2024 CMBEC46
Toronto, ON
May 28-30, 2024
CMBES/SCGB

Deep Learning for the Chronic Obstructive Pulmonary Disease Assessment Using Lung CT

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

Abstract— **COPD is a prevalent and progressive lung disease characterized by airflow limitation due to inflammation in the airways and/or alveoli. To manage disease progression and prevent exacerbations, GOLD recently proposed a staging system to guide treatment plans. This system relies on predictors assessed through comprehensive or simple questionnaires pertaining to disease history. In contrast to questionnaires, lung CT images offer quantitative assessment of disease severity. In this project, we developed a classification tool that uses deep learning to analyze lung CT images and classify the severity of two predictors of COPD: symptoms and exacerbation. We created two CNN models, Symptoms-CNN and Exacerbation-CNN, to identify the severity of these predictors. The Symptoms-CNN model achieved a classification accuracy of 0.7 for symptom severity, while the Exacerbation-CNN model achieved an accuracy of 0.75 for exacerbation severity. These models were integrated to design a novel classification tool based on the GOLD2023 staging system, which achieved a classification accuracy score of 0.72 using only a single 3D CT scan. These results indicate a reasonable level of accuracy, making this tool a promising approach to COPD diagnosis and management.**

Keywords— **COPD, Exacerbation, classification, CT, CNN.**

I. Introduction

Chronic Obstructive Pulmonary Disease (COPD) is a heterogeneous lung disease that is characterized by airflow limitation, which causes difficulty breathing. COPD has high morbidity and mortality worldwide. It was the third leading cause of death, causing 3.23 million deaths in 2019 [1]. Forced expiratory volume in one second (FEV1) and forced vital capacity (FVC) are often measured by pulmonary function tests (PFTs) to diagnose COPD. However, spirometry has a poor correlation with the degree of breathlessness or other COPD symptoms [2].

In 2023, the Global Initiative for Chronic Lung Disease guidelines (GOLD) committee proposed a new ABE staging system for COPD [3]. This system relies on the level of symptoms and COPD Exacerbation history to assess the disease severity (Fig. 1). The symptom level can be measured using a dyspnea measure (the modified Medical Research Council [mMRC] dyspnea score) or a health status measure (the COPD Assessment Test [CAT] score), or the SGRQ (St. George's Respiratory Questionnaire). COPD Exacerbation is

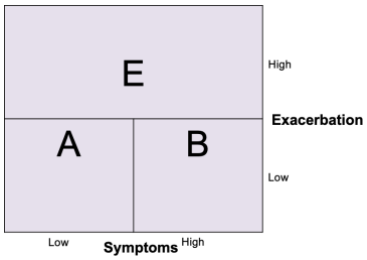

Fig. 1: GOLD2023 COPD staging system

defined as acute worsening of symptoms resulting in additional therapy; and it is measured as the frequency of requiring assessment in the Emergency Department or hospital admission in the prior year. Patients' symptoms and exacerbation history allow clinicians to initiate a treatment plan at the individual patient level. Lange et al. stated that symptoms and exacerbation history are key contributors to preventing or slowing down COPD progression as it has a better prediction for future COPD Exacerbation [4]. The risk of mortality increases with the frequency of severe COPD exacerbations, especially those requiring hospitalization [5]. Therefore, it is crucial to assess the patient's risk of experiencing an exacerbation.

The criterion currently used to assess the severity of exacerbation is not reliable. Not all patients who experience COPD exacerbation visit the hospital. The rate of hospitalization for COPD exacerbations was reduced by 53% compared to the pre-COVID pandemic. Avoiding visiting emergency rooms during COVID-19 possibly caused a change in the treatment of COPD exacerbations from inpatient to outpatient. Thus, other assessment tools are essential to detect the severity of COPD exacerbation to assess disease severity based on GOLD2023.

In this research project, we developed a classification technique based on the GOLD2023 staging system. To develop this model, we used lung CT images integrated with deep learning. Two Convolutional Neural Network (CNN) models have been developed. The first was trained to detect the symptoms level (Symptoms-CNN) while the other was trained to detect the severity of COPD exacerbation (Exacerbation-CNN). A classifier was then developed using

the two models to classify COPD patients based on the GOLD2023 staging system.

