A Dual-Branch Network for Classifying Pulmonary Disease Using Flow-Volume Loops

line 1: Jaeun Kim
line 2: *Department of Computer Engineering*
line 3: *Daejeon University*
line 4: Daejeon, Republic of Korea
line 5: kju0048@gmail.com

line 1: Hongjun Kim
line 2: *Department of Computer Engineering*
line 3: *Daejeon University*
line 4: Daejaeon, Republic of Korea
line 5: hjkim99@dju.kr

***Abstract*— Although flow-volume loops from spirometry are fundamental for diagnosing pulmonary diseases, distinguishing subtle morphological patterns remains challenging—even for experienced clinicians, especially when it comes to differentiating between obstructive and restrictive types. In this paper, a deep learning framework is proposed to automatically classify lung diseases based on these loops. As its core methodology, this work presents a dual-branch architecture that decomposes the complex four-class classification task—comprising normal, obstructive, restrictive, and mixed types—into two independent binary classifiers: one for detecting obstructive types and the other for restrictive types. Additionally, to ensure robust performance despite severe class imbalance in real-world medical data, a generative data augmentation strategy operating within the learned latent space of minority classes is employed. By modeling the underlying data manifold with a Convolutional Variational Autoencoder (VAE) and synthesizing high-fidelity samples, the training distribution is effectively rebalanced. This synergistic combination of a dual-branch structure and generative data augmentation significantly enhances the classifier's generalization performance, demonstrated by an increase of 0.026 in accuracy and 0.126 in F1-score over baseline models, marking a substantial advancement towards robust and scalable automated screening for pulmonary diseases.**

Keywords—pulmonary disease classification, flow-volume loops, dual-branch CNN, VAE, data augmentation

I. INTRODUCTION

Chronic obstructive pulmonary disease (COPD) represents a significant global health burden, ranking as the fourth leading cause of death worldwide. In 2021, it was responsible for 3.5 million deaths, accounting for approximately 5% of all global fatalities. Beyond mortality, COPD is also the eighth leading cause of poor health as measured by disability-adjusted life years (DALYs). The impact of this disease is disproportionately concentrated in low- and middle-income countries (LMICs), where nearly 90% of deaths among individuals under 70 years of age occur. The primary risk factors also differ by economic setting; while tobacco smoking accounts for over 70% of COPD cases in high-income countries, it is responsible for 30–40% of cases in LMICs, where household air pollution is also a major contributing risk factor [1].

The standard diagnostic and classification tool for COPD is spirometry, which measures forced expiratory volume in one second (FEV_1) and forced vital capacity (FVC) and computes their ratio (FEV_1/FVC). A ratio below 0.70 is diagnostic of obstructive impairment [2], [3], and further stratification into normal, obstructive, restrictive and mixed types guides treatment planning. Predicted values for FEV1_pred and FVCpred are conventionally obtained from multivariate linear regression models—such as the Global Lung Function Initiative (GLI-Global) equations—that incorporate gender, age, ethnicity, height and weight [3]. However, these global reference equations may not accurately reflect the genetic and environmental characteristics of specific populations, particularly within East Asia, and they fail to account for individual factors such as smoking history, occupational exposures or obesity [4], [5]. Consequently, substantial discrepancies can arise between measured and predicted values. To address this issue, the Choi regression equations, developed on Korean patient data, are employed to generate more accurate FEV_1_pred and FVCpred values for this cohort and thereby improve the reliability of COPD type labeling [6].

Spirometric results are often visualized as flow–volume loops, in which the x-axis represents lung volume (L) and the y-axis represents airflow rate (L/s). Normal loops exhibit a smooth rise to peak flow followed by a gradual decline, whereas obstructive loops display a markedly reduced peak flow and concave descending limb, and restrictive loops show narrowed loop width with a lower peak. Prior studies have leveraged these morphological features for automated classification: attempts have been made to classify these features using machine learning techniques. Nevertheless, these approaches still rely on expert interpretation at the final decision stage, require patients to undergo in-person testing, and are subject to inter-observer variability in loop interpretation [7], [8]. Particularly, the differentiation between restrictive and obstructive lung diseases remains challenging due to overlapping morphological features and subtle variations in loop patterns that may not be easily distinguishable even by experienced clinicians [9].

