Journal of Associated Medical Sciences 2024; 57(2): 149-156

JAMS
Journal of Associated Medical Sciences

Scopus Indexed & Thai-Journal Citation Index Centre (TCI)

Journal of Associated Medical Sciences

Journal homepage: https://www.tci-thaijo.org/index.php/bulletinAMS/index

AI-based diagnosis of chronic obstructive pulmonary disease from low-dose CT images

Chayanon Pamarapa[1] Salisa Kemlek[1] Wichasa Sukumwattana[1] Pharinda Sitthikul[1] Sichon Khuanrubsuan[1] Akkarawat Chaikhampa[1] Paritt Wongtrakool[2] Ammarut Chuajak[3] Monchai Phonlakrai[1] Ruedeerat Keerativittayayut[1*]

[1]School of Radiological Technology, Faculty of Health Science Technology, Chulabhorn Royal Academy, Bangkok, Thailand.
[2]Department of Diagnostic and Therapeutic Radiology, Faculty of Medicine, Ramathibodi Hospital, Mahidol University, Bangkok, Thailand.
[3]Queen Savang Vadhana Memorial Hospital, Chonburi Province, Thailand.

ARTICLE INFO

Article history:
Received 21 November 2023
Accepted as revised 9 April 2024
Available online 10 April 2024

Keywords:
Chronic obstructive pulmonary disease, Low dose computed tomography, ResNet, convolutional neural network.

ABSTRACT

Background: Chronic obstructive pulmonary disease (COPD) is a group of diseases characterized by airflow blockage. It is one of the leading causes of global mortality and is primarily attributed to smoking. COPD patients are usually diagnosed by spirometry test. Although regarded as the gold standard for COPD diagnosis, the spirometry test carries contraindications, thus prompting the development of low-dose computed tomography (low-dose CT) scan as an alternative for COPD screening. However, a practical limitation of diagnosing COPD from CT images is its reliance on the expertise of a skilled radiologist.

Objective: To address this limitation, we aimed to develop a deep-learning model for the automated classification of COPD and non-COPD from low-dose CT images.

Materials and methods: We examined the potential of a convolutional neural network for identifying COPD. Our dataset consisted of 10,000 low-dose CT images obtained from a lung cancer screening program, involving both ex-smokers and current smokers deemed at high risk of lung cancer. Spirometry data served as the ground truth for defining COPD. We used 90% of the datasets for training and 10% for testing.

Results: Our developed model achieved notable performance metrics: an area under the receiver operating characteristic curve (AUC) of 0.97, an accuracy of 0.89, a precision of 0.85, a recall of 0.96, and an F1-score of 0.90.

Conclusion: Our study demonstrates the potential of deep learning models to augment clinical assessments and improve the diagnosis of COPD, thereby enhancing diagnostic accuracy and efficiency. The findings suggest the feasibility of integrating this technology into routine lung cancer screening programs for COPD detection.

Introduction

Chronic obstructive pulmonary disease (COPD) is a term used to denote a group of diseases that result in the obstruction of airflow and consequent respiratory difficulties.[1,2] In 2024, the Global Initiative for Chronic Obstructive Lung Disease (GOLD) highlighted COPD as a leading cause of mortality globally, ranking among the top three.[3] In the large majority of COPD patients, this pathology is associated with smoking.[4] The absence of symptoms does not represent a dependable marker of disease and may instead delay diagnosis until more pronounced airflow obstruction becomes manifest.[4] Accurate detection of COPD is vital for promptly introducing therapies that reduce the risk of future exacerbation, delay disease

* **Corresponding contributor.**
Author's Address: *School of Radiological Technology, Faculty of Health Science Technology, Chulabhorn Royal Academy, Bangkok, Thailand.*
E-mail address: ruedeerat.kee@cra.ac.th
doi: 10.12982/JAMS.2024.037
E-ISSN: 2539-6056

progression, and enhance patient prognosis. However, a large amount of literature indicates that a significant proportion of patients are incidentally diagnosed with COPD during lung cancer screening, resulting in delayed treatment.[4]

Spirometry represents a widely accepted pulmonary function test and the gold standard for evaluating bronchial obstruction and lung elasticity.[5] However, spirometry presents certain limitations in patients with specific conditions such as aneurysms, respiratory tract infections, untreated pneumothorax, unstable vascular or cardiovascular systems, and those who have undergone thoracic or abdominal surgery. Additionally, spirometry may not be appropriate for patients with untreated or poorly controlled high blood pressure, low blood pressure, recent myocardial infarction, pulmonary embolism, or hemoptysis.[6,7] Therefore, physicians frequently rely on low-dose computed tomography (low-dose CT) as an alternative diagnostic technique, particularly in cases where spirometry may not be advisable.[8,9] Low-dose CT scanning is attractive because it reduces radiation exposure by nearly fourfold compared with conventional chest computed tomography,[10] however, COPD diagnosis through this technology demands the expertise of a skilled radiologist, thus putting pressure on clinical resources.

