mathbf{t}_i))]$$

- $\mathbf{p}_i$ : Predicted probability for  $i^{th}$  sample.
- $\mathbf{t}_i$ : True label for  $i^{th}$  sample.
- $\beta$ : Weight amplifying the impact of false negatives.
- $N$ : Total number of samples.

### ► BCE (Binary Cross-Entropy) Loss

- Used in the sensitivity enhanced loss.
- Measures the difference between predicted and true labels.

### **Metrics used for Evaluating Our Pipeline's performance**

- 1. Dice Score (DSC):** The Dice Score quantifies the spatial overlap between predicted and ground truth segmentation.
- 2. Sensitivity:** Sensitivity measures the model's ability to correctly identify positive instances.
- 3. Specificity:** Specificity gauges the model's capacity to accurately identify negative instances.

**Each metric is calculated per slice, and the maximum value across all slices is reported.**

## Metrics Formulation

### 1. Dice Score (DSC):

$$DSC = \frac{2 \times TP}{2 \times TP + FP + FN}$$

where  $TP$  is true positives,  $FN$  is false negatives.

### 2. Sensitivity:

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

where  $TP$  is true positives,  $FN$  is false negatives.

### 3. Specificity:

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

where  $TN$  is true negatives,  $FP$  is false positives.

Each metric is calculated per slice, and the reported value represents the maximum across all slices.

## Results and Discussion

- ▶ **Preprocessing of CT Images:** Adjusted size and focus for clarity and consistency. Dataset comprises 50 COVID-positive and 51 COVID-negative CT scans.
- ▶ **Cropped Lung Images:** Focused on lung areas to enhance visibility of potential COVID-19 indicators.
- ▶ **Data Organization:** Categorized images into two groups: COVID-19 positive and normal scans.
- ▶ **Format Conversion:** Transformed images and masks into numpy arrays for efficient computer analysis.
- ▶ **Data Storage:** Saved numpy arrays separately, including healthy lung images, infection images, lung masks, and original CT scans. Organization critical for model training and testing with TransUNET-based segmentation.

## Preprocessing Mosmed CT scans and Hyperparameters Setting

- ▶ Resizing followed, where we adjusted each scan to a uniform size while maintaining the aspect ratio to prevent image distortion.
- ▶ We performed resampling of CT pixels to standardize the pixel spacing across all scans.
- ▶ Ensure that each voxel represented consistent physical dimensions ,a key factor for accurate analysis.

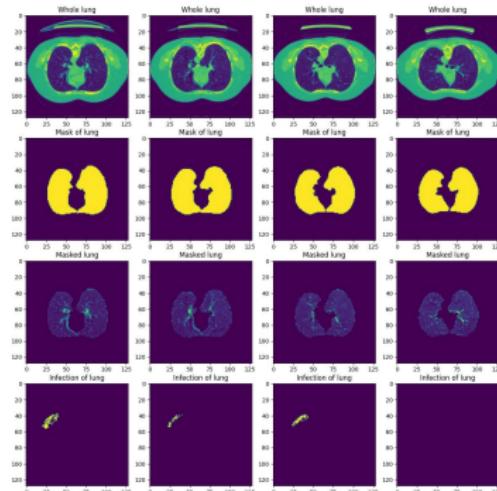


Figure 9:Cropped Lungs by Extracted Masks - Mosmed Data Preprocessing Phase

### 3D GAN Model Training Results

The Figure shows the result of the 3D GAN synthesizing a pseudo-healthy image, to later subtract it from the covid-infected image in order to obtain the pseudo-mask highlighting regions of infection for the training of the 2D segmentation model.