To address the challenges of subjective interpretation and inaccurate reference equations, a deep learning framework is proposed that takes flow–volume loop images as input and automatically classifies them into normal, obstructive, restrictive, and mixed types. To handle the class imbalance common in medical datasets, SMOTE (Synthetic Minority Over-sampling Technique)-based minority oversampling is utilized. The core approach employs a VGG16-based Dual-Branch convolutional neural network [10]. The major contributions are summarized below:

- A novel deep learning framework is proposed that provides an end-to-end solution for objective, automated classification of spirometric types directly from flow–volume loop images, thereby reducing the need for subjective expert judgment.

- The Choi regression equations, developed specifically on Korean patient data, are utilized to generate more reliable ground-truth labels. This addresses the limitations of global reference equations and improves the accuracy of type classification for this cohort.
- A VGG16-based Dual-Branch CNN is designed as the primary model, leading to more robust and accurate classification.
- SMOTE-based oversampling is implemented to effectively address the class imbalance problem, significantly improving predictive performance for underrepresented spirometric types [11].

II. DATASET

The pulmonary function test data used in this study were measured using a spirometry kit manufactured by TR Co., and the raw measurements were provided by Hanaro Leaders Healthcare Co., Ltd. for analysis. The dataset comprises a total of 4,023 records.

TABLE I. SPIROMETRY DATASET VARIABLES

Column	*Unit*	*Description*
Date	YYYY-MM-DD	Date of examination
id	-	Patient identifier
Birth	YYYY-MM-DD	Patient's date of birth
Sex	Male/Female	Patient's biological sex
Height	cm	Patient's height
Weight	kg	Patient's weight
FVC	L	Forced vital capacity
FEV_1	L	Forced expiratory volume in 1 second
FEV_1/FVC	%	Ratio of FEV_1 to FVC
FEF25-75%	L/s	Forced expiratory flow at 25-75% of FVC
PEF	L/s	Peak expiratory flow

A. Data Structure

Each patient's test result is represented by a flow–volume loop image file matched one-to-one with a corresponding row in an Excel spreadsheet. The spreadsheet contains the following columns, and the total number of entries is 4,023. The detailed layout of the Excel file is shown in Table 1.

Each patient's test result image is saved as a flow–volume loop, where the x-axis corresponds to lung volume (L) and the y-axis to instantaneous flow rate (L/s). For example, the loop shown in Figure 1 rises sharply as volume increases and then declines smoothly beyond the peak. In this case, the relatively high peak flow and gradual descending limb closely resemble a normal type [12].

Classification of each subject's flow–volume loop into one of four types (normal, obstructive, restrictive, mixed) was performed based on the computation of two primary metrics:

$$ratio = \frac{FEV_1}{FVC} \quad (1)$$

$$FVC\%Pred = \frac{FEV_{measured}}{FVC_{pred}} \quad (2)$$

where $FVC_{measured}$ is the forced vital capacity measured in liters, and FVC_{pred} is the predicted normal value in liters obtained via Choi's regression equations.

Fig. 1. Example Flow-Volume Loop Image. (Upscaled)

Clinically, an FEV_1/FVC ratio below 0.70 indicates airflow obstruction, while an FVC %Pred below 80 % suggests a restrictive ventilatory defect, according to ATS/ERS and GOLD guidelines. The classification rules are summarized in Table 2.

TABLE II. SPIROMETRIC TYPE CLASSIFICATION CRITERIA

Types	*FEV_1/FVC(ratio)*	*FVC%Pred*
Normal	≥ 0.70	≥ 80%
Obstructive	< 0.70	≥ 80%
Restrictive	≥ 0.70	< 80%
Mixed	< 0.70	< 80%

Choi et al. derived population-specific regression equations for spirometric indices by analyzing data from healthy Korean adults. Unlike the Global Lung Function Initiative (GLI) equations, which pool multi-ethnic cohorts, the Choi regressions reflect Korea's unique demographic and environmental characteristics, providing more accurate predicted values for forced vital capacity (FVC) in this population. In their original study, Choi and colleagues also observed that the lower limit of normal (LLN)—defined as the 5th percentile of the predicted distribution—could be closely approximated by multiplying the mean predicted value by 0.95.