The objective of this study was to develop and evaluate a residual network-based artificial intelligence model that could effectively screen individuals for COPD from low-dose CT images. We hypothesized that only images from a specific portion of the upper lobe may carry reliable information for detecting COPD. Without human guidance, deep neural networks can identify patterns at various spatial scales that could distinguish between COPD and non-COPD.

Materials and methods

Participants

We retrospectively obtained low-dose CT images and spirometry data from individuals enrolled in a lung cancer screening program at Chulabhorn Hospital, Thailand. The study encompassed 295 participants who underwent follow-up assessments over 8 years from 2012 to 2020. Our analysis focused on a subset of this cohort, yielding 1,000 datasets comprising 500 chronic obstructive pulmonary disease (COPD) and 500 instances of non-COPD. Participants in the lung cancer screening initiative had to meet specific criteria: age exceeding 40 years, current smokers or individuals who had refrained from smoking for less than 15 years, and maintaining a minimum consumption of ten packs of cigarettes per year. The quantification of pack-years was based on the average daily packs smoked multiplied by the number of years smoked, with one pack equivalent to 20 cigarettes.

Additionally, prospective participants were required to have no prior history of lung disease or cancer. They were excluded if they self-reported history of lung cancer or congenital lung diseases, had lung surgery, used oral steroids, had blood transfer or dyspnea that required hospitalization for 4 weeks before the study, were unable to walk, or carried metallic implants in their thoracic region. Before engagement in the lung cancer screening program, all participants provided informed consent. Approval for the lung cancer screening program and this retrospective study was obtained from the institutional review board of Chulabhorn Royal Academy, Thailand, and complied with the Declaration of Helsinki.

Spirometry test

Spirometry is the gold standard for testing pulmonary or breathing functions during the identification and classification of COPD severity.[11] Participants were asked to perform the spirometry test before low-dose CT examination. COPD evaluation can be conducted using forced expiratory volume within one second (FEV1), which measures the air expelled in the first moments of a rapid exhalation, and forced vital capacity (FVC), which represents the total amount of air exhaled during the FEV test. Spirometry information served as the ground truth for this study. COPD can be defined using post-bronchodilator spirometry (FEV1/FVC ratio of less than 70%). COPD severity can be graded into four stages based on GOLD criteria: stage I (mild) for FEV1 values ≥80%, stage II (moderate) for FEV1 values between 50% and 79%, stage III (severe) for FEV1 values between 30% and 49%, and stage IV (very severe) for FEV1 values <30%.[4] Based on these criteria, 500 datasets from 164 participants were labeled as COPD, while 500 datasets from 131 regular participants were labeled as non-COPD. In the COPD cohort, 59%, 35%, 5%, and 1% of participants were categorized into GOLD stages 1, 2, 3, and 4, respectively.

Low-dose CT scans

Low-dose CT data were acquired using 192-detector row CT (SOMATOM Force, Erlangen, Germany) following the chest low-dose CT Protocol at Chulabhorn Hospital, Thailand, as outlined in Table 1. Scanning was performed in a head-first supine position. Participants were asked to raise their arms above their heads, fully inspire, and suspend breathing during scanning.

Table 1 Low-dose CT scanning parameters.

Scanning Parameters	Details
Surview image coverage	Cover Mid Neck to Anterior Superior Iliac Spine (ASIS)
Chest low-dose image coverage	Cover Supraclavicular to Adrenal glands
Reconstruction	Lung Window
Matrix size	512×512 pixels
Detector width	192 mm × 0.625 mm
Slice thickness	0.6-1.0 mm.
kV Mode	120 kV
mA Mode	25-50 mA
Pitch	1
Radiation dose ($CTDI_{vol}$)	≤1.9 mGy

Note: mA: milliampere; kV: kilovoltage, $CTDI_{vol}$: Computed Tomography dose index volume

Image pre-processing

Lung segmentation

All low-dose CT image volumes with lung window underwent the same processing pipeline, which consisted of approximate lung mask generation and extraction of regions of interest. Lung volumes were segmented from other organs and tissue in the thoracic region using the U-net model developed by Ronneberger and collaborators.[12] The U-net model is a convolutional U-shaped symmetrical neural network that won first place in the 2015 ISBI cell monitoring competition and is currently the most popular model for segmenting chest CT images.[13] Lung segmentation involved the following steps: 1) fully automated contouring of lung volume using the U-net model, 2) creation of a binary mask of lung volume, and 3) application of the acquired binary mask to the original CT images to isolate pixels associated with lung volume from those associated with other organs.