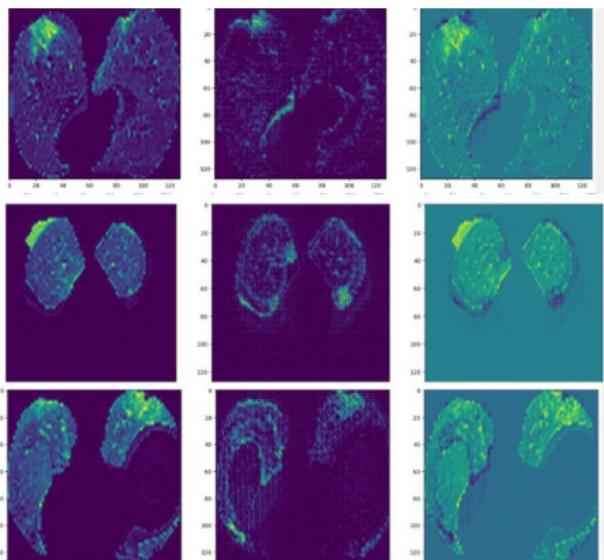


Figure 10:Synthesizing semi-healthy medical images

## TransUNET Segmentation Encoder-Decoder results

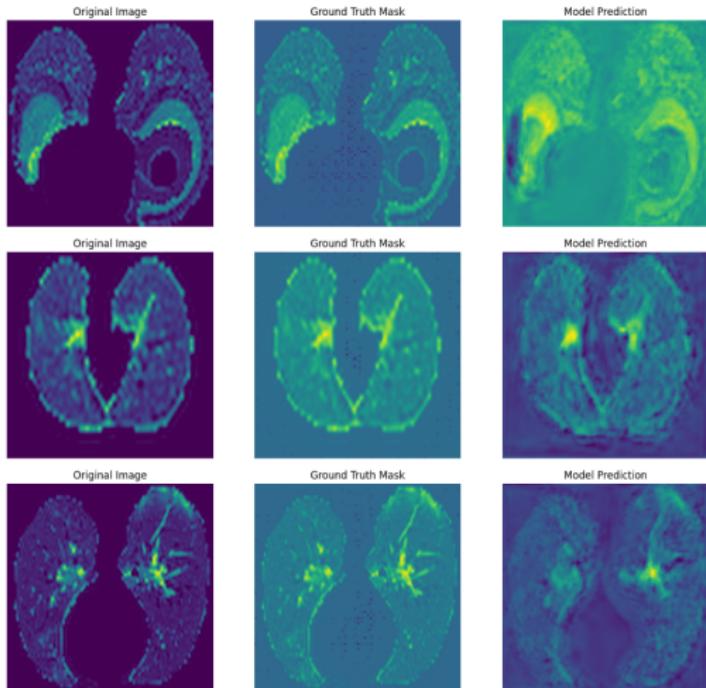


Figure 11: TransUNET Segmentation results

## Validation Set Overview

In the Mosmed Dataset, our validation set is comprised of 50 COVID-19 CT scans. These scans have ground-truth masking of the infection regions, which is critical for assessing the performance of our image segmentation models.

## Evaluation Methodology

### Approach:

- ▶ Iterate over each CT scan slice in the dataset.
- ▶ Calculate Infection Percent: Ratio of infected area to total lung area per slice.
- ▶ Apply a thresholding method to segment the infected regions.
- ▶ Reshape the masks for metric calculations.
- ▶ Compute Dice Score, Sensitivity, and Specificity for each slice.
- ▶ Store these scores to analyze overall model performance.

## Results and Discussion

### Model Performance Comparison

Comparison of Model Performances

Metric	GAN + Perlin Noise + TransUNET		GAN + Perlin Noise + UNET	
	Mean (%)	Std (%)	Mean (%)	Std (%)
Dice Score	50.09	45.44	<b>72.89</b>	46.69
Sensitivity	<b>98.52</b>	11.29	72.89	46.69
Specificity	<b>99.99</b>	1.11e-14	99.63	6.05

We implemented and tested the complete pipeline on [Google Colab Pro](#).

## Conclusion

Notably, our investigation has outperformed comparable weakly supervised models, demonstrating superior performance on a test dataset, particularly in regions of low infection within CT slices. Despite the utilization of a 3D GAN during training, our final 2D segmentation model offers increased flexibility and faster processing times for clinicians. The model exhibits potential applications in diverse medical imaging scenarios beyond the scope of COVID-19 CT segmentation.

## Future Work

- ▶ **Enhancement of the GAN Model:** Consider refining our GAN model, possibly exploring advanced architectures such as TransGAN to further improve its performance.
- ▶ **Exploration of Alternative Loss Functions:** Investigate the incorporation of alternative loss functions to potentially enhance the model's learning capabilities and segmentation accuracy.
- ▶ **Integration of Multi-Modal Information:** Explore the integration of multi-modal information, such as combining CT scans with other imaging modalities, to enhance the overall robustness of the segmentation model.
- ▶ **Clinical Validation and Deployment:** Conduct thorough clinical validation studies and explore possibilities for deploying the model in real-world clinical settings, ensuring its efficacy and reliability.
- ▶ **Collaboration with Domain Experts:** Collaborate with domain experts and medical professionals to refine the model based on clinical feedback and insights, ensuring alignment with practical medical requirements.

## References

- ▶ Azad, R., Khodapanah Aghdam, E., Rauland, A., Jia, Y., Haddadi Avval, A., Bozorgpour, A., ... Merhof, D. (2022). *Medical Image Segmentation Review: The Success of U-Net*. <https://arxiv.org/pdf/2211.14830.pdf>
- ▶ Moriya, T., Roth, H. R., Nakamura, S., Oda, H., Nagara, K., Oda, M., Mori, K. (2022). *Unsupervised Segmentation of 3D Medical Images Based on Clustering and Deep Representation Learning*. <https://arxiv.org/pdf/1804.03830.pdf>
- ▶ Shabani, S., Homayounfar, M., Vardhanabhuti, V., Nikouei Mahani, M.-A., Koohi-Moghadam, M. (2021). *Self-supervised region-aware segmentation of COVID-19 CT images using 3D GAN and contrastive learning*. *Computers in Biology and Medicine*, 149.
- ▶ mkara44. (2023). *transUNet implementation for PyTorch by mkara44 on GitHub*. Available at: [https://github.com/mkara44/transunet\\_pytorch](https://github.com/mkara44/transunet_pytorch) (Accessed on 2023-12-13).
- ▶ Mosmed AI. (2023). *MosMedData-CT-COVID19-type VII-v 2*. Available at: <https://mosmed.ai/en/datasets/covid191110/> (Accessed on 2023-10-26).