

Infection Area Segmentation on X-ray Images

Phan Thanh Bình BI12-059

I. INTRODUCTION

X-ray refers to the electromagnetic sound wave which has the wavelengths ranging from $10pm$ to $10nm$. This technology works based on the tissue differential contrast generated by X-ray and tissue interaction. Using X-ray for diagnostics is an extremely essential tool in modern medicine. One of the applications of X-ray is in medical imaging. X-ray is generated by bombarding a metal target with energetic electrons. Upon impact with the metal target, the incident electrons collide with the electrons and nuclei of the metal atoms. The collisions slow down and deflect the incident electrons. The rapid increase in the momentum of the incident electrons produce large acceleration values and deflected incident electrons emit X-rays. The generated X-rays are then passed through the patient's body and is recorded by the 2D sensor. X-ray image is finally produced with the usage of projection radiography.

COVID-19 is a highly infectious lung disease caused by the corona virus. It was first identified in Wuhan, China, and rapidly spread worldwide, leading to a global pandemic. COVID-19 can cause a wide range of symptoms, from mild flu-like symptoms to severe respiratory distress and death, particularly in older adults and those with declined health conditions. Moreover, the disease can be easily spread through droplets when infected people cough, sneeze, talks or by touching contaminated surfaces. Thus it is vital to isolate infected people as quickly as possible in order to reduce the chance of diffusing the disease unintentionally. One of many ways to achieve this is by distinguishing X-ray lung image of infected versus normal person. In this work, we will:

- Explore the COVID-QU-Ex Dataset
- Build a deep learning model to perform segmentation of the infection area from the input lung image.

II. DATASET

This dataset contains 2 sub datasets: *Infection Segmentation* and *Lung Segmentation* data. The files structure can be view in figure 1. To summarize, both sub datasets contains *train-val-test* splits each split with chest images and ground truth lung segmentation map for 3 types of patients: COVID, non-COVID(other diseases), normal. However, only the *Infection Segmentation* provide the infected area map, that is why for our work, we focus on this sub dataset. Figure 2 shows examples of the *Infection Segmentation* dataset. One key note is that the infected map segmentation for both the non-COVID and normal patients are filled with *zeros*. The statistics about the number of images are shown in table I. Fortunately, the number of positive and negative samples are equal.

```
Infection\ Segmentation\ Data
-- Infection\ Segmentation\ Data
-- Test
-- COVID-19
-- Non-COVID
-- Normal
-- Train
-- COVID-19
-- Non-COVID
-- Normal
-- Val
-- COVID-19
-- Non-COVID
-- Normal
Lung\ Segmentation\ Data
-- Lung\ Segmentation\ Data
-- Test
-- COVID-19
-- Non-COVID
-- Normal
-- Train
-- COVID-19
-- Non-COVID
-- Normal
-- Val
-- COVID-19
-- Non-COVID
-- Normal
```

Fig. 1. Files Structure

	Train	Val	Test
COVID	1864	466	583
Non-COVID	932	233	292
Normal	932	233	291

TABLE I
NUMBER OF IMAGES STATISTICS

III. METHOD

A. Architecture

Our task is quite straightforward: given a chest X-ray image, we want to have a segmentation of the infected area. For that reason, we utilize the architecture design similar to U-Net[1]. We changes the input image size to 256×256 and only perform 3 times up/down sampling. The details can be viewed in figure 3. This architecture follows the *encoder-decoder* scheme to obtain the segmentation map. The features extractor mechanisms is done by two convolutional layers combine with one maxpool layer similar to VGG [2]. The concatenation of features used in upsampling phase before continue performing convolution works as a method to avoid vanishing gradients problem.

B. Training

We train the model on the train and val splits and validate on the test split. The model is trained on 10 epochs with *binary*



Fig. 2. Dataset visualizations

cross entropy loss, using *Adam optimizer* with *learning rate* 0.0001 and *batch size* 16.

IV. EVALUATION

This step is the evaluation of the segmentation model. The metric we use is the average of *Root Mean Squared Error (RMSE)*. *RMSE* metric simply determines how well the model's prediction similar to the ground truth segmented map by calculating the distance between the prediction and the ground truth.

$$RMSE = \frac{1}{N} \sum_{i=1}^N \sqrt{(x_p - x_t)^2}$$

where x_p is the prediction x_t is the ground truth.

Our model achieve the score of **0.0695**

V. CONCLUSION

In this work, we have examine the dataset from COVID-QU-Ex Dataset. However, due to the inconsistency of our proposed formula with the ground truth annotations, we can only evaluate on the segmentation task in our pipeline.

REFERENCES

- [1] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. "U-net: Convolutional networks for biomedical image segmentation". In: vol. 9351. 2015. DOI: 10.1007/978-3-319-24574-4_28.
- [2] Karen Simonyan and Andrew Zisserman. "Very deep convolutional networks for large-scale image recognition". In: 2015.

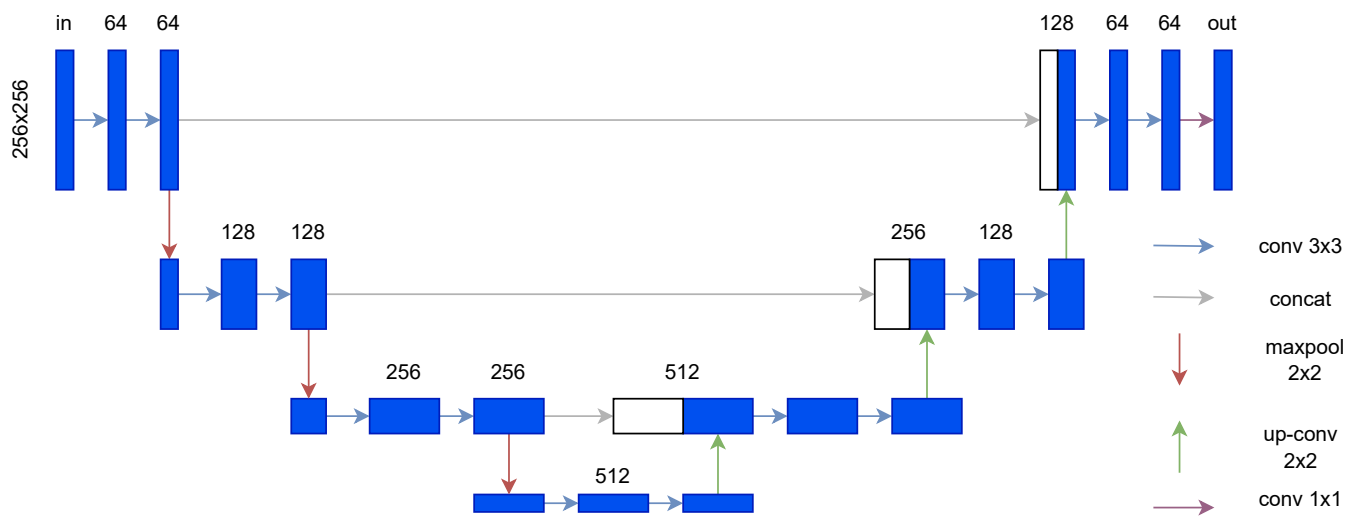


Fig. 3. Model architecture