II. Previous works

Several studies have been conducted to assess COPD severity using thoracic CT. Moghadas et al. developed a classification model to assess the COPD severity based on the previous GOLD staging system [6]. Features used to build the model were extracted from lung CT scans pertaining to exhalation and inhalation; they were extracted based on lung air volume variation and distribution throughout the respiration cycle. Some other approaches used deep learning, where they developed a Convolutional Neural Network (CNN) to assess COPD severity using CT image data. In the model developed by Ho et al. paired 3D volumes were extracted from the 4D-CT data and used to create a 3D Parametric Response Map (PRM) [7] before a 3D-CNN model was trained. More recently, an unsupervised 3D convolutional autoencoder model was developed and integrated with a feature constructor as a classifier. An exploratory factor analysis was applied to explore the latent traits (factors) among pattern clusters. Two of the factors were used to train a logistic regression model to predict the severity of COPD exacerbation [9].

III. Methods

Dataset:

The dataset used in this project is available through the NIH COPDGene (COPD Genetic Epidemiology) study [10]. It is one of the largest studies dedicated to investigating COPD, comprising volumetric lung CT data, PFT measurements of subjects, exacerbation frequency, and symptom levels of 10,192 participants. For this project, we focused on thoracic CT scans taken during the expiration phase. A total of 562 subjects were utilized to develop the Symptoms-CNN model. Out of these, 284 belonged to the high symptom class and the remaining 278 cases belonged to the low symptom class. Similarly, the Exacerbation-CNN model was developed using 468 subjects. In this study, 234 cases represented the high exacerbation class while the other 234 cases belonged to the low exacerbation class.

Thoracic CT Preprocessing

Several preprocessing steps were applied to each lung CT volume in the dataset. First, each lung volume was segmented using a fully automated image segmentation method [11]. The segmented lung volume was used to apply volume cropping and remove the background area. Volumes were resampled to a size of 128×128×128 using cubic spline interpolation. Lastly, the Hounsfield unit (HU) of each voxel was normalized to values in the range of [0, 1].

A. Symptoms-CNN

The level of symptoms in the COPDGene study was assessed using two questionnaires. The mMRC questionnaire and the SGRQ questionnaire; the former is a simple questionnaire with 5 items while the latter is comprehensive with many questions. To avoid errors in the assessment, we used the following criteria: If both questionnaires indicated a high level of symptoms, we considered the symptoms to be high. If both questionnaires indicated a low level of symptoms, then we considered the symptoms to be low. However, if one questionnaire indicated a high level of symptoms while the other questionnaire indicated a low level of symptoms, we considered other data such as increased medication in the prior year and the frequency of using antibiotics or steroids at home to assess the level of symptoms accurately. *Symptoms-CNN Model Structure and Parameters:* The symptoms Convolutional Neural Network (CNN) used in this study begins with a 3D convolutional layer (3D_CONV) consisting of 8 filters and a kernel size of 7×7×7. This is followed by 4 Residual Blocks, each of which contains two 3D_CONV layers with a kernel size of 3×3×3, two ReLU activations, and one addition. The numbers of filters in these blocks are 16, 16, 32, and 32, respectively. The features obtained from these blocks are flattened and then passed to 3 fully connected layers, which include the output layer. The total number of trainable parameters in the model is approximately 2 million, and it was trained using the Adam optimizer with an initial learning rate of 0.001 for a total of 89 epochs. A batch size of 16 was used during training. The model's hyperparameters were assigned using the Hyperband optimization method, which includes activation function, learning rate, number of filters, and neurons in fully connected layers. The dataset was split into 75% for Training (422 CT scans) and 25% for Validation and Testing (140 CT scans) to train the model. The model was trained to classify subjects with COPD based on the level of symptoms, which were categorized as high or low.

B. Exacerbation-CNN

The severity of COPD exacerbation is determined by the number of exacerbations a patient experiences in a year. If the patient has been admitted to the hospital due to at least one exacerbation or has visited the emergency room due to

more than two exacerbations, it is considered a severe exacerbation. Otherwise, it is considered a low exacerbation.

Exacerbation-CNN Model Structure and Parameters: The lung CT scan is first passed through a 3D_CONV layer with a size of 7×7×7 and 16 filters. Then, 4 Residual Blocks are utilized with 16, 16, 32, and 128 filters respectively. The resulting features are flattened and fed into three fully connected layers, including the output layer, for performing the classification. The model was trained with the Adam optimizer with an initial learning rate of 0.001 and has approximately 8 million trainable parameters. Hyperband optimization was used to determine the model's hyperparameters, such as the activation function, learning rate, number of filters, and neurons in each fully connected layer. For training the model, 356 CT scans were used, while the other 112 samples were used for validation and testing. The model was trained for binary classification of high or low risk of future exacerbation.

