A deep learning approach for automatic segmentation of ground-glass opacities and consolidation in small Chest-CT dataset

Alvaro A. Sandino[a], Luisa Fernanda Vargas[a], Sebastian Maglioni[a], Mariajosé Rangel[a], Ana Maria Alvarado[b], Jose Luis Quintana[b], Andres Vargas[b], Jorge Carrillo[b], Jorge Marín[a, b], and Jorge Rudas[a, c]

[a]Visualization & Artificial Intelligent Team, ImexHS, Bogotá, Colombia
[b]Department of Diagnostics Medicine, RIMAB, Bogotá, Colombia
[c]Computational Modeling of Biological System Group, Department of Mathematics, Universidad Nacional de Colombia, Bogotá, Colombia

ABSTRACT

Ground-glass opacities (GGO) and consolidation are findings typically observed in pneumonia and other acute lung diseases. For radiologists, segmentation and analysis of GGO and consolidation on chest computed tomography (CT) images are critical to quantify the disease severity and assess stages of recovery. However, these disease areas are difficult to detect, differentiate and segment on CT scans as these usually do not have clear boundaries. Furthermore, radiologists' manual segmentations of lung lesions are a highly subjective and time-consuming task considering that a high-resolution Chest-CT study comprises hundreds of images. In this paper, we present a deep learning approach that automatically segments GGO and consolidation in Chest-CT studies. Our approach first segments the region of interest (ROI), lung parenchyma, achieving a Dice similarity coefficient of $0,97(\pm 0.02)$ and a Jaccard coefficient of $0,95(\pm 0.04)$. Subsequently, to train our deep learning model with a small dataset of GGO and consolidation CT images, the model training is performed using an active learning approach. The results for semantic segmentation achieve a DSC of $0,81(\pm 0.12)$ and a Jaccard coefficient of $0,70(\pm 0.15)$ outperforming the supervised learning approach. Our results suggest that active learning reduced the amount of labeled data required for medical image segmentation without significant loss of accuracy.

Keywords: Consolidation, ground-glass opacity, lung infections, lung segmentation, deep learning, active learning

1. INTRODUCTION

Ground-glass opacities (GGO) are a common and highly significant finding on high-resolution computed tomography (HRCT) scans of parenchymal lungs.[1] GGO are characterized by areas of hazy increased attenuation of the lung,[2] not associated with obscuration of the underlying vessels or bronchial walls; if vessels are obscured, the term consolidation is preferred.[3] GGO are usually correlated with the presence of several lung diseases, such as alveolar collapse, interstitial thickening, or air-space disease.[2] Additionally, GGO can result from fibrotic processes,[4] and their growth rate can be a determinant measure to assess whether the pulmonary nodule is malignant, especially in cases of non-solitary nodule.[1] Furthermore, in the novel coronavirus disease (COVID-19), some studies have shown that GGOs increase in density and heterogeneity in the most severe cases, thus evolving in a consolidation pattern.[5,6] Sometimes, the areas of GGO may be admixed with areas of focal consolidation and can be confused, which makes segmenting these two findings independently a problematic task. However, their differentiation and measurement are crucial to quantifying the disease severity[7],[5] which will help physicians to manage better, treat, and prognosticate patients.[8]

Further author information: (Send correspondence to jorge.rudas@imexhs.com)

CT imaging has played an essential role in assessing parenchymal abnormalities in lung diseases such as chronic obstructive pulmonary disease (COPD) and, more recently, COVID-19.[9] CT-based quantification of lung disease usually requires annotation of the areas of infection consisting of a label map containing lesion types, in this case, GGO and consolidation. However, this manual segmentation task, performed by expert radiologists, is challenging, subjective, and high time-consumed due to the unclear boundaries of the infection area, the volume of images in a CT scan, and the influence of individual bias and clinical experience.[5] In addition, these segmentation masks are usually marked in just a few slices of the whole CT volume leading to have a reduced number of samples per patient which can be challenging for training deep learning models.

