MDPI

*Data Descriptor*

# Datasets of Simulated Exhaled Aerosol Images from Normal and Diseased Lungs with Multi-Level Similarities for Neural Network Training/Testing and Continuous Learning

Mohamed Talaat [1], Xiuhua Si [2] and Jinxiang Xi [1,*]

1   Department of Biomedical Engineering, University of Massachusetts, Lowell, MA 01854, USA;
    mohamed_talaat@student.uml.edu
2   Department of Mechanical Engineering, California Baptist University, Riverside, CA 92504, USA;
    asi@calbaptist.edu
*   Correspondence: jinxiang_xi@uml.edu; Tel.: +1-978-934-3259

**Abstract:** Although exhaled aerosols and their patterns may seem chaotic in appearance, they inherently contain information related to the underlying respiratory physiology and anatomy. This study presented a multi-level database of simulated exhaled aerosol images from both normal and diseased lungs. An anatomically accurate mouth-lung geometry extending to G9 was modified to model two stages of obstructions in small airways and physiology-based simulations were utilized to capture the fluid-particle dynamics and exhaled aerosol images from varying breath tests. The dataset was designed to test two performance metrics of convolutional neural network (CNN) models when used for transfer learning: interpolation and extrapolation. To this aim, three testing datasets with decreasing image similarities were developed (i.e., level 1, inbox, and outbox). Four network models (AlexNet, ResNet-50, MobileNet, and EfficientNet) were tested and the performances of all models decreased for the outbox test images, which were outside the design space. The effect of continuous learning was also assessed for each model by adding new images into the training dataset and the newly trained network was tested at multiple levels. Among the four network models, ResNet-50 excelled in performance in both multi-level testing and continuous learning, the latter of which enhanced the accuracy of the most challenging classification task (i.e., 3-class with outbox test images) from 60.65% to 98.92%. The datasets can serve as a benchmark training/testing database for validating existent CNN models or quantifying the performance metrics of new CNN models.

**Dataset:** The dataset is submitted and will be published as a supplement to this paper in the journal Data, this field will be filled by the editors of the journal. CFD Lung Diagnosis | Kaggle.

**Dataset License:** license under (CC0, CC-BY, CC-BY-SA, CC-BYaeor-NC, etc.)

## 1. Summary

Several modalities exist in the diagnosis of obstructive lung diseases, including X-ray, low-dose CT scanning, and more recently, the exhaled breath test. X-ray images are a common diagnostic tool used to identify various respiratory conditions. They can reveal certain patterns that can indicate whether a person has a healthy lung or is experiencing certain types of respiratory illness or infection [1]. Low-dose (LD) spiral CT scanning has been recommended in recent years to screen lung cancers in heavy smokers who are at high risk. It checks the signs of neoplasm (i.e., shape, size, and location of abnormal airways) in the lungs to detect carcinogenicity before symptoms appear in patients [2–4]. Increasing efforts have been undertaken to diagnose lung disease using chest X-ray and CT images

with machine learning or a convolutional neural network (CNN), which have shown great potential in achieving automatic diagnosis with satisfactory accuracy [5–8]. However, both X-ray and CT scanning present patients with radiological risks. In addition, images from both modalities are two-dimensional and thus can miss local nodules. Expiratory airflow and particles hold a wealth of information regarding lung health, serving as indicators of either cell metabolisms or airway structure [9]. Exhaled breath in liquid or gaseous phases has been analyzed for physicochemical properties to detect abnormalities in the lungs [10]. In contrast to the high availability of chest X-rays and CT scans, exhaled aerosol images are still not available in the literature.

The dataset in this study was designed to test the hypothesis that the exhaled aerosol distributions could be correlated to the underlying lung structures and associated airway abnormalities [11]. In doing so, breath tests were simulated that inhaled and exhaled an aerosol bolus under varying breathing conditions, which was intended to train and validate convolutional neural network (CNN) models. To develop an AI-aided lung diagnostic system, the accuracy and robustness of four CNN models in classifying the test images were evaluated. Specifically, the CNN models' capacities for interpolation and extrapolation were quantitively evaluated. To this end, three levels of test datasets with decreasing image similarities were developed. The first level was randomly taken from the same source images (i.e., 90% for training, 10% for testing, but with no overlapping). The second level included new images with operating conditions that were different from the original source, but were still within the design space (i.e., termed 'Inbox', with decreased similarities to the original source images). The third level included images with operating conditions outside of the design space (i.e., termed 'Outbox', having the least similarities to the source). The images belonging to Level 1, Level 2 (Inbox), and Level 3 (Outbox) can be downloaded from [12].

The dataset was generated by means of physiology-based modeling and simulations. An anatomically accurate mouth-lung geometry constructed in previous studies [13] was modified to represent the diseased lung geometries with bronchiolar constrictions in small airways at G7–9. The current radiological imaging techniques, such as low-dose CT scanning (LDCT), detect lung nodules only 3 mm and larger [14–17]. To test the diseases in small airways with structural remodeling of 1 mm, which is smaller than the current LDCT detection limit, is highly desirable for the purpose of the early detection of lung tumors. Considering that the performance of a trained CNN model generally depends on the quality of the training data, great efforts were undertaken to ensure the simulated aerosol images were (1) physiologically realistic and (2) numerically consistent. One breath test contained an inhalation of aerosol bolus and a collection of exhaled aerosols at the mouth opening during exhalation. To ensure numerical consistency, control parameters were precisely prescribed for all airflow simulations and particle tracking. As a result, the compounding effects of numerical errors could be avoided.

The dataset usage was demonstrated using four popular CNN models, including AlexNet, ResNet, MobileNet, and EfficientNet, either for their proven performance superiority, or their simplicity and efficiency [18–20]. The detailed classification matrices were presented in case replication of the model training/testing was attempted on the image dataset. These models can be downloaded from the link [12].

This dataset offers several unique benefits. First, it is a multi-level image dataset. It can be useful not only to medical practitioners to develop AI-assisted systems for lung diagnostics and staging, but also to the AI community who want to test the performance of CNN models on test data both within and outside the design space. Secondly, the images in this dataset were generated by systematically varying the operating variables, which preserves the inherent features associated with different variables and their evolution. This provides a clean and systematic dataset to test a model's ability to extract features due to individual or multi-factor interactions. Thirdly, physiology-based modeling and simulation provide detailed information on airflow and particle dynamics, which can be further examined to gain a better understanding of the CNN model's performance.