The examples of a low-dose chest CT input image are shown in Figure 1: the raw image before segmentation (a), the associated binary mask image (b), and the resulting segmented lungs (c). Initially, the U-net model classifies each pixel of the input image, resulting in a mask of lung regions to be extracted. As shown in Figure 1b, this mask assigns a value of 1 to pixels associated with the lung volume (white), and 0 to the remaining areas. Information within the segmented lung volume is utilized for model training in the next step.

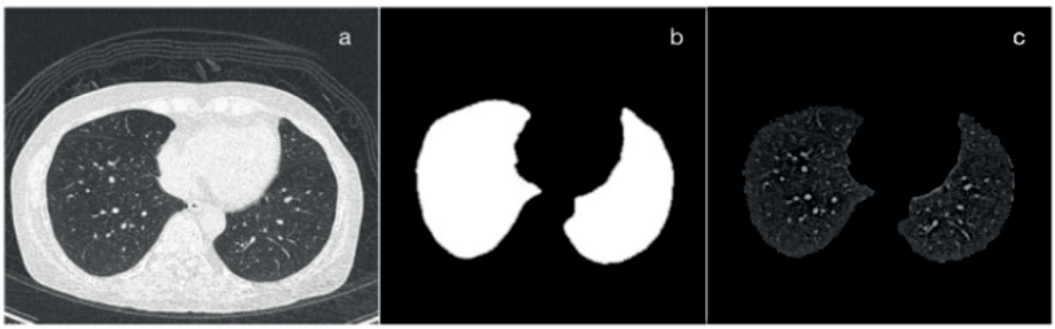

Figure 1 *Example of lung volume segmentation: (a) original image, (b) lung mask derived from the input image, (c) lung volume after segmentation.*

Slice selection

Previous studies identified the upper lobe of the lungs as the region with the highest likelihood of COPD presence.[14,15] Based on this evidence, we focused only on this region to optimize our computational resources. Slice selection was performed using in-house Python software. The first slice of segmented lung volume is determined by the presence of less than 0.5% black pixels of the selected slice. With a slice thickness of 1.0 mm and a slice interval of 0.7 mm, we uniformly sectioned slice numbers 52 to 61 in the axial direction (at a distance of 37.0-43.4 mm from the apex of the lung) based on the average lung dimensions of Thai individuals (240 mm).[16] These slices served as the center of the upper lobe. After slice selection, we obtained ten CT images for each dataset.

Following the above procedures, we obtained 10,000 input images (5,000 COPD and 5,000 non-COPD) for model training and testing. We used 90% of this dataset (9,000 images: 4,500 COPD and 4,500 non-COPD) for training and 10% (1,000 images: 500 COPD and 500 non-COPD) for testing.

The images, which initially measured 512×512 pixels, were resized to 224×224 pixels to comply with the input requirements of the model.

Model development

ResNet-50 architecture

This study adopted a pretrained ResNet-50 model, a convolutional neural network trained on the ImageNet dataset. ResNet-50 is employed as the default model to minimize the necessary input datasets and accelerate the construction of a new model for efficient classification.[17] We utilized the learned weighting factors from pre-training on ImageNet, where these factors dynamically adjust during model training, capturing the relationship between input and output data. The proposed architecture is illustrated in Figure 2.