Thus, novel strategies to aid in the automatic segmentation of GGO and consolidation have played a relevant role during the last years. Recently, deep learning (DL) with convolutional neural networks (CNN) have shown promising results in several biomedical computer vision tasks, due to efficient feature extraction, especially for highlight or detect findings on CT images[910].[5] In particular, U-Net-based architecture has been developed for lung lesions segmentation.[11,12] Whole 3D volumes of CT chest were employed for training 3D deep learning-based approaches for classifying lung lesions.[10,13] Furthermore, unsupervised machine learning and deep reinforcement learning approaches to detect and quantify GGO in CT scans have been previously proposed.[8,14] All of these previous works aim to aid the radiologist's diagnostic process. Nonetheless, all these methods require the availability of large annotated data to train DL models, which can be challenging in the biomedical imaging domain.[15] To tackle this issue, our method aims to produce refined lesion segmentation with limited labeled data using an active learning approach.

Active learning is a paradigm that seeks to minimize the need for generating large labeled datasets through a query process that selects, in an iterative way, a subset of the most informative examples from an unlabeled images pool to be annotated.[1617]The underlying hypothesis is that this deep active learning approach could potentially achieve the same or higher accuracy using fewer training labels if it were allowed to choose the training data it wants to learn from.[18] In this work, we propose a two-stage deep learning model for GGO and lung consolidation semantic segmentation. As a first step, we extract our region of interest (ROI), the lung parenchyma, from each axial CT slice. Subsequently, from the extracted ROI, we develop a deep active learning framework for the segmentation task of GGO and consolidation.

This paper is organized as follows. In Section 2, we present our main contribution. In Section 3, we provide methodological details of this work. The experimental results and the discussion are presented in Section 4. Finally, concluding remarks are given in Section 5.

2. NOVEL CONTRIBUTIONS

Most of the deep learning approaches require the availability of large annotated data, which can be challenging in medical imaging. Herein, data annotation requires extensive expertise on the part of the radiologist, and it is a time-consuming and subjective task. To deal with this issue, we perform an active learning framework in a small Chest-CT, which aims to segment GGO and consolidation. Thus, our main contribution is to reduce annotation costs for highly skilled radiologists in the task of segmenting and detecting lung lesions such as GGO and consolidation using active learning. Our method has the potential to become an aid tool for radiologists and clinicians to annotate a large amount of CT images.

3. METHODOLOGY

3.1 Preprocessing

Typically, in clinical practice, radiologists and clinicians use some techniques to adjust pixel values that can modify the image appearance with the aim of highlighting certain structures in CT images. Therefore, as the first step of the workflow, as shown in Figure 1a, we propose an image preprocessing step to enhance the visualization of lung structures. The preprocessing step is divided into two parts. The first one corresponds to the image contrast and brightness settings. Further, the last step corresponds to the mapping of Hounsfield units (HU) to grayscale units.