While current CNN models can achieve impressively high accuracy rates, identifying the underlying reasons for their performance, such as why one model performs better than another or what key features facilitate high performance, remains challenging. In this case, the simulated image dataset provides a unique advantage in shedding light on the deep architectures of CNN models and how they relate to the problem of interest, namely lung health and disease.

## 2. Data Description

### 2.1. Data Achtiecture

The dataset for exhaled aerosol images considered four variables: the health of small airways (normal: D0, disease stage 1: D1, disease stage 2: D2), throat narrowing (Th0, Th1, Th2, Th3, with decreasing glottal apertures), tracer particle size (dp [μm]), and flow rate (Q [L/min]), as shown in Figure 1. The design space for the aerosol-breath tests was 0.5–10 μm for dp and 10–19 L/min for Q. Three testing datasets with decreasing image similarities were designed (Level 1, Inbox, Outbox) to test the CNN model's capacity for validation, interpolation, and extrapolation, respectively. Inbox images were from operating conditions within the design space but had not been seen by the model before. Outbox images were from operating conditions outside the design space.
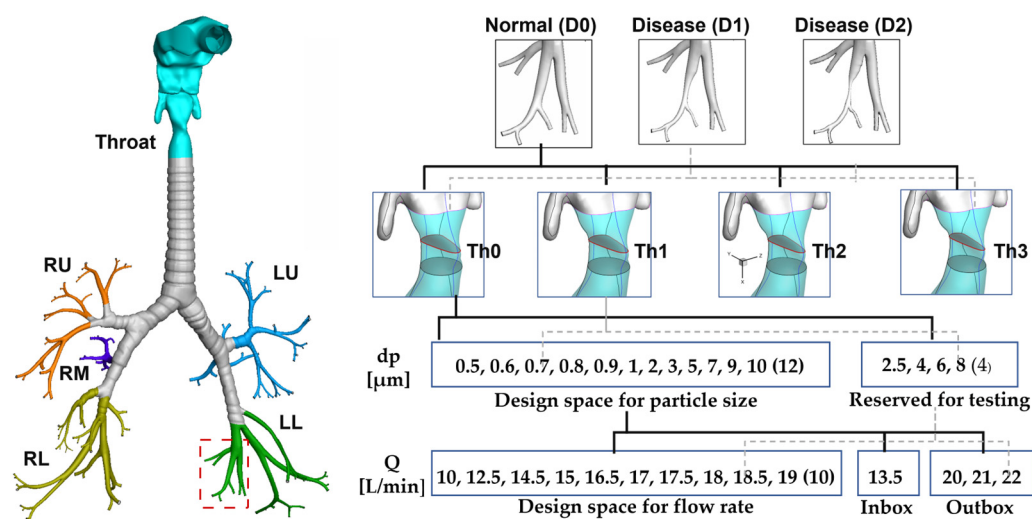


**Figure 1.** Dataset architecture for exhaled aerosol images with four layers of variables: small airway health (D0, D1, D2), throat narrowing, tracer particle size (dp [μm]), and flow rate (Q [L/min}). The design space for the aerosol-breath tests was 0.5–10 μm for dp and 10–19 L/min for Q. Three testing datasets with decreasing image similarities were designed (Level 1, Inbox, Outbox) to test the CNN model's capacity of validation, interpolation, and extrapolation, respectively. Inbox images were from operating conditions within the design space but had never been seen by the model before. Outbox images were from operating conditions outside the design space.

In the baseline dataset (Base), the throat was kept constant in the three airway models (D0, D1, D2). The recommended range (or design space) for the breath tests included particle sizes ranging between 0.5–10 μm and flow rates ranging between 10–19 L/min. Specifically, the particle sizes considered were 0.5, 0.6, 0.7, 0.8, 0.9, 1, 2, 3, 5, 7, 9, and 10 μm, while the respiratory flow rates included 10, 12.5, 14.5, 15, 16.5, 17, 17.5, 18, 18.5, 18.5, and 19 L/min. In total, the baseline dataset contained 1080 images, of which 90% were used for training and 10% for testing in Round 1 (as shown in Table 1). The exhaled aerosol images were organized in different folders, following Table 1, and can be downloaded from the link [12]. The four classification models (AlexNet, ResNet-50, MobileNet, and EfficientNet. AlexNet) are also provided in the same link. Interested readers can train and test the models using the provided dataset and compare their results with ours, which are provided as Supplemental Files.

**Table 1.** Three-round training/testing procedures to evaluate the model capacity of interpolation, extrapolation, and continuous learning. These procedures were tested in four models (AlexNet, ResNet-50, MobileNet, and EfficientNet) for both two-class (normal vs. disease) and three-class (D0 vs. D1, D2) classifications.

| | Training | Testing | | |
|---|---|---|---|---|
| | | **Level 1** | **Level 2** | **Level 3** |
| **Round 1** | 90% Base | 10% Base | Inbox | Outbox |
| **Round 2**: (plus 90% Base) | Th1, Th2, Th3 | 10% Base | Inbox | Outbox |
| **Round 3** (plus 90% Base, and Th1, Th2, Th3) | 25% Outbox | 10% Base | Inbox | Outbox |
| | 50% Outbox | 10% Base | Inbox | Outbox |
| | 75% Outbox | 10% Base | Inbox | Outbox |

The inbox database (Level 2) consisted of 535 exhaled aerosol images with either particle sizes or flow rates that were not present in the baseline database. However, all these particle sizes and flow rates fell within the recommended design space. Particularly, the images generated with particle sizes of 2.5, 4, 6, and 8 μm were exclusively reserved for testing purposes (Figure 1).

The outbox database contained 649 images and represented scenarios outside of the design space. This included different flow rates (i.e., 20, 21, 22 L/min, Outbox) and various throat openings (Figure 1).

