Transfer Learning for Detecting Chronic Obstructive Pulmonary Disease via Computed Tomography Images

Author(s) names withheld

Address withheld

Editors: Under Review for MIDL 2020

Abstract

Chronic obstructive pulmonary disease (COPD) is characterized by irreversible obstruction of the airway tree and progressive damage to the lung parenchyma. A poor quality of patient life, calls for early detection of the disease. In this study, we propose a transfer learning approach for analyzing lung CT images to predict disease status in COPD. Instead of training a deep convolutional neural network from scratch, we use deep features from a pre-trained VGG-19 model. We compare our results to a current state-of-the-art method and demonstrate that a simpler, transfer learning approach can be used to detect COPD in lung CT images by achieving a mean area under the curve of $0.92 \pm 0.03$.

Keywords: Transfer learning, COPD, computed tomography, deep learning.

1. Introduction

Deep learning has shown great promise towards predicting clinically relevant patient outcomes in chronic obstructive pulmonary disease (COPD) (González et al., 2018). Most deep learning methods, however, require large datasets, long training times, and may be difficult to train on smaller datasets. We propose a faster, transfer learning approach for diagnosing COPD within the SubPopulations and InteRmediate Outcome Measures In COPD Study (SPIROMICS) cohort (Couper et al., 2014). Transfer learning offers an efficient way of utilizing accumulated domain knowledge for smaller clinical datasets. It gained widespread attention after successful completion of the ImageNet Large Scale Visual Recognition Challenge (Russakovsky et al., 2015). Several well-known architectures from this challenge have shown promise in identifying disease markers in radiological scans (Lao et al., 2017). We hypothesize that deep features computed from chest computed tomography (CT), using a pre-trained convolutional neural network (CNN), can be utilized effectively for diagnosing a subject with COPD. The proposed model could be utilized for early stage diagnosis of the disease, thereby allowing for timely therapeutic interventions.

2. Methods

A transfer learning approach is used for predicting COPD status from CT images. A well-known, pre-trained VGG-19 network (Simonyan and Zisserman, 2014) is used to extract a rich set of features from the CT images. These features are used as inputs to a linear SVM classifier for predicting the presence of COPD.
2.1. Datasets

In our study, CT scans from the SPIROMICS cohort were used for extracting deep features. SPIROMICS is a prospective cohort study aimed at understanding disease genesis, progression, and prognosis in COPD (Couper et al., 2014). So far, it has enrolled 2,981 participants that have been scanned at both the total lung capacity (TLC) and residual volume (RV). Utilizing that, a total of 1117 paired inspiratory and expiratory scans were used from 540 subjects at three time points: baseline, 1 year and 5 year (see Table 1). Spirometry was used to measure lung function and group subjects into two classes. COPD status was determined using the percentage post bronchodilator FEV1/FVC ratios, with subjects having FEV1/FVC < 70% were labeled to have COPD. From the total of 1,117 scans, 354 were determined to have COPD and the remainder (763) as non-COPD, which clearly highlights the class imbalance present within the data. A heterogeneous set of disease severities were utilized, as shown in Table 1. One COPD subject was missing the FEV1 measurement, resulting in an unknown GOLD stage. Before feature extraction, the images were isotropically resampled to a resolution of 1 mm³.

Table 1: Subject stratification across various GOLD stages, defined in (Vestbo et al., 2013)

<table>
<thead>
<tr>
<th></th>
<th>Baseline</th>
<th>Year - 1</th>
<th>Year - 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-smokers</td>
<td>42</td>
<td>40</td>
<td>21</td>
</tr>
<tr>
<td>GOLD0</td>
<td>307</td>
<td>236</td>
<td>117</td>
</tr>
<tr>
<td>GOLD1</td>
<td>2</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>GOLD2</td>
<td>41</td>
<td>45</td>
<td>27</td>
</tr>
<tr>
<td>GOLD3</td>
<td>87</td>
<td>39</td>
<td>19</td>
</tr>
<tr>
<td>GOLD4</td>
<td>47</td>
<td>35</td>
<td>7</td>
</tr>
<tr>
<td>Unknown</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>526</td>
<td>397</td>
<td>194</td>
</tr>
</tbody>
</table>