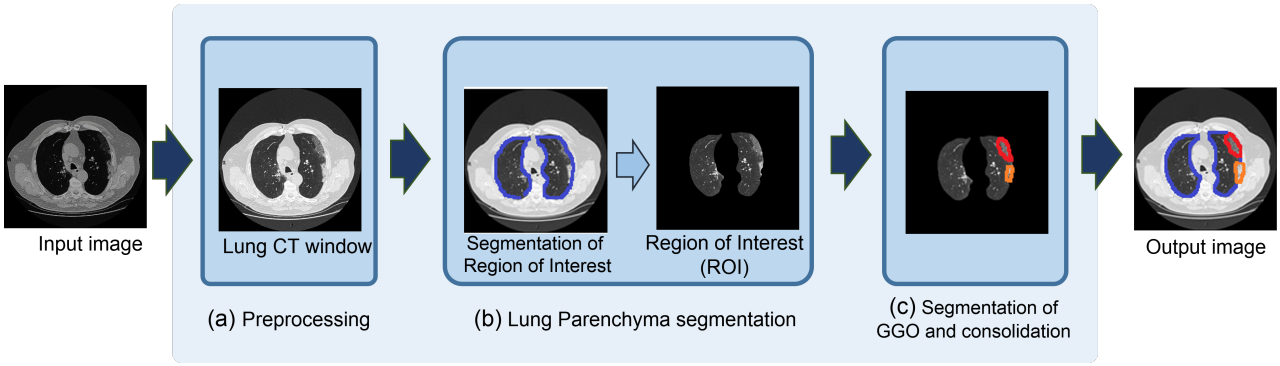

Figure 1. General workflow. As a first step, (a) the preprocessing step aims to enhance some lung structures on CT. Secondly, (b) with the intention of obtaining a region of interest (ROI), the segmentation of the lung parenchyma is performed. As the last step, (c) the identification and segmentation of the GGO and consolidation within the ROI are accomplished.

Contrast and brightness set-up

Windowing is the process in which the gray level of the CT image is mapped to modify the histogram by changing the image appearance to highlight particular structures. In other words, it is the process in which the gray image can be adjusted by manipulating the dynamic range through image processing to compress it into a limited number of grayscale values, maximizing the useful visual information that can be displayed on the image.[19] In CT images, the grayscale range is called the window width (WW), and the center of this range is the window level (WL). Both parameters can adjust brightness and contrast. In this work, aiming to highlight lung structures in CT, as most of the literature suggests, the WW and WL are set at 1500 and -500 respectively.[20]

Hounsfield units to gray-scale mapping

The absorption or attenuation coefficient of radiation in tissue in CT scans is measured in Hounsfield units (HU). Each pixel is quantitatively represented in terms of Hounsfield units (HU). This unit is proportional to the degree of x-ray attenuation by the tissue.[21] In general, The HU scale in CT images ranges between -1024 to $+3071$. By convention, high-density structures are assigned brighter colors while low-density tissues are assigned darker colors. To obtain a grayscale representation of the image, from 0 to 255, it is necessary to map all pixel values by a linear transformation of the data by using the following expression $G_L = S \times P_v + I_v$. Where G_L is the gray level output value, P_v is the current pixel value. From the DICOM header, the DICOM tag rescale intercept $(0028, 1052)$ and rescale slope $(0028, 1053)$ contains I_v and S values. As a final step in the preprocessing workflow, all images are normalized in the range 0 to 1 by dividing all pixel values by the largest pixel value.

3.2 Lung Segmentation

A critical process in identifying lung lesions is to extract the lung regions from the surrounding thoracic tissue successfully. Semantic segmentation provides information about lung parenchyma by labeling each image pixel with its corresponding class of what is being represented. Aiming to segment the lungs, as can be found in Figure 1b, we trained a U-Net-based model with an EfficientNetB0 architecture as a backbone. We train the model for 100 epochs (picking the best model) with a batch size of 8. Besides, we tune the Adam optimizer with a learning rate of 1×10^{-3}. The model output is a binary image of the same size as the input image, also known as a segmentation mask, in which each pixel is classified into a particular class (lung and background).

3.3 Consolidation and ground-glass opacities segmentation

After lung parenchyma segmentation, the last step in the methodology is GGO and consolidation segmentation as shown in the Figure 1c. In this work, we propose an active learning (AL) approach in order to train deep learning models for semantic segmentation with small labeled datasets. As a starting point, the whole image dataset is defined as a set $\mathcal{D} = \{(x_1, y_1), (x_2, y_2), \cdots, (x_n, y_n)\}$, where x is the CT image and its corresponding ground truth, y, that include GGO and consolidation annotations.

Subsequently, as shown in Figure 2a, we divide the dataset $\mathcal{D}$ into two sub-sets. The first one, defined as $\mathcal{T}$, is referred to as an initial training dataset that includes 20% of the dataset $\mathcal{D}$ whereas the remaining data corresponds to an unlabeled pool set, $\mathcal{U}$. Afterward, we train the U-Net model with an initial training data, $\mathcal{T}$, obtaining a set of segmented images, y.

