

Robust Method for 3D Lesion Segmentation of COVID-19 from Lung CT Scans

Zhaozhong Wang^{1†}, S. Shailja²

¹*Homestead High School, 21370 Homestead Rd, Cupertino, CA 95014*

²*Department of Electrical and Computer Engineering, University of California, Santa Barbara, CA 93106*

[†]*corresponding author: alex-wang@berkeley.edu*

Computed Tomography (CT) is an imaging technique to obtain internal body images and can reflect infections due to COVID-19, which primarily affects the human lungs. To improve the accuracy of CT classification processes distinguishing COVID-19 from other viral pneumonia, precise segmentation of the region of interest is essential. We propose a 3D nnU-Net-based segmentation approach targeting lung lesions to free radiologists from completing time-consuming manual segmentation. With CT data from publicly available datasets and the Santa Barbara Cottage Hospital, we examine the use of segmented lobe masks and lung masks as post-processing through mode Dice scores and other statistics. We present a robust 3D lesion segmentation pipeline that exploits lobe localizations to improve performance. Our experimental results demonstrate our pipeline's improvements upon the baseline model. Moreover, we observe the model's robustness and transferability to general COVID-19 CT datasets. The source code for our implementation is available on GitHub.*

Keywords: Semantic Segmentation, Computed Tomography, Medical Imaging, CNN, U-Net

I. INTRODUCTION

The COVID-19 pandemic has caused devastating infections, deaths, and economic losses. Early identification of suspected patients can be achieved with rapid screening, which is critical to curb the pandemic's proliferation [1]. Though COVID-19 can manifest in many organs of the human body, the lungs are most commonly affected [2]. These manifestations have been reported across both the adult and pediatric populations [3], [4].

Chest imaging has been critical for COVID-19 identification and patient management [5]. RT-PCR (Reverse Transcription Polymerase Chain Reaction) has been the most commonly used procedure for COVID-19 detection, and functions by detecting nucleotides from specimens obtained by oropharyngeal swab [6]. Computed Tomography (CT) is an imaging technique used in radiology to obtain internal images of the body. Traditionally, CT has been used to yield cross-sectional images, or slices, of the examined area. With the development of 3D imaging, CT can also be used to acquire volumetric datasets one or multiple slices at once. Prior work indicates common CT findings across patients with COVID-19 in the form of characteristic radiologic features. These features include ground-glass opacities, multifocal patchy consolidation, and interstitial changes with a peripheral distribution [7].

Compared to RT-PCR testing, lung CT offers greater sensitivity for COVID-19 diagnosis, and has demonstrated

promise to be considered as the primary tool for detection in epidemic areas [8]. Prior studies suggest that CT holds capability to reveal pulmonary abnormalities in COVID-19-infected patients with initial negative RT-PCR results [9]. These results place CT as a promising approach for robust COVID-19 detection, compared to RT-PCR with its limited sensitivity.

In most procedures in the domain of biomedical image processing, segmentation serves as the first step towards analysis [10]. Precise definition of the region of interest can simplify the classification task and substantially improve classification accuracy, achieved through robust segmentation algorithms.

Deep neural networks have achieved remarkable performance in segmentation tasks, with convolutional neural networks (CNNs) serving as state-of-the-art models in fields such as tumor segmentation [11], and vessel segmentation [12]. In particular, the U-Net architecture has gained recognition for its excellent segmentation performance [13]. U-Net provides a baseline for image segmentation models, and has inspired numerous modified architectures. Namely, the nnU-Net is a proposed framework that automatically adapts to arbitrary biomedical datasets, and therefore serves as a reliable baseline to compare future models [14].

In this paper, we improve upon U-Net and compare our model with the baseline model. Utilizing a novel dataset from the Santa Barbara Cottage Hospital, we examine a