The architecture of the dataset is shown in Figure 2 and the dataset can be downloaded from [12]. Under the head folder "Dataset", there are three folders: Round-1, Round-2, and Round-3, which represent three rounds of model training/testing, or continuous learning. In each round, there are four subfolders, with one Training dataset and three testing datasets (Level 1, Level 2 which is also called "Inbox", and Level 3 which is also called "Outbox"). Note that the images in the three testing datasets share decreasing levels of similarity to those in the Training dataset from Level 1 to Level 3, thus representing increasing challenges in model performances.
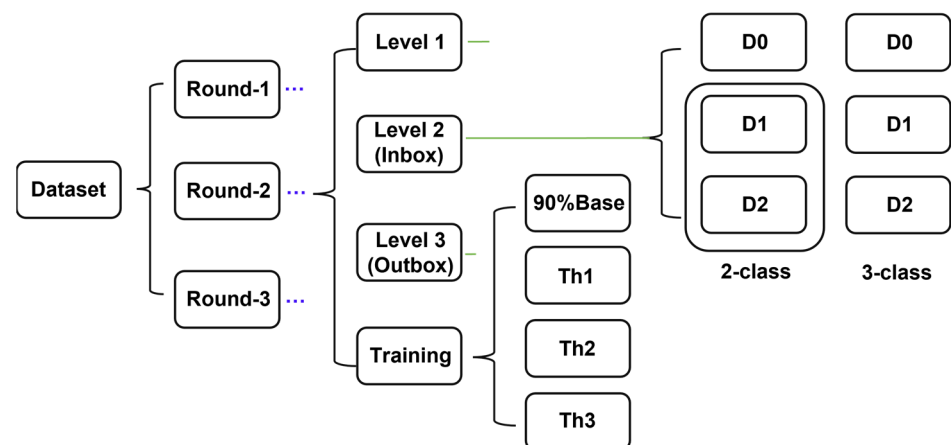


**Figure 2.** Dataset architecture.

The training dataset contains an increasing number of images from round 1 to round 3 by including new images (i.e., continuous learning). In Figure 2, the training dataset in round 2 contains 90% base plus Th1, Th2, and Th3, which is also shown in Table 1.

Each testing database contains three subfolders: D0, D1, and D2, which can be used for both 2-class (normal D0 vs. disease D1 and D2) and 3-class (D0 vs. D1 vs. D2) classifications (Figure 2). From a clinical point of view, the 2-class classification can differentiate healthy subjects from patients with lung diseases, and thus is suitable for screening purposes.

Likewise, the 3-class classification not only differentiates the normal from diseased, but also provides the level of lung obstructions, thus being suitable for disease staging. For those who are interested in testing the CNN model performance, the 2-class and 3-class classification tasks represent different levels of challenges, and thus can provide an additional measure of model performance.

### 2.2. Disease-Induced Disturbances to Particle Distributions

Sample images of the exhaled aerosols predicted from various breath test conditions are shown in Figure 3 from normal (D0) and diseased lungs (D1, D2), respectively. The effects of the particle size (0.5, 1.0, and 5 μm), flow rate (10, 15, 20 L/min), and throat narrowing (Th0 vs. Th3) are shown in each category.
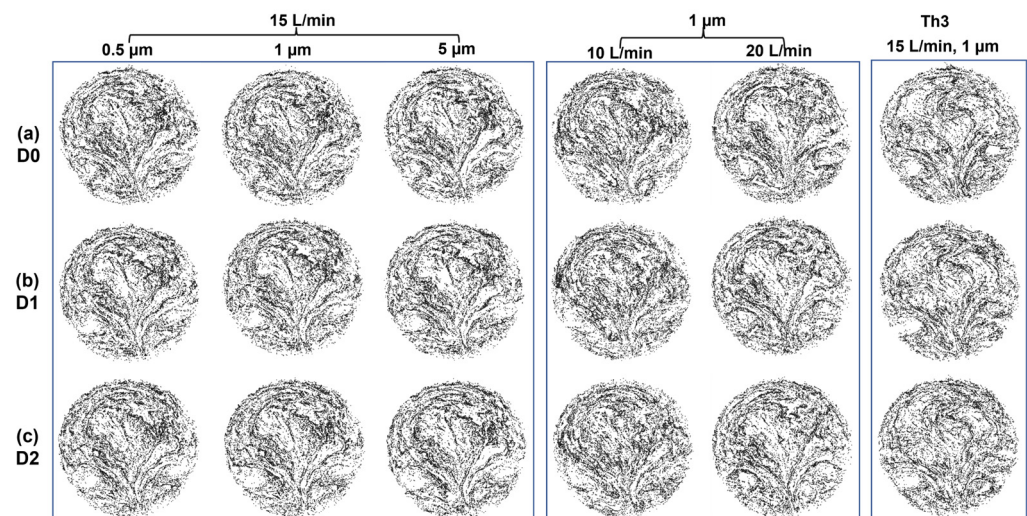


**Figure 3.** Sample images of exhaled aerosols predicted from normal (**a**) and diseased lungs with mild (D1) and severe (D2) constrictions (**b**,**c**), with varying particle sizes, flow rates, and throat openings.

Due to the complex and chaotic patterns, distinguishing between normal lung images and images of diseased lungs is challenging at first glance. One contributing factor is the more pronounced differences observed within the same category among images at different flow rates (10 vs. 15 vs. 20 L/min). Additionally, subtle differences can be observed among images with different particle sizes, suggesting that particle size has a smaller impact compared to the flow rate. At 15 L/min, the exhaled aerosol image associated with a throat narrowing, Th3 (last column, Figure 3), shares more similarities with its counterpart without throat narrowing (second column, Figure 3) than with the images under different flow rates (i.e., 10 and 20 L/min). This indicates that the throat narrowing exerts a smaller impact on particle dynamics than the flow rate.

