JOURNAL OF AEROSOL MEDICINE AND PULMONARY DRUG DELIVERY
Volume 36, Number 1, 2023
Mary Ann Liebert, Inc.
Pp. 27–33
DOI: 10.1089/jamp.2022.0051

Open camera or QR reader and scan code to access this article and other resources online.

Application of Machine Learning for Segmentation of the Pulmonary Acinus Imaged by Synchrotron X-Ray Tomography

Branko Arsic, PhD,[1,2] Igor Saveljic, PhD,[1,2] Frank S. Henry, PhD,[3] Nenad Filipovic, PhD,[1,2] and Akira Tsuda, PhD[4]

Abstract

Background: To assess the effectiveness of inhalation therapy, it is important to evaluate the lungs' structure; thus, visualization of the entire lungs at the level of the alveoli is necessary. To achieve this goal, the applied visualization technique must satisfy the following two conditions simultaneously: (1) it has to obtain images of the entire lungs, since one part of the lungs is influenced by the other parts, and (2) the images have to capture the detailed structure of the alveolus/acinus in which gas exchange occurs. However, current visualization techniques do not fulfill these two conditions simultaneously. Segmentation is a process in which each pixel of the obtained high-resolution images is simplified (i.e., the representation of an image is changed by categorizing and modifying each pixel) so that we can perform three-dimensional volume rendering. One of the bottlenecks of current approaches is that the accuracy of the segmentation of each image has to be evaluated on the outcome of the process (mainly by an expert). It is a formidable task to evaluate the astronomically large numbers of images that would be required to resolve the entire lungs in high resolution.
Methods: To overcome this challenge, we propose a new approach based on machine learning (ML) techniques for the validation step.
Results: We demonstrate the accuracy of the segmentation process itself by comparison with previously validated images. In this ML approach, to achieve a reasonable accuracy, millions/billions of parameters used for segmentation have to be optimized. This computationally demanding new approach is achievable only due to recent dramatic increases in computation power.
Conclusion: The objective of this article is to explain the advantages of ML over the classical approach for acinar imaging.

Keywords: machine learning, lung, alveolus, U-Net, segmentation, image, artificial intelligence, computer, threshold method

[1]Department for Applied Mechanics, Faculty of Engineering, University of Kragujevac, Kragujevac, Serbia.
[2]BIOIRC Bioengineering Research and Development Center, Kragujevac, Serbia.
[3]Department of Mechanical Engineering, Manhattan College, Riverdale, New York, USA.
[4]Tsuda Lung Research, Shrewsbury, Massachusetts, USA.

Introduction

THE BEHAVIOR OF DRUG PARTICLES IN THE LUNGS depends largely on airflow patterns, which are chiefly conditioned by the anatomy of the airways. Therefore, knowledge of the configuration of the airways is fundamental to the understanding of the efficacy of inhalation therapy.

The structure of the lung parenchyma depends on a delicate force balance between pressure inside the lungs (i.e., alveolar pressure) and pressure outside the lungs (i.e., pleural pressure) (West, 2000) and the mechanical properties of lung tissue (i.e., various cells, interstitium, etc.) (Mead, 1961). Moreover, there is a strong interdependence between various regions of the lungs (Mead et al, 1970); that is, the local conditions (e.g., local tissue stiffness due to disease lesions, etc.) can affect the conditions of other parts. For instance, the flow distribution in the lungs, tissue expansion distribution of the lungs, and so on cannot be determined without knowing the mutual interdependence of various parts of the organ.

Therefore, it is important to visualize the whole lungs, instead of focusing only on a small portion of the lungs, to assess the condition of the lungs and their function as a gas exchange organ. This is a formidable task, especially with the high resolution required to visualize the alveolar structure. In this study, we propose a machine learning (ML) approach to circumvent the fundamental shortcoming of classical segmentation approaches; namely, the requirement of human judgment on the accuracy of each segmented image.

In our previous study (Tsuda et al, 2008), we used one of the most popular approaches for lung image processing, namely, threshold banalization methods [i.e., grays below the threshold are replaced with black (air), and grays above the threshold are replaced with white (tissue)]. Although we could successfully segment the images of the lung parenchyma (i.e., imaged by a high-resolution synchrotron radiation-based X-ray tomography and rendered to three-dimensional (3D) structure at micron scale), the area we could segment and thus deal with in our previous study was only a small portion of the parenchyma—a few acini—due to the limitation of the segmentation method (explained below) (Tsuda et al, 2008). To visualize the whole lung at the same micron-length scale, we need a more general and robust segmentation approach.