2.2. Deep Feature Extraction

For feature extraction, we used a fixed VGG-19 network (Simonyan and Zisserman, 2014), which was trained using the ImageNet dataset. The input dimensions of VGG-19 are $224 \times 224 \times 3$, where the last dimension corresponds to the number of channels. To match this input image configuration, we selected three coronal slices consisting of the one with the maximum lung cross-sectional area and the two adjacent slices. Slices were resampled to $224 \times 224$ and stacked to make up the three channel input to the VGG-19 network. We use the last two fully-connected layers ($f_1$ and $f_2$) to extract a total of $16,384 = 4,096$ (layer size) × 2 ($f_1$ and $f_2$) × 2 (TLC and RV), features. CT scans from the baseline, year 1 and year 5 visits were used for generating the complete feature set, which was then mean normalized.
2.3. COPD Classification

A linear support vector machine (SVM) was trained to classify subjects into COPD and non-COPD groups. After feature extraction, a low-variance filter was used to reduce the feature set from 16,384 to 1,366 features. To address the class imbalance, synthetic minority oversampling technique (SMOTE) (Chawla et al., 2002) was used to oversample the minority class (i.e., COPD). Lastly, the model was evaluated using 10-fold cross validation and the area under the receiver operating characteristic (AUC - ROC) curve served as the metric to evaluate performance of our model.

![ROC curve](image)

Figure 1: (left) Performance evaluation of linear SVM using 10-fold cross validation. ROC curve for over each fold, along with their mean is depicted. (right) A confusion matrix highlighting classifier’s performance over 10 - folds

3. Results and Discussion

The ROC curves for each of the folds and the overall average ROC curve, along with confusion matrix, are plotted in Figure 1. Area under the curve was 0.92±0.03 averaged over all 10 folds. Our model achieves a higher accuracy when compared to a recently developed deep learning model (AUC - ROC = 0.856) trained from scratch (González et al., 2018), although on a different, but similar, dataset.

Transfer learning reduces the need for large sets of data and time-consuming model training. In our experiments, we use three different coronal slices for extracting features, because they perceptually offer a larger cross-sectional area of the parenchyma. It would, however be interesting to see how the axial and sagittal slices contribute towards disease status prediction. Another unique aspect of our approach is that we use both the expiratory and inspiratory scans in our analysis, while (González et al., 2018) used only inspiration images. One limitation of using transfer learning is that we are limited to 2D image inputs only, thereby limiting our model to 2D lung texture features. A potential bias in our experiments may arise due to the assumption that multiple visits by a single subject can be treated independently and requires further evaluation. As another future direction, 2D CNNs could be augmented with 2.5D and 3D networks for improving model performance.
4. Acknowledgments

This work was supported in part by XXXXXXXX XXXXX XXXXX XXXXXXXX. SPIROMICS was supported by contracts from the NIH NHLBI (HHSN26820090013C, HHSN26820090014C, HHSN26820090015C, HHSN26820090016C, HHSN26820090017C, HHSN26820090018C, HHSN26820090019C, HHSN26820090020C), grants from the NIH NHLBI (U01 HL137880 and U24 HL141762), and supplemented by contributions made through the Foundation for the NIH and the COPD Foundation from AstraZeneca MedImmune; Bayer; Boehringer-Ingelheim Pharmaceuticals, Inc.; Chiesi Farmaceutici S.p.A.; Forest Research Institute, Inc.; GlaxoSmithKline; Grifols Therapeutics, Inc.; Ikaria, Inc.; Novartis Pharmaceuticals Corporation; Nycomed GmbH; ProterixBio; Regeneron Pharmaceuticals, Inc.; Sanofi; Sunovion; Takeda Pharmaceutical Company; and Theravance Biopharma and Mylan.

References


Detecting COPD using Transfer Learning

Detecting COPD using Transfer Learning