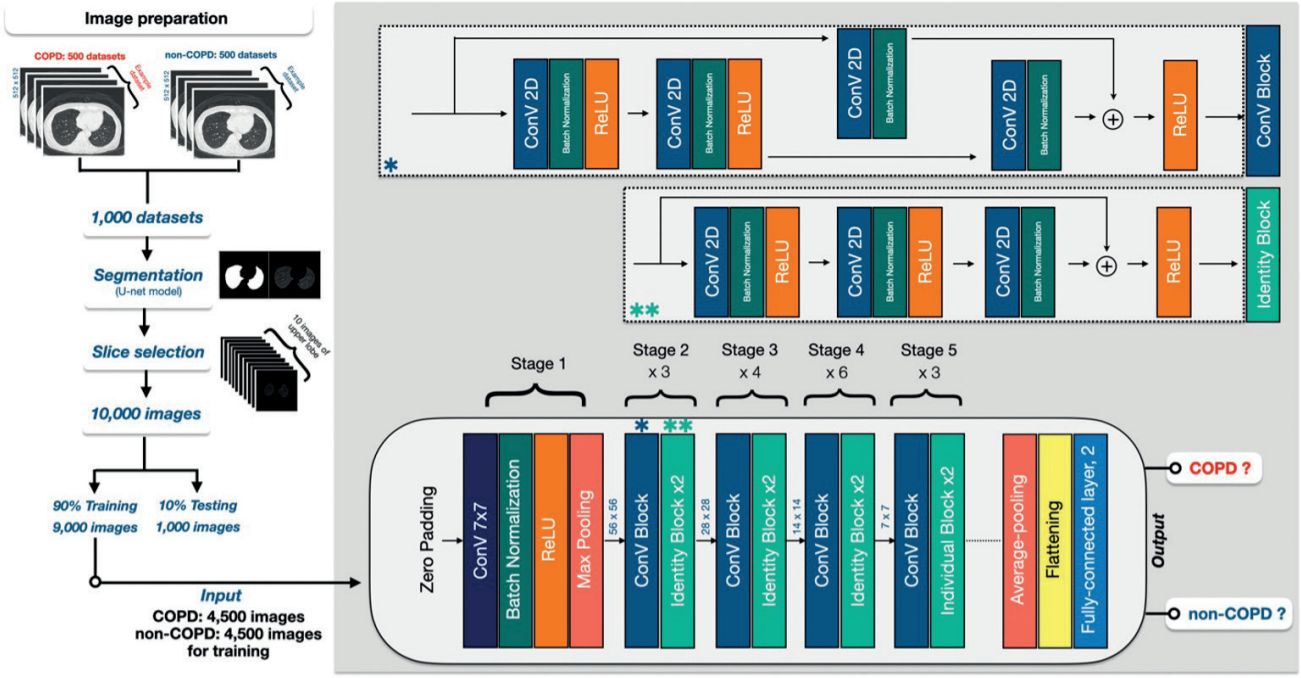

Figure 2 *Overall framework of the proposed architecture.*

Evaluation of model performance

We assessed the performance of the artificial intelligence model by analyzing the following metrics:

- A confusion matrix is a method for describing the performance of classification outcomes obtained when applying the model to the testing dataset. It consists of four fundamental characteristic values: True Positive (TP) represents the number of participants who are correctly classified as having COPD lesions under the spirometry-based diagnosis of COPD; True Negative (TN) represents the number of participants who are correctly classified as not having COPD lesions, by the spirometry-based diagnosis of non-COPD (healthy); False Positive (FP) represents the number of participants that are misclassified as suffering from COPD, in disagreement with the spirometry-based diagnosis of non-COPD (healthy); False Negative (FN) represents the number of participants misclassified as healthy, in dispute with the spirometry-based diagnosis of COPD.
- The accuracy of our model can be assessed by analyzing the values derived from the confusion matrix as follows:

a) Accuracy reflects the agreement between model predictions and spirometry-based diagnoses as detailed by equation 1:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

b) Precision reflects model accuracy by comparing TP with FP, as detailed in equation 2:

$$\text{Precision} = \frac{TP}{TP + FP} \quad (2)$$

c) Recall reflects model accuracy by comparing TP with FN, as detailed in equation 3:

$$\text{Recall} = \frac{TP}{TP + FN} \quad (3)$$

d) F1-score is the average between precision and recall as indicated in equation 4:

$$\text{F1 score} = 2 \times \frac{(\text{Precision} \times \text{Recall})}{(\text{Precision} + \text{Recall})} \quad (4)$$

- Area under the receiver operating characteristic curve (AUC)

The relationship between the True Positive Rate (TPR) and False Positive Rate (FPR) is indicated by the area under the Receiver Operating Characteristic (ROC) curve. An area under the ROC curve (AUC) close to 1 indicates excellent model performance.

Results

Model implementation

We performed hyperparameter optimization to identify the highest network classification performance. This tuning procedure optimized the use of weights returned by pre-training. The best-performing pre-trained ResNet-50 network was trained with the following hyperparameters: the optimization algorithm (sgdm) for optimization algorithm, batch size (32), maximum of epochs (100), learning rate (0.0001), momentum (0.9), L2 regularization (0.0001), and L2 normalization with a gradient threshold method. The model training accuracy and loss function of ResNet-50 are presented in Figure 3.