TABLE III. FVC PREDICTION COEFFICIENTS BY GENDER

Variable	*Male*	*Female*
Constant (β0)	-4.8434	-3.0006
Age2(year2)	-0.00008633	-0.0001273
Height(cm)	0.05292	0.03951
Weight(kg)	0.01095	0.006892

Accordingly, each subject's LLN-corrected predicted values were computed using Choi's gender-specific formulas (coefficients shown in Table 3). Based on the diagnostic

criteria summarized in Table 2, each flow–volume loop was classified into one of four types. Table 4 summarizes the sample counts for each type, and Figure 4 presents representative flow–volume loop images illustrating the Normal, Obstructive, Restrictive, and Mixed types.

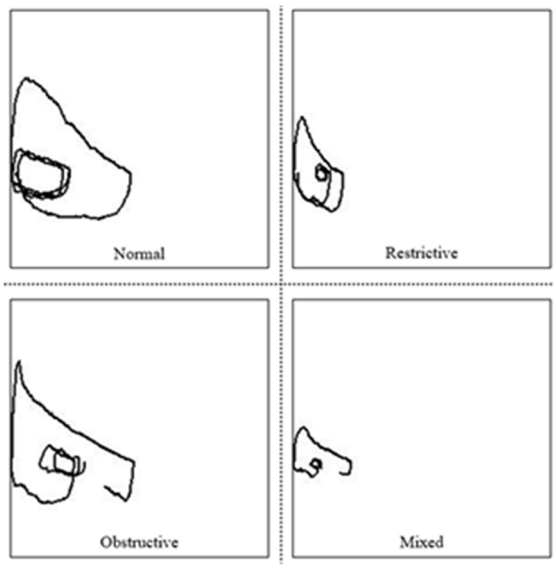

Fig. 2. Flow-Volume Loop Images Classified by Type

TABLE IV. DATASET DISTRIBUTION BY TYPE

Types	*Sample Count*
Normal	2,242
Obstructive	60
Restrictive	1,641
Mixed	80

As shown in Figure 2, the top-left loop exhibits a smooth, full expansion to a high peak flow followed by a gradual decline—hallmarks of the Normal type. The top-right loop is noticeably narrowed in overall volume while maintaining its descending limb shape, indicative of a Restrictive defect. In the bottom-left loop, the peak flow is markedly reduced and the descending limb is concave, characteristic of an Obstructive type. Finally, the bottom-right loop displays both a lowered peak flow and a narrowed loop width, consistent with a Mixed type.

As shown in Table 4, the Normal and Restrictive classes together comprise approximately 90 % of all samples, while the Obstructive and Mixed classes each represent only about 5 %. This severe class imbalance risks underrepresentation of minority types during model training and can degrade predictive performance. Therefore, creating a more balanced training set through data augmentation is a crucial step to mitigate this issue.

III. DATA AUGMENTATION

To augment data, we present the latent-space augmentation strategies developed to address the severe underrepresentation of the Obstructive and Mixed classes. Standard image augmentation techniques, such as rotations, translations, or cropping, are avoided because they can distort crucial axis and scale information within flow-volume loop images, thereby undermining their clinical interpretability.

The ConvVAE encoder processes each 3×224×224 RGB input through four sequential Conv2D+ReLU blocks. This process progressively expands the channel dimension from 3→32→64→128→256 while simultaneously reducing the spatial resolution from 224→112→56→28→14. Following these convolutional layers, the output is flattened and then projected into a mean (μ) and log-variance ($\log\sigma^2$) via two distinct linear layers [13].

The D-dimensional latent vector zi is then sampled using the reparameterization trick, as described in Equation 3:

$$\mathbf{Z_i} = \mu + \exp(0.5\,\log\sigma^2) \odot \epsilon, \epsilon \sim N(0, I) \quad (3)$$

Here, ϵ is a random sample drawn from a standard normal distribution *N(0, I)*, and denotes element-wise multiplication. This trick enables backpropagation through the stochastic sampling process, which is crucial for training VAEs effectively.