Lastly, as shown in figure 2b, we feed the U-Net model with the data from $\mathcal{U}$ obtaining a set of segmented images, y. By using similarity metrics to compare ground truth and predicted segmentation, the set $\mathcal{U}$ is sorted from miss-segmented to well-segmented images. From the sorted set $\mathcal{U}$, we create a sub-set, query pool set ($\mathcal{Q}$) that includes the first sorted images from the set, which in our case, are the most informative images. From this set $\mathcal{Q}$, the label is corrected by a human expert (Figure 2c).

Nonetheless, in this work, we adopt a supervised learning approach to correct the labeled data for later adding these data to the initial training dataset, $\mathcal{T}$, as shown in Figure 2c. The size of the query pool set can be adjusted to a percentage of data from the set $\mathcal{U}$. In this work, the query pool size can vary from 5% to 20% of the set $\mathcal{U}$. Finally, we re-train the model and compute the evaluation metrics by performing an iterative process.

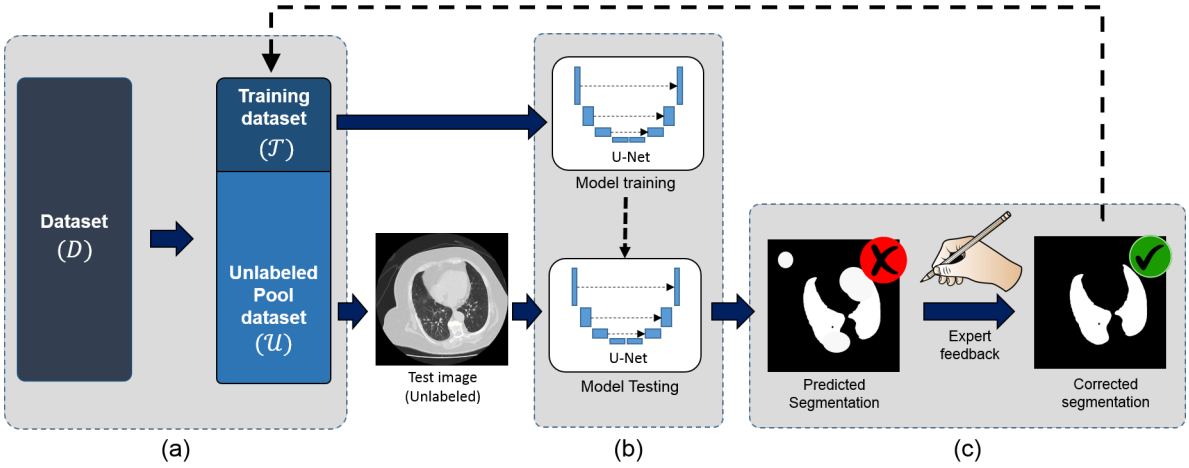

Figure 2. Active Learning approach. (a) Dataset partition for (b) model training and testing. Subsequently, (c) label correction to feedback corrected labels to training dataset

For the purposes of training the model, the numbers of epochs (picking the best model) and batch size were 100 and 8, respectively. Besides, we tunned the optimization algorithm, Adam, with an initial learning rate of 1×10^{-3}. In the last layer of the model, i.e., the output is a black and white image of the same size as the input image, in which each pixel is classified into a particular class. Considering that our segmentation is binary, consolidation and GGO label correspond to one (white area), whereas zero (black area) corresponds to the lung parenchyma.

4. EXPERIMENTS AND RESULTS

4.1 Dataset and Evaluation Criteria

Dataset

We conduct experiments in two different datasets. The first one is a private dataset called *Lungmexhs* which contains axial CT volumes from 45 patients. All the images were acquired in diagnostic imaging centers in Colombia. Ground truth annotations images were performed by expert radiologists of the Department of Diagnostics Medicine, RIMAB in Bogotá, Colombia. Each CT slice has a size of 512×512 pixels. We divided the dataset into 3 subsets; 30 patients for training ($\sim$ 3100 slices), 8 patients for validation ($\sim$ 2600 slices), and 7 patients for testing ($\sim$ 1860 slices). We perform a data augmentation for training with image rotation, flipping, and cropping. The *Lungmexhs* dataset was used only for the lung segmentation task.