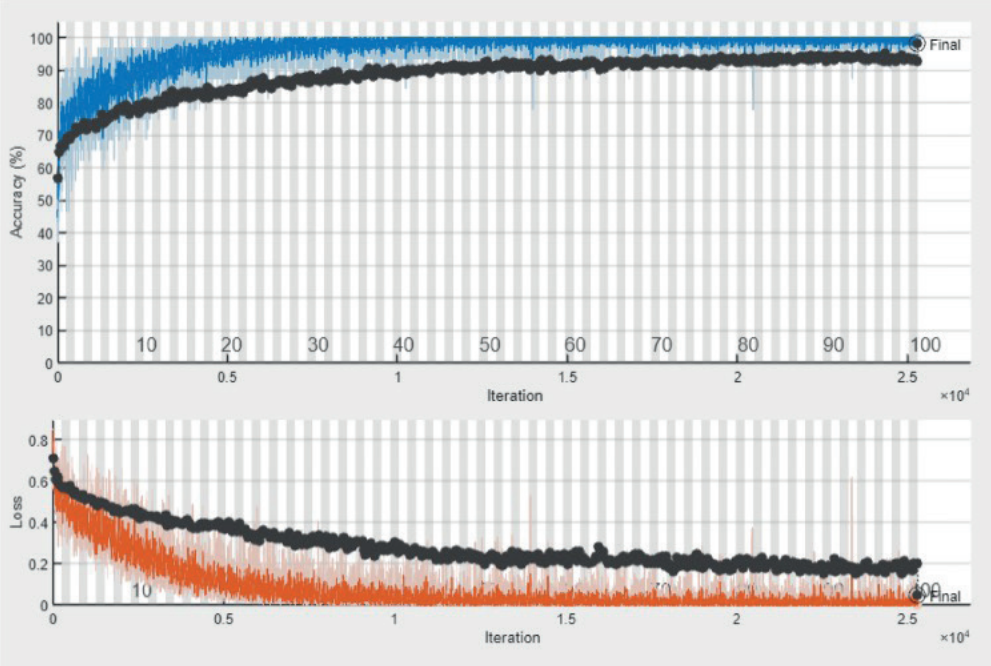

Figure 3 *Model accuracy and loss function. Upper: model training accuracy (blue) with validation accuracy (black) in the upper graph, lower: loss function (orange) with validation loss (black) after 100 epochs of training*

Model testing performance

Table 2 presents the classification results of testing images into COPD or non-COPD. The consistency between model prediction and ground truth (spirometry test) is notable, encompassing 894 images, with 481 identified as TP and 413 as TN. However, inconsistencies were observed in 106 images, with 19 identified as FN and 87 FP.

The performance evaluation of ResNet-50 achieved an accuracy of 0.89, precision of 0.85, recall of 0.96, and F1-score of 0.90. Figure 4 demonstrates the ROC curve from ResNet-50 model with AUC of 0.97. Notably, ResNet-50 demonstrated superior model performance compared with VGG16 and Inception V3 (see Table 3 and Figure 4).

Table 2 Confusion matrix of model prediction versus spirometry test.

		Spirometry test	
		COPD	non-COPD
Model prediction	COPD	481 images (TP)	87 images (FP)
	non-COPD	19 images (FN)	413 images (TN)