The decoder then takes this latent vector zi and restores it to a 256×14×14 tensor through a linear layer. It then applies four rounds of bilinear upsampling (with a scale factor of 2) each followed by a Conv2D+BatchNorm+ReLU block. Ultimately, a 3×224×224 image is reconstructed via a Sigmoid activation function. This specific architecture is designed to ensure that the learned latent space faithfully preserves the manifold of valid volume-flow loops, maintaining the critical clinical information. A detailed schematic of the model architecture is provided in Figure 3.

A. SMOTE Interpolation

SMOTE (Synthetic Minority Over-sampling Technique) is a widely adopted method for augmenting minority class samples. In this study, SMOTE is adapted for the latent space of ConvVAE to effectively increase sample counts while preserving the semantic structure of the data.

F or the latent space mean (μ_i) and log-variance ($\log\sigma_i^2$) of each original sample, two random original samples (corresponding to latent distribution parameters z_i and z_j) are selected. The latent space mean (μ) and log-variance ($\log\sigma^2$) of these two samples are then linearly interpolated separately to generate a new synthetic mean (μ_{synth}) and log-variance ($\log\sigma_{synth}^2$). In this process, λ is a random value between 0 and 1, which determines the interpolation point along the linear path between the two original samples. The new synthetic mean and log-variance are calculated according to Equations 4 and 5, respectively:

$$\mu_{synth} = \lambda \cdot \mu_i + (1 - \lambda) \cdot \mu_j \quad (4)$$

$$log\sigma^2_{synth} = \lambda \cdot log\sigma_i^2 + (1 - \lambda) \cdot log\sigma_j^2 \quad (5)$$

The final synthetic latent code Zsynth is then sampled through reparameterization using these interpolated mean and log-variance values, as shown in Equation 6:

$$\mathbf{Z_{synth}} = \mu_{synth} + \exp\left(0.5\,\log\sigma^2_{synth}\right) \odot \epsilon, \epsilon \sim N(0, I), \lambda \sim U(0,1) \quad (6)$$

B. Gaussion-Noise Augmentation

To capture fine-grained local variations that interpolation alone may miss, we add isotropic Gaussian noise to each

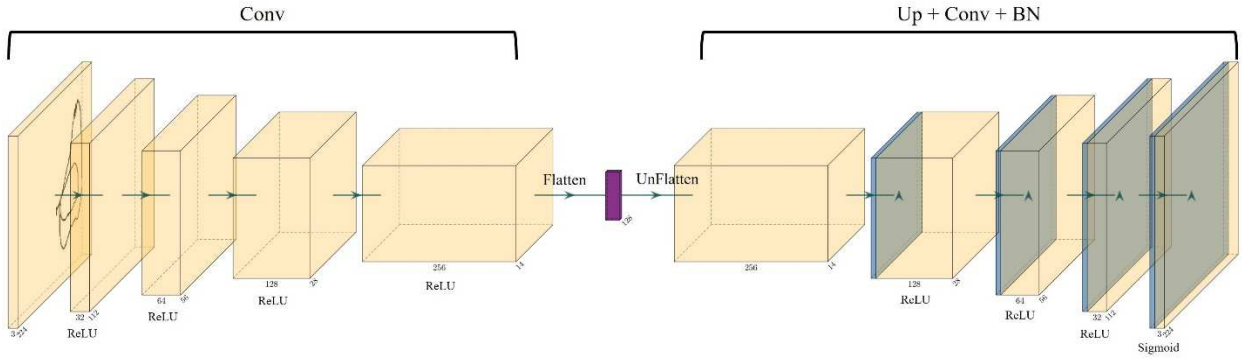

Fig. 3. Architecture of the Proposed ConvVAE Model

original code. Generating N noise-perturbed codes simulates natural variability—such as slight measurement noise or subtle morphological differences—and enriches the neighborhood around each latent centroid [14].

$$\mathbf{Z_{noise}} = z_i + \delta,\ \delta \sim N(0, (0.05)^2 I) \quad (7)$$

Finally, the 2N synthetic latent vectors, generated by Equations 6 and 7, are decoded through the ConvVAE decoder to obtain high-resolution flow–volume loop images. For each of the Obstructive and Mixed classes, the N original samples are combined with the N SMOTE-interpolated and N noise-augmented images, yielding 3N training examples per class. These are then merged with the actual Normal and Restrictive samples to construct a balanced training set. This approach effectively prevents underfitting on rare types and substantially boosts the classifier's overall performance.