On the other hand, for the GGO and consolidation segmentation task, we use the publicly available COVID-19 CT segmentation dataset from Radiopaedia.[22] This dataset includes 373 positives (including GGO, consolidation, and pleural effusion) axial CT slices from 9 volumes. Each sample, sized 512×512 pixels, has its corresponding segmentation mask as our ground truth evaluated by a radiologist. In this work, in order to ensure that we had CT slices with segmented and labeled areas, either GGO/Consolidation or both (on the same slide), 328 slices out of 373 were used. For our baseline, the supervised learning approach, the dataset has been randomly divided into three subsets; 201 slices for training, 61 for validation, and the remaining 66 for testing. Meanwhile, for the active learning approach, the dataset has been randomly divided into three subsets; 53 slices for training, 209 corresponds to the unlabeled pool set, and the remaining 66 for testing. Note that the test set for both methods is the same.

Evaluation criteria

In order to evaluate the segmentation performance, two similarity metrics are calculated to compare the predicted labels and the provided ground truth. The first one is the Intersection-Over-Union (IoU) index, and the second is the Dice similarity coefficient or commonly known as Dice score (DS).[23] The IoU, also known as the Jaccard Index, is the overlapped area between the predicted segmentation (set A) and the ground truth (set B) divided by the area of union between the predicted segmentation and the ground truth[24] as defined by the following expression:

$$IoU(A, B) = \frac{|A \cap B|}{|A \cup B|} = \frac{|A \cap B|}{|A| + |B| - |A \cap B|} \tag{1}$$

On the other hand, the Dice score (DS) is a spatial overlap index that measures the similarity between ground truth (set A) and prediction (set B) as defined as follows:

$$DS(A, B) = \frac{2\,|A \cap B|}{|A + B|} \tag{2}$$

In the expressions 1 and 2, A and B are the cardinalities of the two sets, i.e., the number of elements in each set. Both IoU and DS are in an interval from 0 to 1, where 0 indicates no spatial overlap between the ground truth and the predicted segmentation, and 1 indicates a complete overlap.

4.2 Results: Lung segmentation

In the first stage of our approach, lung segmentation, the U-Net model (EfficientNetB0 as a backbone) for distinguishing lung parenchyma yielded a training Dice score of $0,98$ and IoU coefficient of $0,96$. On the testing set ($n = 457$ Lung CT images), Dice score of $0,97(\pm 0.02)$ and IoU coefficient of $0,95(\pm 0.04)$. Based on a small image set containing a few hundreds of manually segmented lung images. The resulting image in the classification layer is a binary segmentation mask, in which each pixel is classified into a particular class as shown in Figure 3

4.3 Results: Consolidation and GGO segmentation

Initially, as a baseline, we performed a supervised learning approach to deep learning (training, validation, and testing) with the intention of comparing the performance of semantic segmentation using the active learning approach. The core of the experiments consists of training the U-Net model (with EfficientNetB0 as a backbone) for distinguishing GGO and consolidation from lung parenchyma. Initially, the first experiment consisted of train and testing the model with the whole slide CT (Baseline 1). Subsequently, another experiment consisted of training the model with a segmented ROI, i.e., segmented lung parenchyma (Baseline 2). The results in terms of IoU and Dice score are shown in Table 1.