To gain insights into the correlation between exhaled aerosol distributions and lung structures, one sample image (D0, 15L/min, and 1 μm, first row and second column) in Figure 3 was further examined. Figure 4 shows the distribution of aerosols exhaled from the region of interest (ROI) at G7–9 of the left lower lobe, as well as from the five lobes. Note that in Figure 4, the LL region does not include the ROI. Distinct patterns in the distribution of aerosols are observed in the ROI and five lobes, characterized by the presence of concentrated particles arranged in stripes or vortices. This indicates a strong influence from respiratory flows. Different from the quick response of airflow dynamics, the trajectory of a particle is integrative in nature, depending not only on local airflows but also on the motion history of the particle. There was an apparent asymmetry between the left and right lungs, which was observed in the comparison of LL vs. RL or LU vs. RU. The right panel of Figure 4 reveals an even more intricate pattern when all particles were combined.
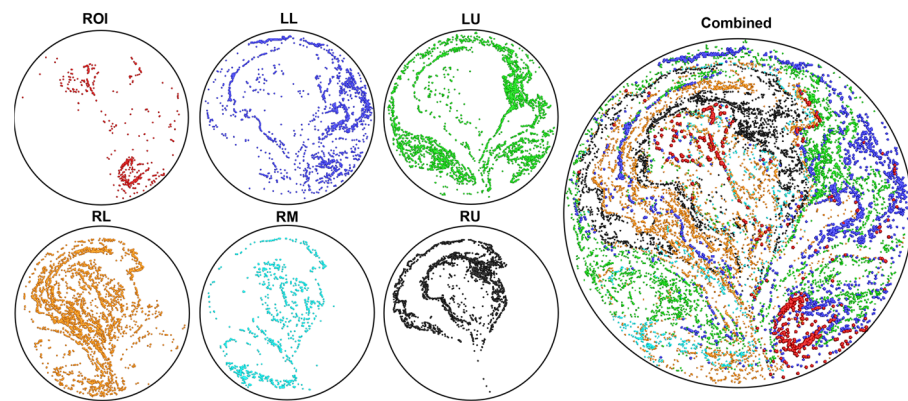
**Figure 4.** Distributions of aerosols exhaled form the region of interest (ROI) and five lobes of the normal lung D0 for a particle size of 1 μm and a flow rate of 15 L/min.

To emphasize the aerosol distribution from the ROI bronchioles, exhaled particles from the LR, RU, RM, and RL were grouped together in green color due to their peripheral location relative to the ROI (Figure 5). On the other hand, aerosols from the LL (excluding ROI) remained as a separate group (blue), considering their proximity to the ROI, which presumably made them more sensitive to disturbances from the ROI. Exhaled particle distributions under other test conditions (0.5 μm, 5 μm, 10 L/min, 20 L/min, throat narrowing) are also presented for both normal (D0, Figure 5a) and diseased conditions (D1 and D2 in Figure 5b,c), respectively. By doing so, the evolution of the ROI-related particle distributions from D0 to D1 and D2 can be compared, which not only decreased in extent but also varied in morphology. Particle profiles from the blue group representing the LL lobe showed noticeable variations in their distributions among D0-2 for any given test condition, indicating a significant effect of the disease conditions on the exhaled aerosol patterns. On the other hand, the particle distributions from the green group representing the peripheral lobes remained relatively similar among D0–2 for any given test condition, suggesting a lesser impact of the disease on these regions.
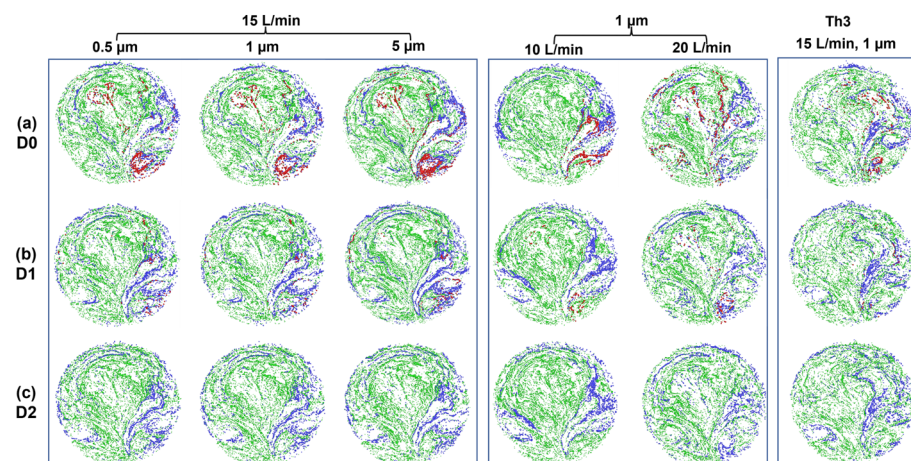


**Figure 5.** Distributions of aerosols exhaled form ROI (red), LL lobe (blue), and peripheral lobes (LR, RU, RM, and RL, green) under varying test conditions for (**a**) normal (D0), (**b**) diseased lung at stage 1 (D1), and (**c**) diseased lung at stage 2 (D2).

### 2.3. Results for Multi-Round Training and Multi-Level Testing

The exhaled aerosol images were used to train and test four CNN models to evaluate their capacity for learning disease-defining features and classifying test samples with varying dissimilarity levels to the training dataset. These models included AlexNet, ResNet-50, MobileNet, and EfficientNet. AlexNet and ResNet-50 were chosen for their wide usage

and proven performance [21,22], while MobileNet and EfficientNet were selected for their simple network architecture and low computational requirements [23–26]. After training, each model was validated by testing a dataset that consisted of a random 10% baseline dataset, with the remaining 90% used as the Round 1 training dataset. To evaluate the model's interpolation capacity, the trained model was tested on images from different test conditions within the design space, known as Inbox images. To test the extrapolation capacity, the trained model was tested on images with operating conditions outside the design space (i.e., Outbox images). This evaluation aimed to determine how well the models could generalize to new and unseen conditions, which is crucial for their applicability in real-world scenarios. The Outbox images were generated by simulating operating conditions beyond the ranges used in training, including different flow rates and particle sizes.

In clinical conditions, new quality data will become available, which can be incorporated into the training dataset to improve the model's performance. Each model underwent two additional training rounds, in addition to the first round. In this study, the breath test variables included throat narrowing, flow rate, and particle size. In Round 2, 720 images with three throat narrowing variants (Th1, Th2, Th3) were added to the training dataset, followed by re-training of the four models. The models were then tested on three datasets with increasing dissimilarities: Level 1, Inbox, and Outbox.

In Round 3, new images with Outbox flow rates (20, 21, 22 L/min) were added to the training dataset to enhance the model performance. To determine the minimum number of images needed to achieve a significant improvement, different proportions of the Outbox images (25%, 50%, and 75%) were added to the training dataset. The newly trained models were then tested on the Level 1, Inbox, and Outbox test datasets to evaluate their performance.