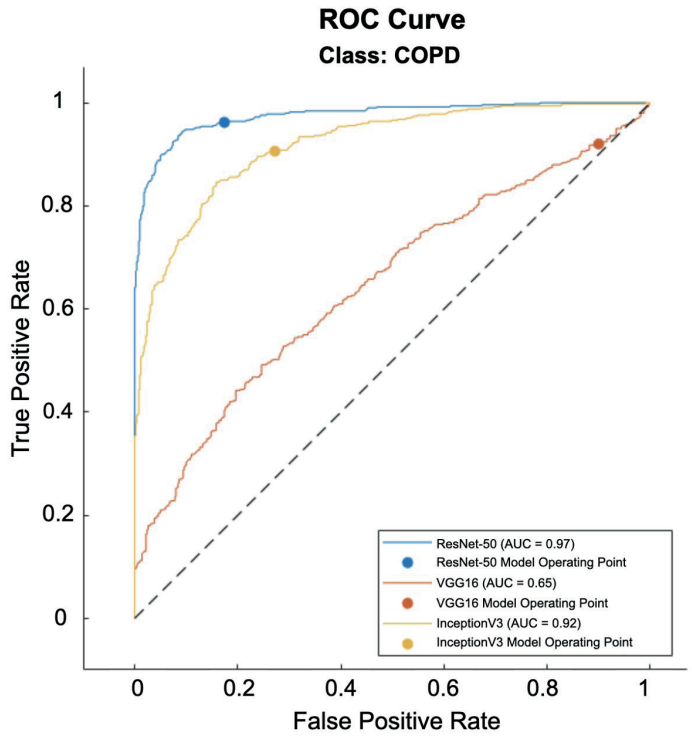

Figure 4 *Receiver operating characteristic (ROC) curve of ResNet-50, Inception V3, and VGG16.*

The performance evaluation of other pre-trained models, including Inception V3 and VGG16, was elaborated. Inception V3 achieved an accuracy, precision, recall, and F1-score of 0.82, 0.77, 0.91, and 0.83, respectively. Meanwhile, VGG16 attained an accuracy, precision, recall, and F1-score of 0.51, 0.51, 0.92, and 0.65, respectively. Each model exhibited an AUC of 0.92 and 0.65 (Figure 4), respectively, lower than that of ResNet-50. The detailed performance characteristics of Inception V3 and VGG16 are provided in supplementary figures 1-2. Furthermore, a comprehensive summary of model performance and comparative analysis with previous studies is presented in Table 3.

Table 3 The comparison of predictive model performance across previous studies

Authors	Models	Number of CT datasets	Model performance				
			AUC	Accuracy (%)	Precision (%)	Recall	F1-score
Present study	ResNet-50	1,000	0.97	89.40	84.68	0.96	0.90
	Inception V3	1,000	0.92	81.70	76.91	0.91	0.83
	VGG16	1,000	0.65	51.00	50.55	0.92	0.65
Ho TT *et al.* [18]	3D-CNN (Naive model)	596	0.93	89.30	82.60	0.88	0.85
Tang *et al.*[15]	Deep residual networks with 152 layers	2,589	0.889	n/a	74.00	0.79	n/a
González *et al.*[1]	CNN analysis	7,963	0.856	74.95	n/a	n/a	n/a

Note: AUC: area under the ROC Curve, CNN: convolutional Neural Network, 2D: two-dimensional image, 3D: three-dimensional image or volumetric data.

Discussion

The present study used a pretrained ResNet-50 to identify COPD from low-dose CT images. Specifically, we capitalized on clinical evidence indicating a prominent presence of COPD lesions in the upper lobe of the lung. This observation prompted us to focus only on upper lung slices, rather than the whole lung, as input to the model, to reduce computational time and workload. Remarkably, the proposed model achieved an AUC value of 0.97, accuracy of 0.89, precision of 0.85, recall of 0.96, and F1-score of 0.90. The performance of our model is comparative with previous studies [1,15,18] (see Table 3). Notably, we achieved comparable predictive efficacy using a significantly smaller number of participants than the previous studies.[1,15]

While our model demonstrated acceptable performance for classifying COPD, the confusion matrix revealed 87 instances of non-COPD images being misclassified as COPD, and 19 misclassified as non-COPD. These instances of misclassification may be attributed to inaccuracies in labeling the COPD input during the training phase. While it is the case that COPD participants were labeled using the gold standard method of spirometry testing, we cannot be sure that all images from these participants necessarily contain COPD lesions. Some slices without lesions were labeled and integrated into the COPD training set. Consequently, the model may have misinterpreted and erroneously classified some non-COPD images as COPD during the testing phase. Future studies should establish robust techniques to improve the accuracy of input labeling.

To develop a more precise and reliable COPD classification, we must acknowledge and address certain limitations of our study. First, given that over half of the COPD patients were categorized into GOLD stage 1 (mild COPD), there is a potential concern that the model could misclassify them as usual. This observation underscores the importance of future studies with larger populations representing each COPD stage. Such studies could facilitate the development of more robust classification models capable of distinguishing between COPD and non-COPD individuals, as well as accurately identifying different COPD stages.