*<https://github.com/AlexWang05/covid-segmentation>

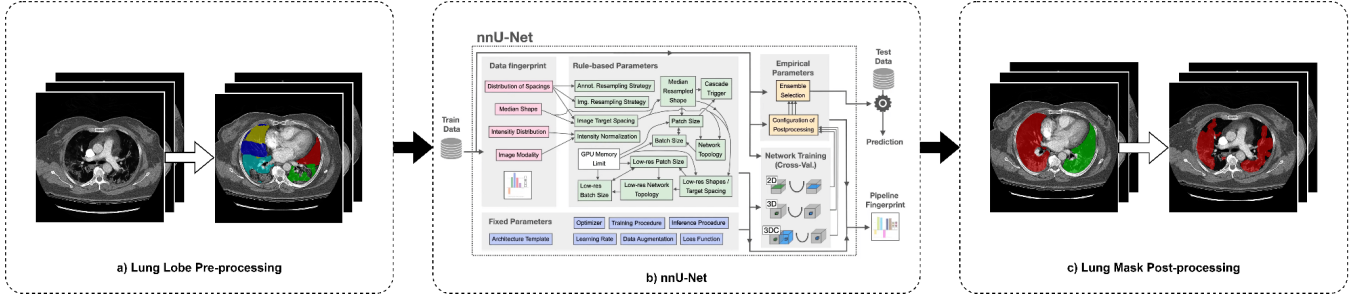


FIG. 1. Overall workflow of our approach. (a) pre-processing step adding lung lobe information to CT dataset. (b) nnU-net architecture selecting the best performing model based on Dice scores [15]. (c) post-processing step removing extraneous lesion segmentation.

nnU-Net based architecture and evaluate the effect of including lobe masks as additional input and as a post-processing layer. We evaluate our methods through collected statistics, including model dice scores, accuracy, precision, sensitivity, and specificity on training, inference, and test datasets.

II. METHODS

We hypothesize that a well-trained U-Net-based network can achieve robustness and satisfactory performance on the novel dataset collected from the Santa Barbara Cottage Hospital. In this section, we describe the details of our proposed methods to achieve a robust and transferable model that improves upon the baseline U-Net architecture. We include the data description, network architecture, data post-processing, and evaluation metrics. The overall workflow is illustrated in Fig. 1. The source code for the implementation is available on GitHub [16].

A. Dataset

We train our networks with datasets collected from the COVID-19 Lung CT Lesion Segmentation Challenge - 2020 (COVID-19-20), which includes CT data provided by the Multi-national NIH Consortium for CT AI in COVID-19 [17]. The training and validation data include unenhanced chest CT scans from 199 and 50 patients, respectively. These CT scans are collected from patients with positive RT-PCR test results for COVID-19; they include ground truth annotations of COVID-19 lesions within the provided CT scans.

Initial evaluation of trained models were completed with the testing dataset provided by the COVID-19-20 challenge, which includes data from 46 unseen patients with positive RT-PCR COVID-19 results; the test dataset includes ground truth annotations of COVID-19 lesions and come from a variety of sources, including those not used for training and validation.

We complete evaluations of our methods with a de-identified dataset from the Santa Barbara Cottage Hospital. This is a retroactive study, with all images

de-identified as specified by the IRB agreement between the medical center and the University of California, Santa Barbara. We examine data from a dataset collected from a group including patients with confirmed COVID-19 by SARS-CoV2 RT-PCR testing and or patients with positive IgM or IgG following a symptomatic period. Scans were acquired as part of the treatment process. Patients with multiple positive respiratory PCR panels are excluded (i.e. patients with influenza and COVID-19). Other exclusion criteria include patients with incomplete data, duplicated patient subsets, and patient's whose data may be difficult or impossible to anonymize. 104 CT scans of 104 patients from the Santa Barbara Cottage Hospital were included in the study, with all patients having a diagnosis of COVID-19; 31 patients had scans that were manually classified to identify 2D slices where lesions were present, while no ground truth lesion annotations were included in the provided dataset. Number of CT slices per subject varied from 45 to 251. The greater the number of slices, the higher the resolution in the z-dimension.

B. Network Architecture

We examine different segmentation architectures adapted from U-Net and nnU-Net in our study [13], [14]. U-Net-based architectures are widely used for volumetric image segmentation tasks [17], [18], [19]. We develop our network based on the MONAI U-Net COVID-19-20 challenge baseline and the dynamic U-Net (DynUNet) [20]. MONAI is a PyTorch-based open-source framework for deep learning in healthcare used to develop advanced training workflows with a focus on medical imaging.

We employ the base U-Net architecture to serve as a point of comparison for future model performance. Intensity and spatial random augmentations were computed using MONAI; data augmentations are randomized at every training iteration. Segmentations are generated by MONAI's sliding window inferior with a window size of $192 \times 192 \times 16$; the segmentation network has two output channels, denoting background and foreground, and one input channel, denoting the intensity value of the input data.