Segmentation of COVID19 X-ray images

Nguyen Minh Dat
Machine Learning in Medicine
USTH University
Email: datnm.23bi14089@usth.edu.vn

I. INTRODUCTION

The COVID-19 pandemic spread rapidly from 2020 to 2024. According to statistics from the Vietnamese Ministry of Health, by June 2023 there were 11,619,451 confirmed cases, including 43,206 deaths, accounting for about 0.4% of total infections. This was a very dangerous pandemic. Many scientific studies have shown that COVID-19 can cause damage to internal organs, especially the lungs. Therefore, this study aims to identify and mark damaged regions in the lungs of COVID-19 patients.

II. DATA DESCRIPTION

A. Data organization

The dataset used in this study is the COVID-QU-Ex dataset from Kaggle, which contains X-ray images of patients. This dataset consists of two parts: Lung Segmentation Data and COVID-19 Infection Segmentation Data. In this research, only the COVID-19 Infection Segmentation Data is used to detect damaged regions. The dataset is organized with the following directory structure:

```
Infection Segmentation Data/
 train/
      COVID-19/
            images/
            infection masks
            lung masks
      Non-COVID/
            images/
            infection masks
            lung masks
      Normal/
            images/
            infection masks
            lung masks
 val/
      COVID-19/
      Non-COVID/
 test/
      COVID-19/
      Non-COVID/
```

B. Data training

I use two data groups, COVID and non-COVID, for model training. Each group is split into training, validation, and test sets, including X-ray images and corresponding infection

TABLE I: Dataset description

Name	Number of images
Train	1864 COVID and 932 non-COVID
Validation	466 COVID and 233 non-COVID
Test	583 COVID and 292 non-COVID

masks indicating lung damage. In infection masks, COVID patients show white regions representing infected lung areas, whereas non-COVID patients do not show any marked regions. This data selection is well-balanced and suitable for training the proposed model.

III. MODEL SELECTION

In this study, an Attention U-Net model is used to segment COVID-19 infection regions from chest X-ray images. By adding attention gates to the skip connections, the model can focus more on important infected areas and reduce the influence of background and healthy tissues. This architecture is well-suited for medical image segmentation, especially when infected regions are small and imbalanced, helping to improve localization and overall segmentation performance. In addition, I also experimented with another U-Net variant called ResNet U-Net, which was trained by my friend, Pham Minh Hieu. We used the same dataset, and the goal was to compare the differences between his model and mine under the same conditions.

A. Attention U-net

Attention U-Net is built on the standard U-Net architecture. The encoder consists of regular convolution blocks. Attention Gates are integrated into the skip connections. These attention gates combine information from both the encoder and the decoder to create attention maps and filter out unimportant features. The decoder then uses these attention-filtered features to reconstruct the final mask.

Strengths This model can focus well on important lesion regions and reduce the effect of background noise. It is especially effective when the lesions are small, have blurry boundaries, or are unevenly distributed. Therefore, it is very suitable for high-precision medical image segmentation tasks.

B. ResNet U-Net model

Architecture ResNet U-Net is based on the U-Net structure with an encoder, bottleneck, and decoder. The encoder is replaced by a ResNet backbone (for example, ResNet-34) pretrained on ImageNet. It uses residual blocks in the encoder

to extract deeper features. Skip connections directly transfer feature maps from the encoder to the decoder. The decoder contains upsampling and convolution layers to restore the image resolution. The final layer is a 1×1 convolution that generates a binary mask.

Strengths This model has a strong feature extraction ability thanks to the pretrained ResNet. It usually converges faster and more stably during training. It works well with small to medium-sized datasets and is widely used and validated in medical image segmentation tasks.

C. Overall Comparison

ResNet U-Net mainly focuses on strong feature extraction and is suitable when a stable and easy-to-train model is needed. In contrast, Attention U-Net concentrates more on important regions in the image, making it more appropriate when high accuracy is required, especially for small lesion areas.

IV. RESULTS

A. Results of Attention U-Net

1) Result of training: These are the training results after 30 epochs:

TABLE II: Training performance of Attention U-Net over epochs
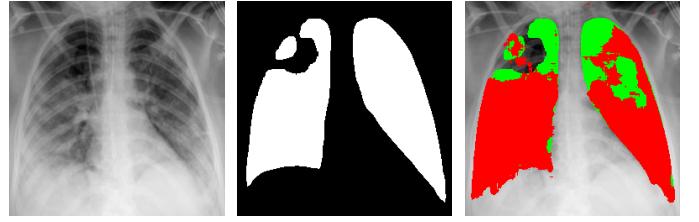
Epoch	Train Loss	Train Dice
1	0.6550	0.3460
2	0.5581	0.3900
3	0.5163	0.4060
4	0.4945	0.4179
5	0.4802	0.4271
⋮	⋮	⋮
25	0.3805	0.5678
26	0.3763	0.5919
27	0.3699	0.5944
28	0.3692	0.5968
29	0.3656	0.6081
30	0.3675	0.5930