Second, the reliance on spirometry to label individuals as COPD or non-COPD introduces potential uncertainties when using 2D CT images as input to the model. While we deliberately tried to incorporate ten slices from the upper lobe of labeled COPD individuals, this approach does not guarantee the presence of COPD lesions in all slices. It is, therefore, advisable to submit all input images to thorough verification by experienced radiologists, thereby ensuring precise and reliable labeling. This verification process may entail the exclusion of numerous images, which may potentially result in an inadequate number of images for training. Training with 3D input data may address this issue. A 3D image of the entire lung or some relevant part from an individual labeled as COPD may guarantee the inclusion of COPD lesion characteristics for each training input. Although the adoption of 3D input data is promising,[1,15,18] it is essential to balance performance optimization and computational feasibility.

Additionally, no practical protocol for selecting specific lung regions has been established yet. Thus, further research and refinement are warranted to determine the most pragmatic and effective means of harnessing 3D input data for improved classification performance. Striking a balance between computational feasibility and the comprehensive representation of essential features remains a central challenge.

In addition to substantial computational demands, implementing 3D input data requires a significantly larger dataset to ensure the availability of a sufficiently large training dataset. A viable solution to this problem may involve collaborations among multiple centers, to collect more diverse and extensive datasets.[19] Such collaborative efforts have the potential to enrich the variety and representation of COPD cases, allowing models to generalize across different patient populations. However, to effectively utilize CT datasets from diverse sources, careful standardization and normalization procedures must be implemented.[20] These essential procedures could address differences in imaging protocols, acquisition parameters, and equipment variations.[19]

Third, we specified the upper lobe of all lungs at slice numbers 52 to 61, guided by the average lung dimensions observed in Thai individuals.[16] Nevertheless, individual lung sizes may vary due to age, ethnicity, and physiological profiles. Future studies should consider implementing advanced segmentation techniques[21,22] for the precise delineation of individual lungs before their division into distinct components. This refined approach would enable the accurate selection of the upper portion of each lung as input for the model, thereby enhancing the precision and customization of assessments.

Lastly, k-fold validation is a fundamental feature that addresses the common challenge of overfitting in deep learning. By adopting batch image processing and implementing k-fold validation, the generalization capacity and accuracy of the model are significantly enhanced. These critical steps contribute to a more robust and reliable classification system, minimizing overfitting concerns and bolstering the ability of the model to handle diverse datasets effectively.[23,24]

Conclusion

In conclusion, our study provides empirical evidence supporting the efficacy of deep learning for classifying COPD from non-COPD cases using low-dose CT images. Notably, our model achieved 89.40% accuracy using less input data than previous studies. This result highlights the potential utility of our model as a pre-screening aid for radiologists, particularly in lung cancer programs for smokers. It may also serve as a valuable screening tool for COPD in participants who may be limited in their ability to undergo spirometry. In the future, this model may be able to classify COPD severity if provided with sufficient data.

Additionally, the incorporation of activation maps shows promise for highlighting precise lesion locations and expediting diagnosis. Ultimately, diagnostic aids in deep learning models may offer invaluable insights for

radiologists and clinicians, enhancing clinical observations and patient follow-up care. These advancements elevate the diagnostic process, rendering it more accurate and efficient.

Acknowledgements

We would like to express our sincere gratitude to Chulabhorn Royal Academy for the generous support through grant number RAA2564/036 and RCP2555/002. In addition, we would like to extend our appreciation to all the participants in Chulabhorn Royal Academy Integrated Lung Cancer Screening Program.

References

[1] González G, Ash SY, Vegas-Sánchez-Ferrero G, Onieva JO, Rahaghi FN, Ross JC, *et al*. Disease staging and prognosis in smokers using deep learning in chest computed tomography. Am J Respir Crit Care Med. 2018; 197(2): 193-203. doi: 10.1164/rccm.201702-0388OC.

[2] MacNee W. Pathology, pathogenesis, and pathophysiology. BMJ. 2006; 332(7551): 1202-4. doi: 10.1136/bmj.332.7551.1202.

[3] Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease [Internet]. 2024. Available from: https://goldcopd.org/2024-gold-report/. Accessed February 18, 2024.

[4] Lamprecht B, Soriano JB, Studnicka M, Kaiser B, Vanfleteren LE, Gnatiuc L, *et al*. Determinants of underdiagnosis of COPD in national and international surveys. Chest. 2015; 148(4): 971-85. doi: 10.1378/chest.14-2535.

[5] Johns DP, Walters JA, Walters EH, *et al*. Diagnosis and early detection of COPD using spirometry. J Thorac Dis. 2014; 6(11): 1557-69. doi: 10.3978/j.issn.2072-1439.2014.10.11.