IV. DUAL-BRANCH CNN

This paper introduces a Dual-Branch Convolutional Neural Network (Dual-Branch CNN) to classify the four types of pulmonary function (Normal, Obstructive, Restrictive, and Mixed) in a more systematic and interpretable manner.

The core idea of this model is to move away from the conventional approach of directly classifying four independent categories and instead mimic the fundamental logic of clinical diagnosis. That is, the model is designed to independently infer two fundamental properties from a given flow-volume loop image: (1) the presence or absence of obstructive impairment and (2) the presence or absence of restrictive impairment. The probabilistic results of these two binary judgments are then combined to perform the final four-class classification.

A. Key components of the model

- Shared Feature Extractor: The pre-trained VGG16 is utilized from the timm library as a backbone to extract rich visual features from the input flow-volume loop images.
- Channel Attention Module: The extracted feature map is passed through a Squeeze-and-Excitation (SE) block [15]. The SE-Block acts as an attention mechanism, dynamically recalibrating the importance of each channel to focus on more informative features for classification.
- Dual Branches & Probability Combination: The feature vector, enhanced by the attention module, is fed into two independent pathways: an 'Obstruction Branch' and a 'Restriction Branch'. While conventional dual-branch architectures often produce separate outputs for each task, the proposed model is distinct in that each branch performs a binary classification, and their probabilistic outputs are subsequently combined to produce a single, final four-class classification result [16], [17], [18].

B. Detailed Components and Training Process

The model's process begins by passing a 224×224×3 input image through the pre-trained VGG16 backbone to extract a high-dimensional visual feature vector. This feature vector is then refined by a SE block, which re-weights channel-wise importance to emphasize useful information for classification.

This enhanced feature vector is simultaneously fed into two independent fully connected layers: the 'Obstruction Branch' and the 'Restriction Branch.' The Obstruction Branch analyzes the feature vector to determine the presence of obstructive impairment, outputting a probability distribution P_obs = [P(¬O), P(O)] for 'non-obstructive' and 'obstructive' via a softmax function. In the same manner, the Restriction Branch assesses for restrictive impairment, calculating a probability distribution P_res = [P(¬R), P(R)] for 'non-restrictive' and 'restrictive'.

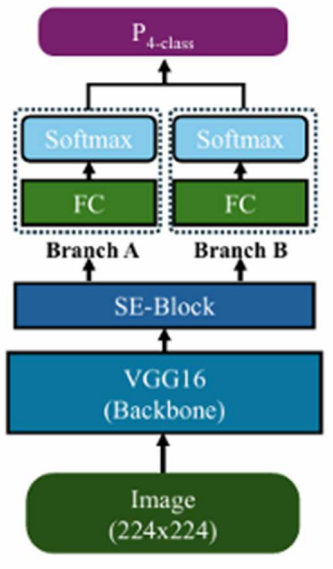

Fig. 4. Overall Architecture of the Proposed Dual-Branch CNN

These two probability distributions are then combined to generate the final four-class classification output through an outer product operation. Specifically, the joint probabilities are computed as follows: P(Normal) = P(¬O) × P(¬R), P(Obstructive) = P(O) × P(¬R), P(Restrictive) = P(¬O) × P(R), and P(Mixed) = P(O) × P(R). These four probabilities form the final class distribution, which is then converted to log-probabilities and optimized using the NLLLoss (Negative Log Likelihood Loss) function to update the model weights during training. A schematic diagram of the Dual-Branch model architecture is presented in Figure 4.