Table 2 compares the 2-class and 3-class classification accuracies among four CNN models (AlexNet, ResNet-50, MobileNet, and EfficientNet) that were tested on samples with decreasing similarities (Level 1, Inbox, Outbox). In doing so, the model's capacity for verification, interpolation, and extrapolation could be assessed. Among the four models, ResNet-50 consistently gave superior performances, especially in classifying external (Outbox) test images, for the purposes of both diagnosis (2-class classification) and disease staging (3-class classification). As expected, training models, by exposing them to more external images (i.e., continuous learning), progressively increased the models' classification accuracy. More details of the model performances are also provided in Tables S1–S5 (Supplemental Files) in terms of AUC (area under the curve), specificity, sensitivity, and precision. More details (step-by-step results of the training and testing) are given in the file "Data_Images_CNN_MultiRound_MultiLevel.pdf", which be downloaded from [12] with the dataset and source code. These results can be used for validating existing CNN models or training/testing new CNN models.

### 2.4. Learned Features from CNN Models

To develop a computer-aided diagnostic system for obstructive lung diseases, it is crucial to understand the features that have been learned by the CNN models and utilize them to differentiate health and disease. In our previous studies that employed machine learning algorithms for disease classification, great efforts were taken to identify and characterize disease-specific features [27,28]. One advantage of CNN models is their ability to automatically learn features at different layers from the training dataset. However, the learned features are often highly abstract and challenging to correlate with the original images. Figure 6 shows the learned features at varying layers by the AlexNet after the first round (i.e., Round 1) of training. Note that AlexNet has five convolutional layers and three fully connected layers. From Figure 6, it can be observed that the features become increasingly abstract and less recognizable as the layers go deeper, the filters become more blurred and have fewer details. It is noteworthy that the number of filters (or features) generally increases at deeper layers. The output layer (FC8) consists of 1000 neurons

connected to 4096 outputs from FC7 and calculates a classification value for a given test image [29].

**Table 2.** Comparison of 2-class and 3-class classification accuracies among four CNN models (AlexNet, ResNet-50, MobileNet, and EfficientNet) tested on datasets with decreasing similarities (Level 1, Inbox, Outbox) to evaluate the model's capacity for verification, interpolation, and extrapolation, respectively. Each model was trained for three rounds, with round 3 adding 25%, 50%, and 75% Outbox images.

| Network | Accuracy (%) | 2—Classes | | | 3—Classes | | |
|---|---|---|---|---|---|---|---|
| | | Level 1 | Inbox | Outbox | Level 1 | Inbox | Outbox |
| AlexNet | Round 1 | 100 | 98.88 | 58.49 | 99.24 | 99.24 | 47.07 |
| | Round 2 | 98.6 | 98.69 | 68.06 | 99.3 | 99.3 | 54.01 |
| | Round 3 | | | | | | |
| | 25% Outbox | 98.73 | 98.13 | 97.38 | 98.09 | 98.09 | 93.06 |
| | 50% Outbox | 99.36 | 97.75 | 99.07 | 92.99 | 92.99 | 94.14 |
| | 75% Outbox | 99.36 | 98.5 | 99.69 | 100 | 100 | 98.46 |
| ResNet–50 | Round 1 | 100 | 99.63 | 65.12 | 99.24 | 82.77 | 60.65 |
| | Round 2 | 99.3 | 99.44 | 65.9 | 100 | 91.01 | 58.18 |
| | Round 3 | | | | | | |
| | 25% Outbox | 99.36 | 99.63 | 97.84 | 97.45 | 84.83 | 95.06 |
| | 50% Outbox | 100 | 97.94 | 99.85 | 99.36 | 88.2 | 97.38 |
| | 75% Outbox | 99.36 | 98.69 | 98.77 | 98.73 | 84.46 | 98.92 |
| MobileNet | Round 1 | 99.24 | 96.63 | 60.19 | 97.73 | 73.4 | 45.8 |
| | Round 2 | 97.9 | 95.88 | 71.14 | 95.1 | 77.34 | 54.48 |
| | Round 3 | | | | | | |
| | 25% Outbox | 98.1 | 97.57 | 94.14 | 94.27 | 78.28 | 87.35 |
| | 50% Outbox | 98.73 | 95.69 | 98.15 | 99.24 | 76.78 | 95.06 |
| | 75% Outbox | 100 | 95.51 | 98.3 | 96.18 | 75.09 | 96.91 |
| EfficientNet | Round 1 | 96.97 | 91.2 | 61.27 | 90.15 | 70.22 | 41.98 |
| | Round 2 | 97.9 | 93.07 | 61.88 | 93 | 68.73 | 56.33 |
| | Round 3 | | | | | | |
| | 25% Outbox | 93.63 | 90.64 | 87.5 | 84.18 | 68.73 | 80.09 |
| | 50% Outbox | 96.18 | 91.76 | 95.68 | 89.81 | 68.35 | 90.28 |
| | 75% Outbox | 98.1 | 89.14 | 97.99 | 89.1 | 65.92 | 95.06 |
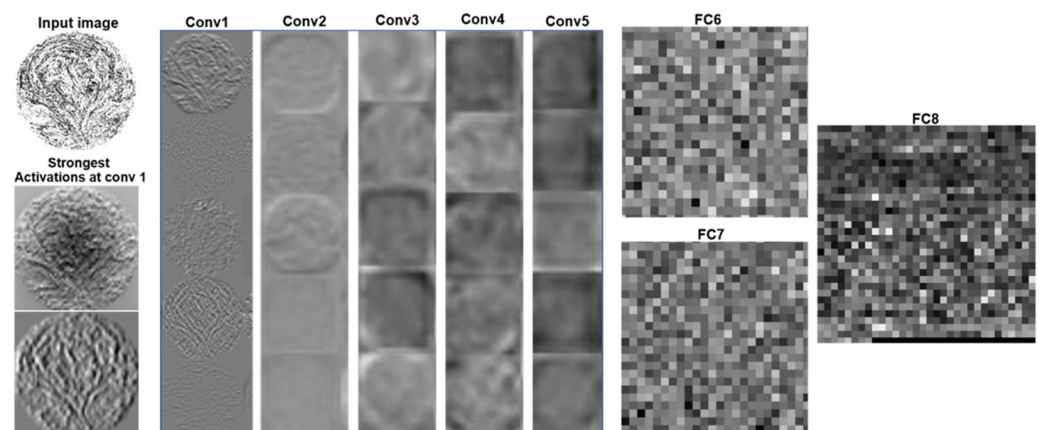


**Figure 6.** Learned features at varying convolutional layers by AlexNet, which has five convolutional (Conv) layers and three fully connected (FC) layers.