[6] Coates AL, Graham BL, McFadden RG. Spirometry in primary care. Can Respir J. 2013; 20(1): 13-22. doi: 10.1155/2013/485906.

[7] Cooper BG. An update on contraindications for lung function testing. Thorax. 2011; 66(8): 714-23 doi: 10.1136/thx.2010.140236.

[8] Gierada DS, Black WC, Chiles C, Pinsky PF, Yankelevitz DF, *et al*. Low-dose CT screening for lung cancer: Evidence from 2 decades of study. Radiol Imaging Cancer. 2020; 2(2): e190058. doi: 10.1148/rycan.2020190058.

[9] Nawa T. Low-dose CT screening for lung cancer reduced lung cancer mortality in Hitachi City. Int J Radiat Biol. 2019; 95(10): 1441-6. doi: 10.1080/09553002.2019.1614235.

[10] Larke FJ, Kruger RL, Cagnon CH, Flynn M, McNitt-Gray MM, Wu X, *et al*. Estimated radiation dose associated with low-dose chest CT of average-size participants in the National Lung Screening Trial. AJR Am J Roentgenol. 2011; 197(5): 1165-9. doi: 10.2214/AJR.11.6707.

[11] Bailey KL. The importance of the assessment of pulmonary function in COPD. Med Clin North Am. 2012; 96(4): 745-52. doi: 10.1016/j.mcna.2012.03.009.

[12] Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation. Springer International Publishing; 2015. p. 234-41.

[13] Tan W, Huang P, Li X, Ren G, Chen Y, Yang. J. Analysis of segmentation of lung parenchyma based on deep learning methods. J Xray Sci Technol. 2021; 29(6): 945-59. doi: 10.3233/XST-218010.

[14] Nemec SF, Bankier AA, Eisenberg RL. Upper lobe-predominant diseases of the lung. AJR Am J Roentgenol. 2013; 200(3): W222-W237. doi: 10.2214/AJR.12.9544.

[15] Tang LYW, Coxson HO, Lam S, Leipsic j, Sin D. Towards large-scale case-finding: Training and validation of residual networks for detection of chronic obstructive pulmonary disease using low-dose CT. Lancet Digit Health. 2020; 2(5): e259–e267. doi: 10.1016/S2589-7500(20)30100-1.

[16] Bradley P, Fuhrman M, Zimmerman M. Pediatric critical care (Fourth Edition). St. Louis; 2011.

[17] He K, Zhang X, Ren S, Sun J, editors. Deep Residual Learning for Image Recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR); 2016 27-30 June 2016.

[18] Ho TT, Kim T, Kim WJ, Lee CH, chae KJ, Bak SH, *et al*. A 3D-CNN model with CT-based parametric response mapping for classifying COPD subjects. Sci Rep. 2021; 11: 34. doi: 10.1038/s41598-020-79336-5.

[19] Gierada DS, Bierhals AJ, Choong CK, Bartel ST, Ritter JH, Das NA. et al. Effects of CT section thickness and reconstruction kernel on emphysema quantification relationship to the magnitude of the CT emphysema index. Acad Radiol. 2010; 17(2): 146-56. doi: 10.1016/j.acra.2009.07.025.

[20] Selim M, Zhang J, Fei B, Zhang GQ, Chen J. STAN-CT: Standardizing CT Image using Generative Adversarial Networks. AMIA Annu Symp Proc. 2020; 2020: 1100-9.

[21] Ait Skourt B, El Hassani A, Majda A. Lung CT Image Segmentation Using Deep Neural Networks. Procedia Computer Science. 2018; 127: 109-13.

[22] Murugappan M, Bourisly AK, Prakash NB, Sumithra MG, Acharya UR. et al. Automated semantic lung segmentation in chest CT images using deep neural network. Neural Comput Appl. 2023; 35(21): 15343-64. doi: 10.1007/s00521-022-07399-7.

[23] Rodriguez JD, Perez A, Lozano JA. Sensitivity analysis of k-fold cross validation in prediction error estimation. IEEE Trans Pattern Anal Mach Intell. 2010; 32(3): 569-75. doi: 10.1109/TPAMI.2009.187.

[24] Anguita D, Ghelardoni L, Ghio A, Oneto L, Ridella S. The 'K' in K-fold cross validation. In: ESANN 2012 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 25-27 April 2012. i6doc.com publ. ISBN 978-2-87419-049-0. Available from: https://www.esann.org/sites/default/files/proceedings/legacy/es2012-62.pdf.