TABLE III: Validation performance of Attention U-Net over epochs

| Epoch	Val Loss	Val COVID Dice	Val Non-COVID Dice
1	0.5712	0.6006	0.0000
2	0.5040	0.6248	0.0000
3	0.4718	0.6502	0.0000
4	0.4657	0.6790	0.0086
5	0.4659	0.6827	0.0687
⋮	⋮	⋮	⋮
25	0.3791	0.7182	0.4893
26	0.3688	0.7192	0.4421
27	0.3893	0.6921	0.0558
28	0.3764	0.7309	0.6137
29	0.4092	0.7028	0.6524
30	0.3946	0.7275	0.4464
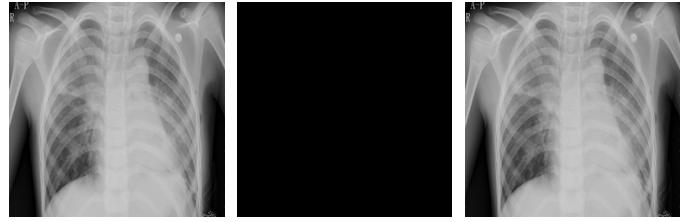
In the final epochs, the model converged stably with low training and validation loss. The training Dice stayed around 0.59–0.61, while the validation COVID Dice remained high and stable at about 0.72–0.73, indicating good segmentation performance and generalization. The validation of the Non-COVID Dice also improved compared to the early stage. No serious overfitting was observed, and the best overall performance appeared in the last epochs, making them suitable for final model selection.

(a) Original image (b) Ground truth mask (c) Prediction overlay

Fig. 1: COVID sample: original image, ground truth mask, and predicted segmentation overlay

(a) Original image (b) Ground truth (c) Prediction overlay

Fig. 2: Non-COVID sample: original image, ground truth mask, and predicted segmentation overlay

TABLE IV: Overall prediction performance on the test set

Metric	Value
Mean Dice	0.6926
Mean IoU	0.6095
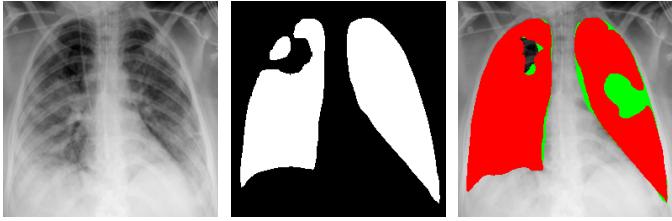
I observed that the red predicted regions do not completely match the ground truth masks when representing the infected areas of COVID-19 patients. For non-COVID patients, the predicted masks usually contain no red regions or only very small ones. This indicates that the model performs better on the non-COVID set than on the COVID cases.

B. Results of ResNet U-Net

TABLE V: Prediction performance for COVID and Non-COVID cases
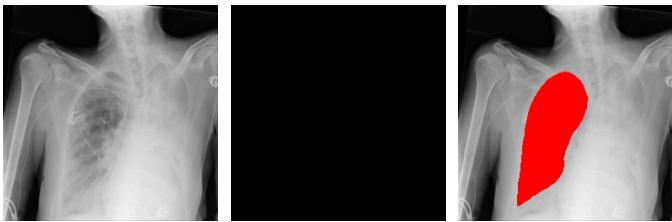
Category	Mean Dice	Mean IoU
COVID	0.7701	0.6650
Non-COVID	0.2911	0.2911

For the ResNet U-Net model, I consider the predicted masks for COVID patients to be better than those from the Attention U-Net. The red-highlighted regions cover the infected areas more completely and are closer to the ground-truth labels. However, on the non-COVID dataset, the model is less accurate and does not learn the patterns as well.

(a) Original image   (b) Ground truth   (c) Prediction overlay

Fig. 3: COVID sample: original CT image, ground truth infection mask, and predicted segmentation overlay

(a) Original image   (b) Ground truth   (c) Prediction overlay

Fig. 4: Non-COVID sample: original CT image, ground truth infection mask, and predicted segmentation overlay

V. CONCLUSION

This study explored COVID-19 lung infection segmentation on chest X-ray images using the COVID-QU-Ex dataset with two models: Attention U-Net and ResNet U-Net. Both models converged stably and learned meaningful features. Attention U-Net showed more balanced performance and better results on non-COVID samples, while ResNet U-Net produced more accurate and complete segmentation on COVID infection regions but was weaker on non-COVID data. Overall, each model has its own strength, and the choice depends on whether the goal is balanced performance or higher accuracy in detecting COVID lesions.