	IoU	**DS**
Baseline 1	0,33 $(\pm 0,22)$	0,45 $(\pm 0,26)$
Baseline 2	0,61 $(\pm 0,22)$	0,73 $(\pm 0,23)$

Table 1. Supervised learning approach. Quantitative results for semantic segmentation

Subsequently, the following experiments consist of the active learning approach. Several experiments based on this approach consisted in varying the percentage of pool size, e.g., 5%, 10%, 15%, and 20% of the unlabeled pool. The metrics evaluated for semantic segmentation were Dice score (DS) and IoU index, as shown in Table 2. Qualitative results are shown in figure 4

Query pool size	IoU	DS
Exp. 1 - 5%	0,70 ($\pm 0,15$)	0,81 ($\pm 0,12$)
Exp. 2 - 10%	0,69 ($\pm 0,15$)	0,80 ($\pm 0,11$)
Exp. 3 - 15%	0,61 ($\pm 0,18$)	0,73 ($\pm 0,18$)
Exp. 4 - 20%	0,65 ($\pm 0,16$)	0,77 ($\pm 0,15$)

Table 2. Active learning approach. Quantitative results for semantic segmentation

5. DISCUSSION AND CONCLUSIONS

This work introduces an active learning approach for training deep learning models with small datasets aiming to perform semantic segmentation. As our experimental results show, we achieve promising results in the segmentation of the infection regions, such as GGO and consolidation regions. Herein, the IoU and dice score performance achieves the best results under the active learning framework compared with traditional supervised learning. These results could reduce the time and costs associated with annotating a large amount of CT images, which could become a valuable tool for the radiologist and other clinicians.

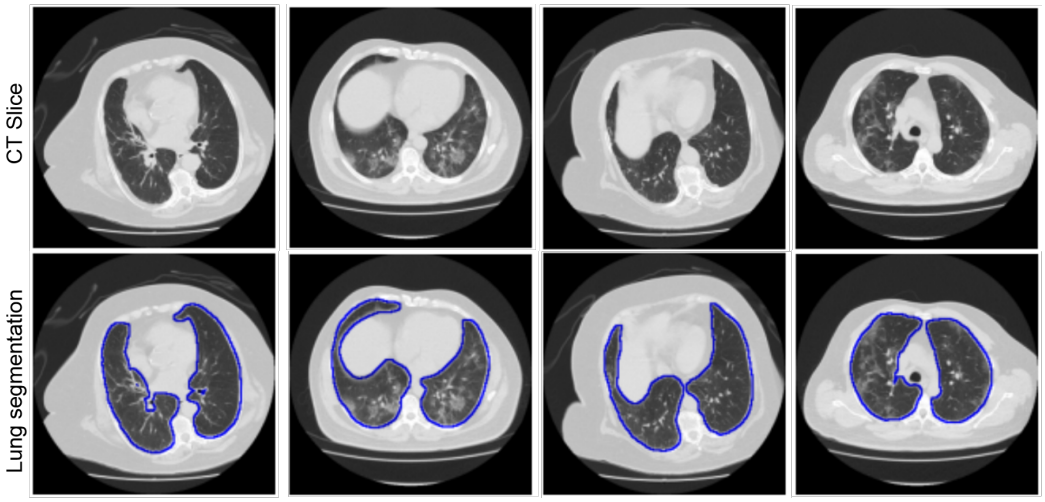

Figure 3. Lung segmentation in CT Chest slice. The first row corresponds to CT slice. In the second row, images correspond to lung segmentation overlapped in whole CT slice.

REFERENCES

[1] Zhu, Y., Tan, Y., Hua, Y., Zhang, G., and Zhang, J., "Automatic segmentation of ground-glass opacities in lung ct images by using markov random field-based algorithms," *Journal of digital imaging* **25**(3), 409–422 (2012).

[2] Battista, G., Sassi, C., Zompatori, M., Palmarini, D., and Canini, R., "Ground-glass opacity: interpretation of high resolution ct findings.," *La Radiologia Medica* **106**(5-6), 425–42 (2003).

[3] Cozzi, D., Cavigli, E., Moroni, C., Smorchkova, O., Zantonelli, G., Pradella, S., and Miele, V., "Ground-glass opacity (ggo): A review of the differential diagnosis in the era of covid-19," *Japanese journal of radiology* **39**(8), 721–732 (2021).