V. EXPERIMENTS

All models were implemented using the PyTorch framework [19]. As detailed in Table 5, the dataset was partitioned into training and testing sets at a 7:3 ratio. The values in parentheses within the table denote the original sample counts for the minority classes before augmentation. To address this class imbalance, the Obstructive and Mixed classes were augmented with synthetic samples equivalent to twice their original count using a latent-space technique. For training, the Adam optimizer was used, and the batch size, epochs, and learning rate were set to 32, 100, and 1e-4, respectively.

TABLE V. TRAIN/TEST SPLIT

Class	*Train*	*Test*
Normal	1569	673
Obstructive	126(42)	18
Restrictive	1148	493
Mixed	168(56)	24

A. Performance

The detailed performance of the final proposed model—the Dual-Branch CNN trained with latent-space data augmentation—on the test set is presented in Table 6.

TABLE VI. MODEL PERFORMANCE METRICS

Class	*Precision*	*Recall*	*F1-score*
Normal	0.9443	0.9316	0.9379
Obstructive	0.9444	0.9444	0.9444
Restrictive	0.9062	0.9209	0.9135
Mixed	0.9200	0.9583	0.9388

The final model achieves a high overall accuracy of 92.80% and a macro F1-score of 0.9337. Notably, the model demonstrates outstanding performance on the underrepresented minority classes, achieving F1-scores of 0.9444 for Obstructive and an impressive 0.9388 for Mixed. This result strongly suggests that the combination of the proposed Dual-Branch architecture and the data augmentation strategy is highly effective in overcoming the challenges posed by severe class imbalance. Additionally, the model maintains robust performance on the majority classes, Normal (F1-score: 0.9379) and Restrictive (F1-score: 0.9135), indicating a well-balanced and reliable classification capability across all types.

B. Comparative Analysis

To dissect the individual contributions of the proposed components, a comparative analysis of model performance was conducted by varying two factors: the application of data augmentation and the use of the Dual-Branch architecture. The results are summarized in Table 7.

The analysis reveals two key insights. First, data augmentation provided the most significant performance improvement. Regardless of the model architecture, applying the latent-space augmentation technique substantially increased the macro F1-score (from 0.8069 to 0.8850 for the standard CNN, and from 0.9144 to 0.9337 for the Dual-Branch CNN). This confirms that the proposed augmentation strategy is highly effective for enhancing the model's ability to recognize minority class types.

Second, the Dual-Branch architecture consistently outperformed the standard CNN baseline under identical data conditions. This suggests that decomposing the multi-class problem into two independent binary classifications (obstructive and restrictive) provides a more effective learning pathway than treating the four types as distinct, unrelated categories.

Ultimately, the combination of both data augmentation and the Dual-Branch CNN yielded the best overall performance (Accuracy: 0.9280, F1-score: 0.9337). This indicates a synergistic effect, where the specialized architecture and the balanced dataset work together to maximize classification accuracy.

TABLE VII. ABLATION STUDY RESULTS

Class		*Recall*	*F1-score*
Dual-Branch	**Data Augment**		
Not Applied	Not Applied	0.9015	0.8069
Applied	Not Applied	0.9106	0.9144
Not Applied	Applied	0.8940	0.8850
Applied	Applied	0.9280	0.9337

VI. CONCLUSION

In this paper, the novel deep learning framework is proposed to address the key challenges in the automated classification of spirometric types from flow-volume loop images: the subtle morphological similarities between types and severe class imbalance. The approach is characterized by the synergistic combination of two core components: the Dual-Branch CNN architecture that mimics clinical diagnostic reasoning by independently assessing obstructive and restrictive features, and the generative data augmentation strategy that uses the Convolutional Variational Autoencoder (VAE) to effectively rebalance the training data by synthesizing high-fidelity samples in the latent space.

The experimental results demonstrate the significant efficacy of this combined approach. The final model achieves a high overall accuracy of 92.80% and a macro F1-score of 0.9337. The ablation studies confirm that both the Dual-Branch architecture and the latent-space augmentation individually contributed to performance gains, with their combination yielding the best results. The model's

exceptional performance on the underrepresented Obstructive and Mixed classes, in particular, validates the hypothesis that rebalancing the data distribution with high-quality synthetic samples is critical for robust classification.