Figure 7 compares the learned features by ResNet-50 from normal (D0) and disease (D1 and D2) images at the first (Conv1) and third (Conv3) convolutional layers. Similar to AlexNet, the level of feature details decreased from Conv1 to Conv3. For a given

convolutional layer, the features seemingly differed, with each capturing one specific aspect of the image. At Conv3, the pixels appeared to be vertical in the right panel vs. horizontal in the left panel. It should be noted that correlating the learned features to visually differentiate the diseases may be challenging.
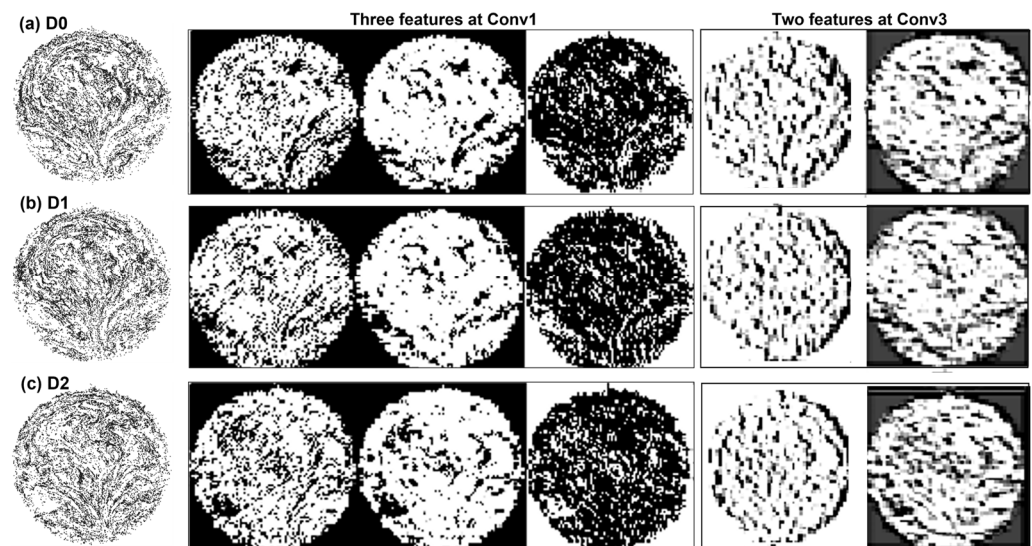


**Figure 7.** Comparison of learned features by ResNet-50 at the first and third conventional layer from images belonging to the category of (**a**) normal (D0), (**b**) disease stage 1 (D1), and (**c**) disease stage 2 (D2).

## 3. Methods

### 3.1. Normal and Diseased Lungs

Physiology-based modeling and simulations were used to generate images of exhaled aerosols from normal and diseased airways under varying breathing conditions (Figure 8). The normal airway model was developed by Xi et al. [30,31], which extends from the mouth up to the ninth generation (G9) lung bifurcations and retains 125 bronchial outlets (Figure 1). In this study, airway obstruction occurred at G7–9 bronchioles, the diameters of which were less than 1 mm (i.e., small airways). Therefore, the obstruction was also smaller than 1 mm in size, which was below the smallest nodule size to be detected using X-rays or CT scanning (3–4 mm) [14]. Note that the model-generated images could be less complex than real-life images and might be less challenging to differentiate. Thus, we hypothesized that by considering the airway lesions that were below the detection limit of current radiological imaging technologies, it was anticipated that the proposed computer-aided diagnostic system could achieve sufficiently high diagnostic accuracy when applied in clinical settings.
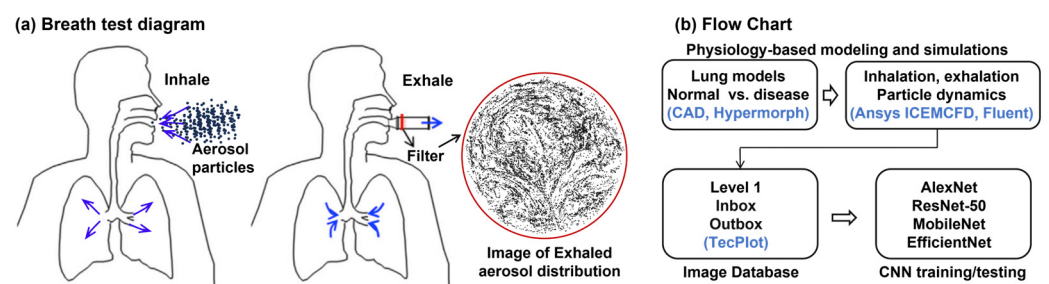


**Figure 8.** Breath tests: (**a**) test diagram with inhalation and exhalation, and (**b**) flow chart of image generation and CNN training/testing.

The morphology of the normal mouth-lung model (D0) was modified to generate two diseased models (D1, D2) in the left lower lobe, as indicated by the red dashed rectangle in Figure 1. To achieve this, the software Hypermorph (Troy, MI) was used to shrink the bronchioles at G7–9 twice, resulting in the two diseased models D1 and D2 (Figure 1). Additionally, the normal glottal aperture, which is the opening between the vocal cords, was progressively reduced by 1mm, 2mm, and 3mm to generate three constricted throats (Th1, Th2, Th3; Figure 1). The normal and modified airway models were then meshed using Ansys ICEMCFD to perform fluid-particle simulations (Figure 1).

### 3.2. Computational Mesh and Grid Independent Study

The computational mesh in the mouth-lung airway geometries was created using ANSYS ICEMCFD. To accurately capture the flow variation near the walls, body-fitted meshes were generated with a five-layer prism mesh. A grid-independent study was conducted by varying the mesh densities from coarse to ultrafine, and the results were achieved at 4.8 million tetrahedral cells with five layers of prismatic cells and a near-wall cell height of 0.05 mm [30,32,33].

