Deep Learning Model for COPD Classification/Staging Using Lung CT

H. Alsurayhi[1] and A. Samani[1,2,3]

[1]Department of Electrical and Computer Engineering, Western University, London, ON, Canada
[2]Departments of Medical Biophysics, Western University, London, ON, Canada
[3] School of Biomedical Engineering, Western University, London, ON, Canada

I. Introduction

Chronic Obstructive Lung Disease (COPD) is a progressive and prevalent lung disease which is associated with airflow obstruction due to inflamed airways and lung parenchyma. The Global Initiative for Chronic Obstructive Lung Disease (GOLD) utilizes a combined assessment of COPD using three factors of lung function, symptoms, and exacerbation history [1]. These are used to evaluate disease severity and prevent future risks through devising patient-specific treatment. This assessment approach led to an Eight COPD Staging System developed in the COPDGene study [2]. In this System, COPD subjects are classified into 8 classes of: A, B, C1, C2, C3, D1, D2, and D3 to provide critical information to guide therapy. In this study, we have developed a deep learning 3D-CNN model using patient's thoracic CT images instead of data pertaining to lung function to assess COPD severity based on the high-resolution 8-stage system.

II. Methods

The Convolutional Neural Network (CNN) is a deep-learning approach that automatically extracts features from image data. A 3D-CNN deep learning model for COPD classification/staging was developed in this investigation based on the recent Eight COPD Staging scheme. It was developed based on the lung CT scans at expiration phase in conjunction with two variables pertaining to COPD exacerbation and symptoms level. COPD exacerbation is the frequency of hospitalization in the prior year due to severe symptoms. The symptoms level is measured with a questionnaire (St. George's Respiratory Questionnaire). Several preprocessing steps were applied to the lung CT volume, including image resampling, cropping, and normalization. All image volumes were resampled to a size of 128×128×128 using cubic spline interpolation while the HU of each voxel was normalized within the range [0, 1].

The 3D-CNN model consists of eight 3D convolutional layers of kernel size 3 × 3 × 3. The feature volume from each convolutional layer was down-sampled by a max-pooling layer with a 2 × 2 × 2 voxels window. The resulting features from the convolutional layers were flattened and concatenated with the exacerbation and symptoms variables to perform the classification in five fully connected layers. Rectified linear unit (ReLU) activation function was used in each fully connected and convolutional layer. The convolutional network structure was optimized by an Adam optimizer in 107 epochs and batch size of 16 samples. Early stopping was used to avoid overfitting. The model was trained using 330 subjects from the NIH COPDGene study [3]. The number of cases used are 40, 42, 40, 40, 40, 42, 42, 45 for classes A, B, C1, C2, C3, D1, D2, and D3, respectively. The dataset was split into three parts of training set 60%, validation set 20%, and testing set 20%. A stratified split was used to preserve the class distribution in the training, validation, and test sets.

III. Results

An accuracy of 80% was achieved to classify the COPD subjects in the test set into the 8-stages. Most of the miss-classifications occur between classes that share two factors. The confusion matrix for the test set prediction is shown in Figure 1. A value of 0.84, 0.81, 0.80 were achieved for model precision, recall, and F1 score as a macro averaging for all classes. The ROC curve for each class is shown in Figure 2, which was calculated using One vs. Rest technique.

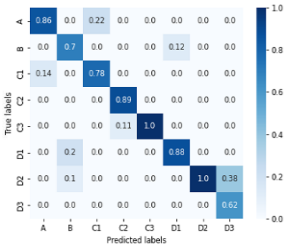

Figure 1: Confusion Matrix

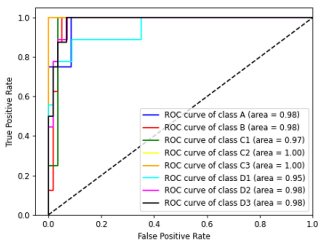

Figure 2: Receiver operating characteristic ROC curve

IV. Conclusions

A deep learning 3D-CNN model was developed using patient's thoracic CT images instead of data pertaining to lung function to assess COPD severity using the novel COPD 8-staging system. The system has demonstrated high accuracy in assessing COPD severity, providing valuable diagnostic information useful for treatment planning.

V. References

[1] "Global Initiative for Chronic Obstructive Lung Disease Global Initiative for Chronic Obstructive Lung Disease pocket guide to copd diagnosis, management, and prevention a guide for health professionals," 2017.

[2] M. L. K. Han *et al.*, "GOLD 2011 disease severity classification in COPDGene: A prospective cohort study," *Lancet Respir Med*, vol. 1, no. 1, pp. 43–50, Mar. 2013.

[3] E. A. Regan *et al.*, "Genetic epidemiology of COPD (COPDGene) study design," *COPD: Journal of Chronic Obstructive Pulmonary Disease*, vol. 7, no. 1, pp. 32–43, 2010.