C. GOLD2023 classification

GOLD2023 staging System is based on the level of symptoms the patient experiences and the risk of future exacerbation. To classify COPD patients based on this staging system, we employ two CNN models: Symptoms-CNN and Exacerbation-CNN. The classification process begins by feeding the lung CT scan to the Exacerbation-CNN model. If the output of the Exacerbation-CNN model is positive (indicating high risk), the patient is classified as GOLD stage E. If the output is negative (low risk), the lung CT scan is passed to the Symptoms-CNN model. If the Symptoms-CNN model output is positive, the patient is categorized as GOLD stage B. If the output is negative, the patient is classified as GOLD stage A. The diagram of this classification scheme is shown in Fig. 2.

IV. RESULTS

D. Symptoms-CNN model evaluation

The accuracy score is used to evaluate the Symptoms-CNN model classification performance. The model achieved a classification accuracy of 0.8, 0.71, and 0.7 in training, validation, and test sets, respectively. The Confusion Matrix and the Receiver Operating Characteristic (ROC) were calculated for the test set. The test set has 70 samples, where 35 samples belong to each of the high and low symptom classes. The normalized Confusion Matrix is shown in Fig. **3**; it shows that 0.74 of the samples with high symptoms are predicted correctly, and 0.69 of the low symptoms samples are also predicted correctly as low symptoms samples. The ROC curve with an area under the curve (AUC) of 0.73 was achieved in the Symptoms-CNN model (

Fig. 4).

E. Exacerbation-CNN model evaluation

The Exacerbation-CNN model was evaluated in terms of accuracy score, Confusion Matrix and ROC curve. The model's classification accuracy scores in training, validation and testing sets are 0.84, 0.76, and 0.75, respectively. The normalized Confusion Matrix for the test set is shown in Fig. 5, where the number of samples in the test sets is 56 samples (28 low risk, 28 high risk).

Fig. 6 shows the ROC curve for the Exacerbation-CNN model. The model achieved an AUC of 0.78 in the test set.

F. GOLD2023 classification evaluation

In order to assess the classification performance of the algorithm used for GOLD2023, a total of 75 samples were chosen. These samples were equally divided into three groups, with 25 samples representing each stage of GOLD2023. The accuracy of the model was measured using a classification accuracy score, which was found to be 0.72.

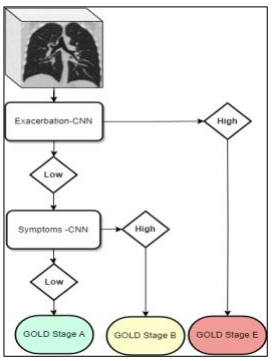

Fig. 2: GOLD2023 classification scheme

V. DISCUSSION

Our project involved applying deep learning techniques to lung CT images to improve the assessment of COPD. We utilized a single lung 3D CT scan to predict the two primary indicators currently used for treatment planning. Our Exacerbation-CNN model outperformed the Symptoms-CNN model in terms of classification accuracy. This could

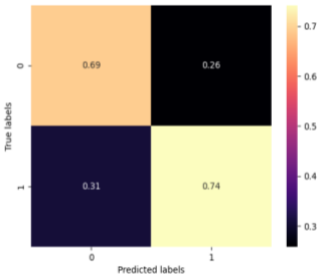

Fig. 3: Symptoms-CNN Confusion Matrix

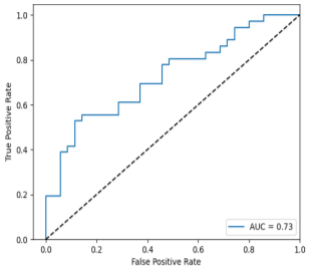

Fig. 4: Symptoms-CNN ROC curve

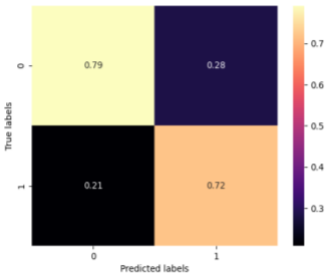

Fig. 5: Exacerbation-CNN Confusion Matrix

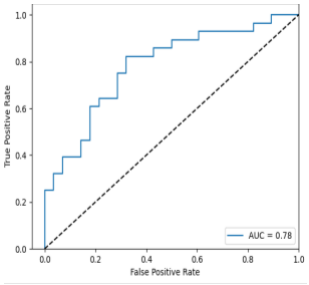

Fig. 6: Exacerbation-CNN ROC curve

be because exacerbation is caused by severe lung damage, which can be more effectively captured through lung CT imaging.

