Deep Learning Based Auto Lung Segmentation and Gross Tumor Volume Detection in NSCLC

Ravindra Patil∗1,2, patil.ravindra@philips.com
Allmin Pradhap Singh Susaiyah∗1, allmin.singh@philips.com
Geetha M∗1,2, geetha.m@philips.com
Andre Dekker∗1, andre.dekker@maastro.nl
Leonard Wee∗1, leonard.wee@maastro.nl
1 Philips Research, Bangalore, India 2 Department of Radiation Oncology, MAASTRO Clinic, Maastricht, The Netherlands

Editors: Under Review for MIDL 2019

Abstract

Segmentation of Lung is the vital first step in radiologic diagnosis of lung cancer. In this work, we present a deep learning based automated technique that overcomes various shortcomings of traditional lung segmentation and reuse code by transferring the model to perform gross tumor detection with little effort. Our approach resulted in 0.99 average dice score for lung segmentation including diseased condition without any seed point initialization and 0.85 dice score in segmentation of Gross Tumor Volume (GTV).

Keywords: Deep Learning, Chest CT, NSCLC, Lung Segmentation, Lung Tumor Detection, GTV

1. Introduction

Lung Cancer (LC) is by far the leading cause of death by any cancer. The American Cancer Society estimates the number of cancer deaths in US in 2019 to be 142,670. Being the second most common cancer in both genders, the chances of developing it is 1:15 and 1:17 for men and women respectively. With 228,150 new cases of lung cancer expected to be diagnosed in 2019 in US (Society, 2018) there is a great need for automating LC diagnosis for screening and treatment planning. The Non-small cell lung cancer (NSCLC) is the most common type of LC (Ohno Yoshiharu, 2018) and Chest CT has been the prime imaging technique to diagnose it. Segmentation of the lung is a vital first step in radiologic pulmonary image analysis. It is followed by detection of malignant findings and generation of clinical report by radiologists.

1.1. Traditional Approaches and Challenges

A number of techniques have been used for segmentation of lungs over the past decade. The techniques ranges from thresholding to more advanced such as neighbouring anatomy,

* Contributed equally
model based segmentation (Sluimer Ingrid, 2005). However, the various challenges involved in this task still makes it an open research field. Lung has a complex structure, and at a time, airways in CT scans are mistaken for lungs due to similar HU values. The lung can be dispersed due to physiological factors and does not have a rigid shape. The CT values of tumor mostly overlap with that of normal tissues and even the presence of calcification might be confused for tumor. For this reason, higher level image features are necessary to make reliable segmentation and diagnostics. In this work, we developed a deep learning (DL) based lung segmentation and gross tumor volume detection algorithm that will serve as a clinical decision support system for NSCLC diagnosis. Our work differs from existing approaches in following aspects

- Ability to generalize well on large dataset
- No need of seed point initialization
- Re-inclusion of nodules
- Intelligent enough to eliminate air ways
- Easy maintainability

2. Methodology

A well annotated public data set from the cancer imaging archive (TCIA) (www.cancerimagingarchive.net) (Aerts HJ, 2014) has been used in this study to train and validate our model. It contains 422 full Chest CT scans subjects with NSCLC, mapping to approx 42000 slices. Expert Radiologist annotation is available for the lung regions and gross tumor volumes (GTV). The training data has the following sub categories of NSCLC: Adenocarcinoma, Large cell carcinoma and Squamous cell carcinoma. The CT scanners used for imaging the subjects were from different vendors (Siemens, CNS inc, Philips, and GE) and the slice thickness was 3mm with 512 * 512 resolution.

2.1. Lung Segmentation Model Training

Our approach of lung segmentation is described in this section and the overall pipeline is shown in Figure 1. There are two phases in creation of the deep learning model, which aids in segmentation of the lung anatomy. First stage is the training phase in which, the data is preprocessed and fed to the model. The model is trained to segment the region of the lung using annotated ground truth mask as reference. Further, the learnt model layers are analyzed using visualization to ascertain, which region of the image, the model looked into, to arrive at the segmentation. The second stage is scoring phase in which the trained model is used to delineate the region of the lung on the unseen/live chest CT scans.

The model architecture that is built was inspired from U-Net Convolution Neural Network (CNN), (Ronneberger Olaf, 2015) which consists of 18 convolutional layers, 4 central pooling layers with the convolutional kernel size of 3X3 in each convolutional layer. The schematic representation of the model architecture is as shown in the Figure 2.