The primary contribution of this work is a robust, end-to-end framework that can enhance the accuracy of automated spirometric analysis. By decomposing the classification task, the proposed model offers a pathway toward more transparent AI-driven diagnostics, and the success of this generative augmentation technique provides a viable solution for the pervasive problem of data imbalance in medical imaging. Ultimately, this study represents a significant step toward developing automated, scalable, and reliable screening tools to aid clinicians in diagnosing pulmonary diseases. Future work should focus on validating the model's performance on larger, multi-ethnic datasets and exploring the integration of other generative models for data augmentation.

ACKNOWLEDGM

This research was supported by a grant of the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare, Republic of Korea (grant number : RS-2025-02263696)

REFERENCES

[1] WHO, "Chronic obstructive pulmonary disease (COPD)," 2024. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/chronic-obstructive-pulmonary-disease-(copd). Accessed: Jul. 31, 2024.

[2] A. S. Buist *et al.*, "International variation in the prevalence of COPD (the BOLD Study): A population-based prevalence study," *Lancet*, vol. 370, no. 9589, pp. 741–750, 2007.

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

[4] M. Kubota *et al.*, "Reference values for spirometry, including vital capacity, in Japanese adults calculated with the LMS method and compared with previous values," *Respir. Investig.*, vol. 52, no. 4, pp. 242–250, 2014.

[5] H. W. Lee *et al.*, "Spirometric interpretation and clinical relevance according to different reference equations," *J. Korean Med. Sci.*, vol. 39, no. 4, p. e20, 2024.

[6] J. K. Choi, D. Y. Paek, and J. O. Lee, "Normal predictive values of spirometry in Korean population," *Tuberc. Respir. Dis.*, vol. 58, no. 3, pp. 230–242, 2005.

[7] J. S. Reynolds *et al.*, "Classification of voluntary cough airflow patterns for prediction of abnormal spirometry," *IEEE J. Biomed. Health Inform.*, vol. 20, no. 3, pp. 963–969, 2015.

[8] S. Bhattacharjee, B. Saha, P. Bhattacharyya, and S. Saha, "Classification of obstructive and non-obstructive pulmonary diseases on the basis of spirometry using machine learning techniques," *J. Comput. Sci.*, vol. 63, p. 101768, 2022.

[9] B. H. Culver *et al.*, "Recommendations for a standardized pulmonary function report. An official American Thoracic Society technical statement," *Am. J. Respir. Crit. Care Med.*, vol. 196, no. 11, pp. 1463–1472, 2017.

[10] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.

[11] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic minority over-sampling technique," *J. Artif. Intell. Res.*, vol. 16, pp. 321–357, 2002.

[12] S. Stanojevic *et al.*, "ERS/ATS technical standard on interpretive strategies for routine lung function tests," *Eur. Respir. J.*, vol. 60, no. 1, p. 2101499, 2022.

[13] D. P. Kingma and M. Welling, "Auto-encoding variational Bayes," *arXiv preprint arXiv:1312.6114*, 2013.

[14] T. DeVries and G. W. Taylor, "Dataset augmentation in feature space," *arXiv preprint arXiv:1702.05538*, 2017.

[15] J. Hu, L. Shen, and G. Sun, "Squeeze-and-excitation networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2018.

[16] H. Si *et al.*, "A dual-branch model integrating CNN and Swin Transformer for efficient apple leaf disease classification," *Agriculture*, vol. 14, no. 1, p. 142, 2024.

[17] C. Zhang, L. Wang, G. Wei, Z. Kong, and M. Qiu, "A dual-branch and dual attention transformer and CNN hybrid network for ultrasound image segmentation," *Front. Physiol.*, vol. 15, p. 1432987, 2024.

[18] K. Rana, P. Goyal, and G. Sharma, "Dual-branch convolutional neural network for robust camera model identification," *Expert Syst. Appl.*, vol. 238, p. 121828, 2024.

[19] A. Paszke *et al.*, "PyTorch: An imperative style, high-performance deep learning library," in *Advances in Neural Information Processing Systems 32*, 2019.