Among the previous different studies that used lung CT to assess COPD severity, the 3D-CNN trained with the PRM maps obtained from the paired lung scans to classify subjects to COPD and non-COPD [7] achieved a classification accuracy of 0.89. The Lung air model developed by Moghadas [6] classifies COPD subjects based on a previous GOLD staging system designed using the severity of PFT. That model achieved a classification accuracy of 0.84. While both previous methods achieved higher accuracy, it should be noted that they use paired 3D CT images pertaining to inhalation and exhalation phases in contrast to the proposed technique that uses a single 3D CT scan which is more likely to be available. The unsupervised model of Li et al. [9] achieved a classification accuracy of 0.66, while our exacerbation model achieved a higher classification accuracy of 0.75. Identifying the severity of COPD exacerbation from images is crucial as not all patients who experience exacerbation visit the hospital, while the tool used to measure COPD exacerbation is based on the history of visiting the hospital.

The recent GOLD 2023 guidelines recommend classifying patients with COPD by the severity of their symptoms and exacerbations before recommending proper medications following these measurements. The recommendations aim at minimizing disease progression. To our knowledge, the proposed classification algorithm is the first method used to perform GOLD2023 COPD staging using a single 3D CT scan which combines the two CNN models for predicting symptoms and exacerbation severity. The developed algorithm may be considered a valuable tool for identifying COPD severity based on image data before devising a therapy plan.

VI. Compliance with Ethical Requirements

All participants in THE COPDGene study provided written informed consent. This research protocol was approved by the institutional ethics and review board at each participating institution.

VII. Conclusions

A classification tool for COPD assessment based on the GOLD2023 was developed using 3D lung CT data and deep learning. Two CNN models, Symptoms-CNN and Exacerbation-CNN were trained and used to develop this classification method. This classifier has the potential to be used in the clinic for effective COPD assessment using image data.

VIII. Conflict of Interest

The authors declare that they have no conflict of interest.

IX. References

[1] World Health Organization WHO, “Chronic obstructive pulmonary disease (COPD).”

[2] A. Agusti *et al.*, “Characterisation of COPD heterogeneity in the ECLIPSE cohort,” *Respir Res*, vol. 11, Sep. 2010, doi: 10.1186/1465-9921-11-122.

[3] “global initiative for chronic obstructive lung disease global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease (2023 report),” 2022. [Online].

[4] P. Lange *et al.*, “Prediction of the clinical course of chronic obstructive pulmonary disease, using the new GOLD classification: A study of the general population,” *Am J Respir Crit Care Med*, vol. 186, no. 10, pp. 975–981, Nov. 2012, doi: 10.1164/rccm.201207-1299OC.

[5] J. J. Soler-Cataluña, M. Á. Martínez-García, P. Román Sánchez, E. Salcedo, M. Navarro, and R. Ochando, “Severe acute exacerbations and mortality in patients with chronic obstructive pulmonary disease,” *Thorax*, vol. 60, no. 11, pp. 925–931, Nov. 2005, doi: 10.1136/thx.2005.040527.

[6] H. Moghadas-Dastjerdi, M. Ahmadzadeh, E. Karami, M. Karami, and A. Samani, “Lung CT image based automatic technique for COPD GOLD stage assessment,” *Expert Syst Appl*, vol. 85, pp. 194–203, Nov. 2017, doi: 10.1016/j.eswa.2017.05.036.

[7] T. T. Ho *et al.*, "A 3D-CNN model with CT-based parametric response mapping for classifying COPD subjects," *Sci Rep*, vol. 11, no. 1, Dec. 2021, doi: 10.1038/s41598-020-79336-5.

[9] F. Li *et al.*, "Latent traits of lung tissue patterns in former smokers derived by dual channel deep learning in computed tomography images," *Sci Rep*, vol. 11, no. 1, Dec. 2021, doi: 10.1038/s41598-021-84547-5.

[10] 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, doi: 10.3109/15412550903499522.

[11] S. M. Ryan, B. Vestal, L. A. Maier, N. E. Carlson, and J. Muschelli, "Template Creation for High-Resolution Computed Tomography Scans of the Lung in R Software," *Acad Radiol*, vol. 27, no. 8, pp. e204–e215, Aug. 2020, doi: 10.1016/j.acra.2019.10.030.