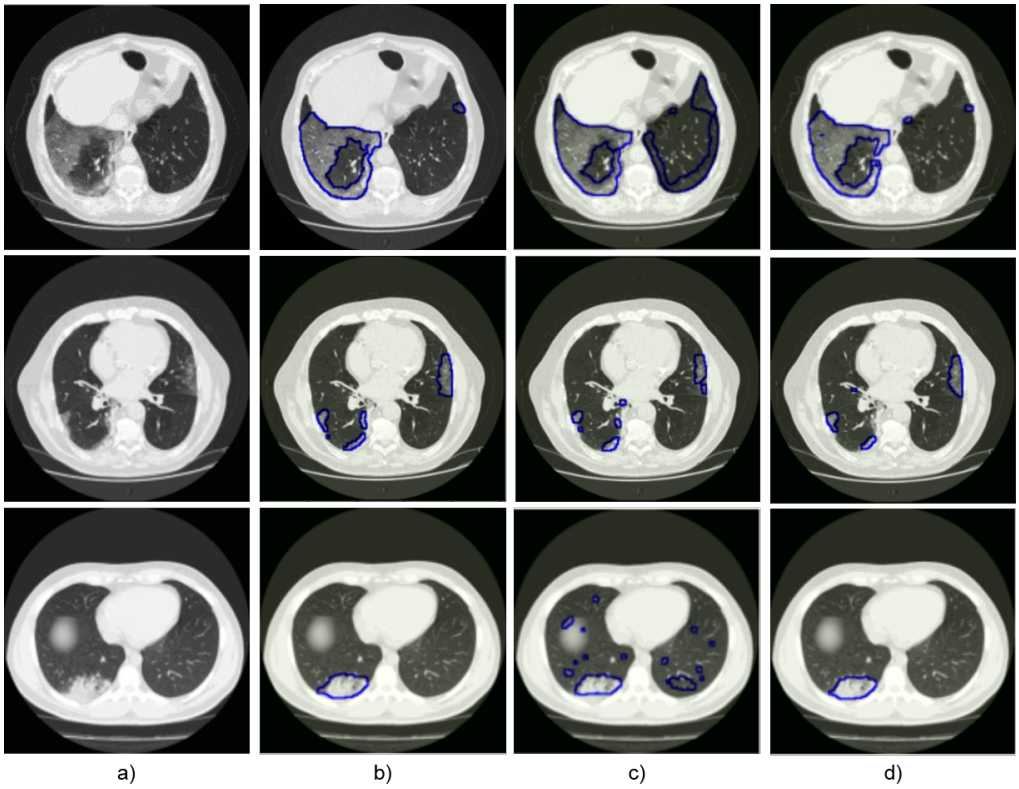

Figure 4. Qualitative results for GGO and consolidation segmentation. a) shows the original chest CT slice, b) shows the ground truth annotations. Segmentation results for the baseline e.g, supervised learning is shown in c) and the results for active learning approach is shown in d)

[4] Shimizu, K., Johkoh, T., Ikezoe, J., Ichikado, K., Arisawa, J., Nakamura, H., Tamura, S., and Nagareda, T., "Fractal analysis for classification of ground-glass opacity on high-resolution ct: an in vitro study," *Journal of computer assisted tomography* **21**(6), 955–962 (1997).

[5] Yao, H.-y., Wan, W.-g., and Li, X., "A deep adversarial model for segmentation-assisted covid-19 diagnosis using ct images," *EURASIP Journal on Advances in Signal Processing* **2022**(1), 1–17 (2022).

[6] Delli Pizzi, A., Chiarelli, A. M., Chiacchiaretta, P., Valdesi, C., Croce, P., Mastrodicasa, D., Villani, M., Trebeschi, S., Serafini, F. L., Rosa, C., et al., "Radiomics-based machine learning differentiates "ground-glass" opacities due to covid-19 from acute non-covid-19 lung disease," *Scientific Reports* **11**(1), 1–9 (2021).

[7] Mittal, B. and Oh, J., "Enhancement of covid-19 segmentation using machine learning analyses of lung imaging," in [*2021 6th International Conference on Biomedical Imaging, Signal Processing*], 24–30 (2021).