### 3.3. Boundary Conditions

The breath tests were envisioned as follows: the subject adopts a sitting position, inhales a bolus of aerosols, and then exhales, as illustrated in Figure 8a [27]. The exhaled aerosols are collected on a filter at the mouth opening. Both the flow rate and particle size can be controlled. Ambient pressure is prescribed at the mouth opening. Negative/positive pressures are specified to generate a prescribed inhalation/exhalation flow rate. For each breath test, a bolus of 100,000 tracer particles is released into the mouth during inhalation. To clarify, during exhalation, the undeposited particles that are recorded at the bronchiolar outlets are released and carried out with the exhaled flow. The velocities of these particles are reversed from the inhalation phase and tracked until they exit the mouth opening and are collected on a filter. The exiting particles are then recorded at the mouth opening and plotted in TecPlot to generate the exhaled aerosol image, as shown in Figures 3–5. The distribution pattern of these particles collectively forms the exhaled aerosol image, which can be used in subsequent CNN training and/or testing (Figure 8b).

### 3.4. Numerical Methods

ANSYS Fluent (Canonsburg, PA, USA) was utilized to simulate the inhalation/exhalation flows and generate the exhaled aerosol images. The turbulent $k$-$\omega$ model was used to simulate the inhalation and exhalation airflows. Particle motion was tracked with a Lagrangian discrete phase model (DPM). Particles were deposited on the airway wall upon contact. Considering the dilute nature of the particles, one-way coupling (i.e., flow to particles) was assumed during the particle tracking. Specific tracking parameters are listed below. The drag law was Stokes-Cunnigham. The maximal tracking number of steps was 5,000,000 with a step length scale of 5. The higher-order tracking scheme was Runge-Kutta and the lower-order scheme was implicit. The SIMPLE scheme was utilized for the pressure-velocity coupling. Regarding spatial discretization, the Green-Gauss cell-based scheme was used for the gradient calculation, a standard scheme for pressure, and a second-order upwind scheme for momentum. All simulated breath tests followed these specifications to ensure results were numerically consistent.

User-defined MATLAB codes were developed to generate particles at the mouth inlet and reverse the particle velocities at the bronchiolar outlets. One breath test required one inhalation, exhalation, and particle tracking after both phases, which required approximately 4 h, 4 h, and 10–90 min depending on the particle size, respectively, in an AMD Ryzen 3960X 24-Core workstation with 3.79 GHz processors, 256G RAM, and 8G GPU. For a total of 2745 images used in this study (eleven flow rates, four geometrical models), a cumulation of 3200 computational hours or so was used.

### 4. User Notes

As an exploratory study, the exhaled aerosol images used in this study were not from human subjects but were generated using computational fluid-particle simulations in physiologically realistic airway models. Currently, no breath tests have been conducted on humans, and thus there are no in vivo images for training/testing. Future in vitro breath tests can be conducted in 3D-printed mouth-lung replicas with normal and diseased airways. Regarding the validity of simulation-generated images, previous studies have demonstrated that physiology-based simulations could sufficiently reproduce in vivo conditions [34].

To the authors' best knowledge, the proposed dataset is the first to use exhaled aerosol images from normal (D0) and diseased (D1 and D2) lungs. The dataset is also unique in providing testing image datasets with decreasing similarities to images in the training datasets. This feature provides an opportunity to train and test CNN models at multiple levels; for instance, for validation, interpolation, and extrapolation. The dataset architecture allows interested readers to evaluate the continuous learning capacity of the CNN models by incrementally adding outbox test images to the training dataset. It also allows for evaluating model performance on 2-class and 3-cass classifications. This database may be of interest to the AI community, who are interested in benchmark testing current and new CNN models, physicians working with the automatic diagnosis of obstructive lung diseases, as well as researchers in the field of respiratory dynamics. In addition to the well-structured dataset, the source code for CNN models, detailed training/testing results, and a result summary were also provided, which can serve as an easy-to-start educational tutorial for AI beginners.

## References

1. Hu, J.; Zhang, C.; Zhou, K.; Gao, S. Chest X-Ray diagnostic quality assessment: How much Is pixel-wise supervision needed? *IEEE Trans. Med. Imaging* **2022**, *41*, 1711–1723. [CrossRef] [PubMed]
2. Hoffman, R.M.; Atallah, R.P.; Struble, R.D.; Badgett, R.G. Lung cancer screening with low-dose CT: A meta-analysis. *J. Gen. Intern. Med.* **2020**, *35*, 3015–3025. [CrossRef] [PubMed]
3. Tanino, A.; Kawamura, T.; Hamaguchi, M.; Tanino, R.; Nakao, M.; Hotta, T.; Tsubata, Y.; Hamaguchi, S.; Isobe, T. A novel model-based questionnaire based on low-dose CT screening data for chronic obstructive pulmonary disease diagnosis in Shimane, Japan. *Int. J. Chron. Obstruct. Pulmon. Dis.* **2021**, *16*, 1823–1833. [CrossRef] [PubMed]
4. Duffy, S.W.; Field, J.K. Mortality reduction with low-dose CT screening for lung cancer. *N. Engl. J. Med.* **2020**, *382*, 572–573. [CrossRef] [PubMed]

5. Das, N.; Topalovic, M.; Janssens, W. Artificial intelligence in diagnosis of obstructive lung disease: Current status and future potential. *Curr. Opin. Pulm. Med.* **2018**, *24*, 117–123. [CrossRef]

6. Kanjanasurat, I.; Tenghongsakul, K.; Purahong, B.; Lasakul, A. CNN-RNN network integration for the diagnosis of COVID-19 using Chest X-ray and CT images. *Sensors* **2023**, *23*, 1356. [CrossRef]

7. Heuvelmans, M.A.; van Ooijen, P.M.A.; Ather, S.; Silva, C.F.; Han, D.; Heussel, C.P.; Hickes, W.; Kauczor, H.U.; Novotny, P.; Peschl, H.; et al. Lung cancer prediction by Deep Learning to identify benign lung nodules. *Lung Cancer* **2021**, *154*, 1–4. [CrossRef]

8. Bharati, S.; Podder, P.; Mondal, M.R.H. Hybrid deep learning for detecting lung diseases from X-ray images. *Inform. Med. Unlocked* **2020**, *20*, 100391. [CrossRef]

9. Xi, J.; Si, X.A.; Kim, J.; Mckee, E.; Lin, E.-B. Exhaled aerosol pattern discloses lung structural abnormality: A sensitivity study using computational modeling and fractal analysis. *PLoS ONE* **2014**, *9*, e104682. [CrossRef]

10. Cavaleiro Rufo, J.; Paciência, I.; Mendes, F.C.; Farraia, M.; Rodolfo, A.; Silva, D.; de Oliveira Fernandes, E.; Delgado, L.; Moreira, A. Exhaled breath condensate volatilome allows sensitive diagnosis of persistent asthma. *Allergy* **2019**, *74*, 527–534. [CrossRef]

11. Talaat, M.; Xi, J.; Tan, K.; Si, X.A.; Xi, J. Convolutional neural network classification of exhaled aerosol images for diagnosis of obstructive respiratory diseases. *J. Nanotheranostics* **2023**, *4*, 228–247. [CrossRef]

12. Talaat, M.; Si, X.A.; Xi, J. CFD Lung Diagnosis|Kaggle. Available online: https://www.kaggle.com/datasets/mohamedtalaat92/cfd-lung-diagnosis. (accessed on 21 July 2023).