The first half of the U-Net act as a feature extractor that learn complex feature representations of the object of interest (here, the lung) and the second half of the U-Net act
DL Based Lung Segmentation and Tumor Detection

Figure 1: Block diagram representation of the model building process

Figure 2: U-Net architecture block diagram
as localizers that map features to locations in a 2D plane. The hyperparameters set for this architecture are Optimizer = Adam, Learning Rate=1.0e-6, Metric = Dice score, Number of Epochs = 50, Batch Size = 2, Weight Initialization Method = Xavier initialization

2.2. Transfer learning Tumor Segmentation Model

On comparison to lung segmentation, GTV segmentation is more challenging as the region of interest contributes to less than 10% of the pixels in the axial slice. However, we overcome this challenge by following a transfer learning approach where the tumor segmentation model is trained on top of the already trained lung segmentation model. This helps in achieving better performance as the model already has a knowledge about the image features generally present in chest CT during the start of training.

3. Results

Figure 3 shows the result of lung segmentation on an axial slice of chest CT scan. It is observed that the segmentation mask generated by the DL network closely resembles the ground truth mask annotated by the radiologist. Quantitatively, a Dice score of 0.99 is observed on these slice. Overall, the lung segmentation was trained on 10,000 slices and tested on 30,000 unseen slices and an average Dice Coefficient of 0.99 was obtained. Comparison with other lung segmentation approach is provided in Table 1. The approaches that are mentioned are either with seed point initialization or semi automatic in nature, also have added domain specific constraints on HU values and post processing to arrive at segmentation of lung region. Further, the data set that is used in other studies for training and validation is not exhaustive so to comment on the generalization of the approach in diseased condition such as NSCLC. Our approach doesn’t employ postprocessing or any domain based adaptation to arrive at the segmentation, further it is validated exhaustively on 30,000 slices of diseased as well as non diseased to arrive at the dice score.

A representative output of the tumor detection model is shown in Figure 4.

Table 1: Comparison with other Lung Segmentation Approaches.

<table>
<thead>
<tr>
<th>Study</th>
<th>Approach</th>
<th>Number of Dataset</th>
<th>Dice Coeff.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y.Wei et.al (Y.Wei, 2013)</td>
<td>Bresenham algorithm</td>
<td>97 subjects</td>
<td>0.95</td>
</tr>
<tr>
<td>Dai et.al (S.Dai, 2015)</td>
<td>Gaussian Mixture</td>
<td>Not mentioned</td>
<td>0.98</td>
</tr>
<tr>
<td>Noor et.al (N.M.Noor, 2015)</td>
<td>Thresholding and Morphological operation</td>
<td>96 subjects</td>
<td>0.98</td>
</tr>
<tr>
<td>Zhang et al (W.Zhang, 2015)</td>
<td>Active Contour based approach</td>
<td>60 subjects</td>
<td>0.97</td>
</tr>
<tr>
<td>Hasan et.al (Koyuncu, 2018)</td>
<td>Region growing and framework approach</td>
<td>170 slices</td>
<td>0.98</td>
</tr>
<tr>
<td>Our Approach</td>
<td>Deep Learning</td>
<td>422 Subjects</td>
<td>0.99</td>
</tr>
</tbody>
</table>

From Figure 5 it can be observed that the dice score for the tumor detection is 0.83 for tumors with area more than 900 mm². This transfer learning model was trained using 3991 tumor masks and validated across 997 unseen images. Scoring on a Nvidia GTX 1080 GPU took on an average, 15 seconds per subject.
Figure 3: Lung Segmentation on an Axial Slice of Chest CT

Figure 4: GTV Segmentation on an Axial Slice of Chest CT
4. Conclusion

The results show the robust nature of the DL algorithm to segment the lung and tumor regions in the diseased subjects. The same algorithm can also be extended to other anatomical structures such as liver and skull. The GTV segmentation was performed using a GTX 1080 GPU in 15 seconds per scan on an average. This is great improvement from manual delineation that generally takes 25 minutes per scan. This automated approach without manual involvement of radiologist leads to saving significant diagnosis/treatment time.

References

Leijenaar RTl Aerts HJ, Velazquez ER. Decoding tumour pheno-type by noninvasive imaging using a quantitative radiomics approach. 2014.