[8] Chaganti, S., Grenier, P., Balachandran, A., Chabin, G., Cohen, S., Flohr, T., Georgescu, B., Grbic, S., Liu, S., Mellot, F., et al., "Automated quantification of ct patterns associated with covid-19 from chest ct," *Radiology: Artificial Intelligence* **2**(4), e200048 (2020).

[9] Gerard, S. E., Herrmann, J., Xin, Y., Martin, K. T., Rezoagli, E., Ippolito, D., Bellani, G., Cereda, M., Guo, J., Hoffman, E. A., et al., "Ct image segmentation for inflamed and fibrotic lungs using a multi-resolution convolutional neural network," *Scientific reports* **11**(1), 1–12 (2021).

[10] Wang, D., Zhang, T., Li, M., Bueno, R., and Jayender, J., "3d deep learning based classification of pulmonary ground glass opacity nodules with automatic segmentation," *Computerized Medical Imaging and Graphics* **88**, 101814 (2021).

[11] Zhou, T., Canu, S., and Ruan, S., "Automatic covid-19 ct segmentation using u-net integrated spatial and channel attention mechanism," *International Journal of Imaging Systems and Technology* **31**(1), 16–27 (2021).

[12] Raj, A. N. J., Zhu, H., Khan, A., Zhuang, Z., Yang, Z., Mahesh, V. G., and Karthik, G., "Adid-unet—a segmentation model for covid-19 infection from lung ct scans," *PeerJ Computer Science* **7**, e349 (2021).

[13] Lu, W., Wei, J., Xu, T., Ding, M., Li, X., He, M., Chen, K., Yang, X., She, H., and Huang, B., "Quantitative ct for detecting covid-19 pneumonia in suspected cases," *BMC infectious diseases* **21**(1), 1–8 (2021).

[14] Kalia, R. K., Sharma, A., Amin, S. B., Saha, M., and Thittamaranahalli, S. K., "Ai-driven quantification of ground glass opacities in lungs of covid-19 patients using 3d computed tomography imaging," *medRxiv* (2021).

[15] Sharma, D., Shanis, Z., Reddy, C. K., Gerber, S., and Enquobahrie, A., "Active learning technique for multimodal brain tumor segmentation using limited labeled images," in [*Domain Adaptation and Representation Transfer and Medical Image Learning with Less Labels and Imperfect Data*], 148–156, Springer (2019).

[16] Gorriz, M., Carlier, A., Faure, E., and Giro-i Nieto, X., "Cost-effective active learning for melanoma segmentation," *arXiv preprint arXiv:1711.09168* (2017).

[17] Sharma, D., Shanis, Z., Reddy, C. K., Gerber, S., and Enquobahrie, A., "Active learning technique for multimodal brain tumor segmentation using limited labeled images," in [*Domain Adaptation and Representation Transfer and Medical Image Learning with Less Labels and Imperfect Data*], 148–156, Springer (2019).

[18] Settles, B., "Active learning literature survey," (2009).

[19] Gomori, J. M. and Steiner, I., "Non-linear ct windows," *Computerized radiology* **11**(1), 21–27 (1987).

[20] Whalen, B., "The slice is right (an exercise in ct windowing)," *Canadian Journal of Medical Radiation Technology* **34**(4), 5–10 (2003).

[21] DenOtter, T. D. and Schubert, J., "Hounsfield unit," (2019).

[22] Gaillard, F. et al., "Radiopaedia: building an online radiology resource," European Congress of Radiology-RANZCR ASM 2011 (2011).

[23] Eelbode, T., Bertels, J., Berman, M., Vandermeulen, D., Maes, F., Bisschops, R., and Blaschko, M. B., "Optimization for medical image segmentation: Theory and practice when evaluating with dice score or jaccard index," *IEEE Transactions on Medical Imaging* **39**(11), 3679–3690 (2020).

[24] Jaccard, P., "Étude comparative de la distribution florale dans une portion des alpes et des jura," *Bull Soc Vaudoise Sci Nat* **37**, 547–579 (1901).