13. Lu, J.; Xi, J.; Langenderfer, J.E. Sensitivity analysis and uncertainty quantification in pulmonary drug delivery of orally inhaled pharmaceuticals. *J. Pharm. Sci.* **2017**, *106*, 3303–3315. [CrossRef]

14. Force, U.P.S.T. Screening for lung cancer: US preventive services task force recommendation statement. *JAMA* **2021**, *325*, 962–970.

15. Henschke, C.; McCauley, D.I.; Yankelevitz, D.F. Early lung cancer action project: Overall design and findings from baseline screening. *Lancet* **1999**, *354*, 99–105. [CrossRef]

16. Sone, S.; Li, F.; Yang, Z.G. Results of three-year mass screening programme for lung cancer using low-dose spiral computed tomography scanner. *Br. J. Cancer* **2001**, *84*, 25–32. [CrossRef]

17. Swensen, S. Screening for cancer with computed tomography. *BMJ* **2003**, *326*, 894–895. [CrossRef]

18. Gao, Y.; Cui, Y. Deep transfer learning for reducing health care disparities arising from biomedical data inequality. *Nat. Commun.* **2020**, *11*, 5131. [CrossRef]

19. Ayana, G.; Dese, K.; Choe, S.W. Transfer learning in breast cancer diagnoses via ultrasound imaging. *Cancers* **2021**, *13*, 738. [CrossRef]

20. Valverde, J.M.; Imani, V.; Abdollahzadeh, A.; De Feo, R.; Prakash, M.; Ciszek, R.; Tohka, J. Transfer learning in magnetic resonance brain imaging: A systematic review. *J. Imaging* **2021**, *7*, 66. [CrossRef]

21. Loey, M.; Manogaran, G.; Taha, M.H.N.; Khalifa, N.E.M. Fighting against COVID-19: A novel deep learning model based on YOLO-v2 with ResNet-50 for medical face mask detection. *Sustain. Cities Soc.* **2021**, *65*, 102600. [CrossRef]

22. Benali Amjoud, A.; Amrouch, M. Convolutional neural networks backbones for object detection. In *Image and Signal Processing*; Springer: Cham, Switzerland, 2020; Volume 12119.

23. Tan, M.; Le, Q. EfficientNet: Rethinking model scaling for convolutional neural networks. In Proceedings of the 36th ICML, Long Beach, CA, USA, 9–15 June 2019; pp. 6105–6114.

24. Tan, M.; Le, Q. Efficientnetv2: Smaller models and faster training. In Proceedings of the 38th ICML, Online, 18–24 July 2021; pp. 10096–10106.

25. Howard, A.G.; Zhu, M.; Chen, B.; Kalenichenko, D.; Wang, W.; Weyand, T.; Andreetto, M.; Adam, H. Mobilenets: Efficient convolutional neural networks for mobile vision applications. *arXiv* **2017**, arXiv:1704.04861.

26. Sandler, M.; Howard, A.; Zhu, M.; Zhmoginov, A.; Chen, L.-C. Mobilenetv2: Inverted residuals and linear bottlenecks. In Proceedings of the IEEE CVPR, Salt Lake City, UT, USA, 18–23 June 2018; pp. 4510–4520.

27. Si, X.A.; Xi, J. Deciphering exhaled aerosol fingerprints for early diagnosis and personalized therapeutics of obstructive respiratory diseases in small airways. *J. Nanotheranostics* **2021**, *2*, 94–117. [CrossRef]

28. Xi, J.; Zhao, W. Correlating exhaled aerosol images to small airway obstructive diseases: A study with dynamic mode decomposition and machine learning. *PLoS ONE* **2019**, *14*, e0211413. [CrossRef] [PubMed]

29. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. Imagenet classification with deep convolutional neural networks. *Commun. ACM* **2017**, *60*, 84–90. [CrossRef]

30. Xi, J.; Zhao, W.; Yuan, J.E.; Cao, B.; Zhao, L. Multi-resolution classification of exhaled aerosol images to detect obstructive lung diseases in small airways. *Comput. Biol. Med.* **2017**, *87*, 57–69. [CrossRef]

31. Xi, J.; Wang, Z.; Talaat, K.; Glide-Hurst, C.; Dong, H. Numerical study of dynamic glottis and tidal breathing on respiratory sounds in a human upper airway model. *Sleep Breath.* **2018**, *22*, 463–479. [CrossRef]

32. Talaat, M.; Si, X.A.; Dong, H.; Xi, J. Leveraging statistical shape modeling in computational respiratory dynamics: Nanomedicine delivery in remodeled airways. *Comput. Methods Programs Biomed.* **2021**, *204*, 106079. [CrossRef]

33. Xi, J.; Talaat, M.; Si, X.A.; Chandra, S. The application of statistical shape modeling for lung morphology in aerosol inhalation dosimetry. *J. Aerosol Sci.* **2021**, *151*, 105623. [CrossRef]
34. Xiao, Q.; Stewart, N.J.; Willmering, M.M.; Gunatilaka, C.C.; Thomen, R.P.; Schuh, A.; Krishnamoorthy, G.; Wang, H.; Amin, R.S.; Dumoulin, C.L.; et al. Human upper-airway respiratory airflow: In vivo comparison of computational fluid dynamics simulations and hyperpolarized 129Xe phase contrast MRI velocimetry. *PLoS ONE* **2021**, *16*, e0256460. [CrossRef]