One of the bottlenecks of the classical segmentation approach is that the evaluation is performed on the results of the segmentation process (Haberthür et al, 2021; Haberthür et al, 2013; Haberthür et al, 2009; Sznitman et al, 2010; Tsuda et al, 2008; Vasilescua et al, 2012; Xiao et al, 2013); the classical approach requires the evaluation of the accuracy of the segmentation process on astronomically large numbers of images if the entire lungs were imaged to high resolution (Fig. 1, top panel). Since this approach (i.e., based on a posteriori evaluation), although theoretically possible, is not practical, we need a different approach to visualize the entire lungs in high resolution.

An ML approach (LeCun et al, 2015; Schmidhuber, 2015) uses a different concept of segmentation compared with the classical approach. Namely, in the classical concept, the evaluation is performed on the results of the segmentation process (i.e., on the output by *a posteriori* evaluation shown in Fig. 1, top), whereas in the ML approach, the function is treated as a black box (the relationship between the input and the output is treated as unknown, just expressed using a group of parameters), and by using a vast amount of past data, the expression of this relationship is optimized and evaluated (Fig. 1, bottom).

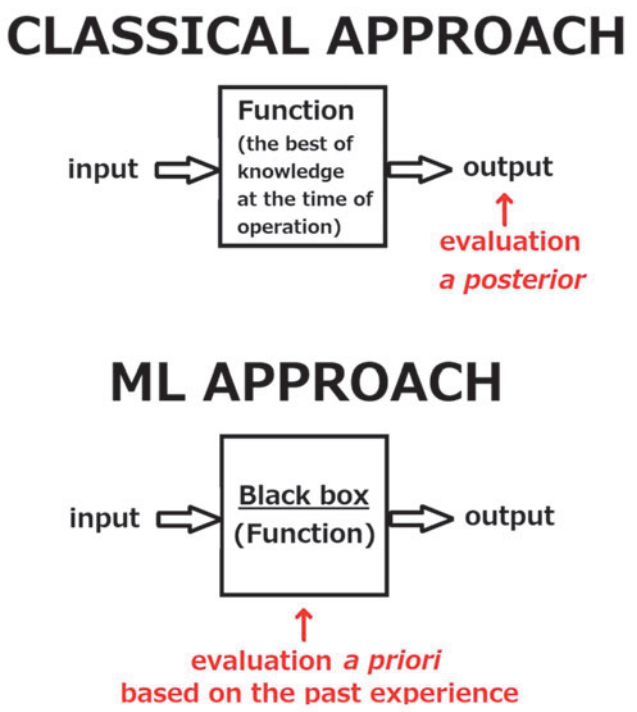

FIG. 1. Classical approach versus ML approach. In the classical approach, the evaluation is performed *a posteriori* on the output. However, in the ML approach, the function itself is evaluated *a priori* based on experience. ML, machine learning.

Subsequently, the optimized relationship is used to yield the prediction (output) for a given input. In other words, in the ML approach, the optimization is evaluated, that is, the evaluation is performed at the function level. This leads to a significant advantage of this new approach over the classical approach; namely, there is no limitation on the productivity once the optimal relationship between the input and the output (i.e., a group of the parameters to achieve the optimal prediction based on past experience) is established. Hence, the productivity of the ML approach does not depend on *a posteriori* evaluation.

The use of ML has dramatically increased in many fields (Hamet and Tremblay, 2017; Hussain and Zeadally, 2018; Huynh et al, 2020) as new techniques have been developed and computational power has increased, but this has not been the case for acinar imaging. The lack of uptake of ML for acinar imaging maybe due to the fact that this technique relies heavily on complex mathematics and uses a great deal of computer-science jargon. Hence, the technique is difficult for non-mathematicians to understand. The objective of this article is to demystify ML by explaining in simple terms the basic concepts of the ML approach (Supplementary Data) and its advantages over the classical approach for acinar imaging. We note that what follows is largely restricted to two subsets of ML; namely, deep learning and neural networks (NNs), the latter being the backbone of ML.

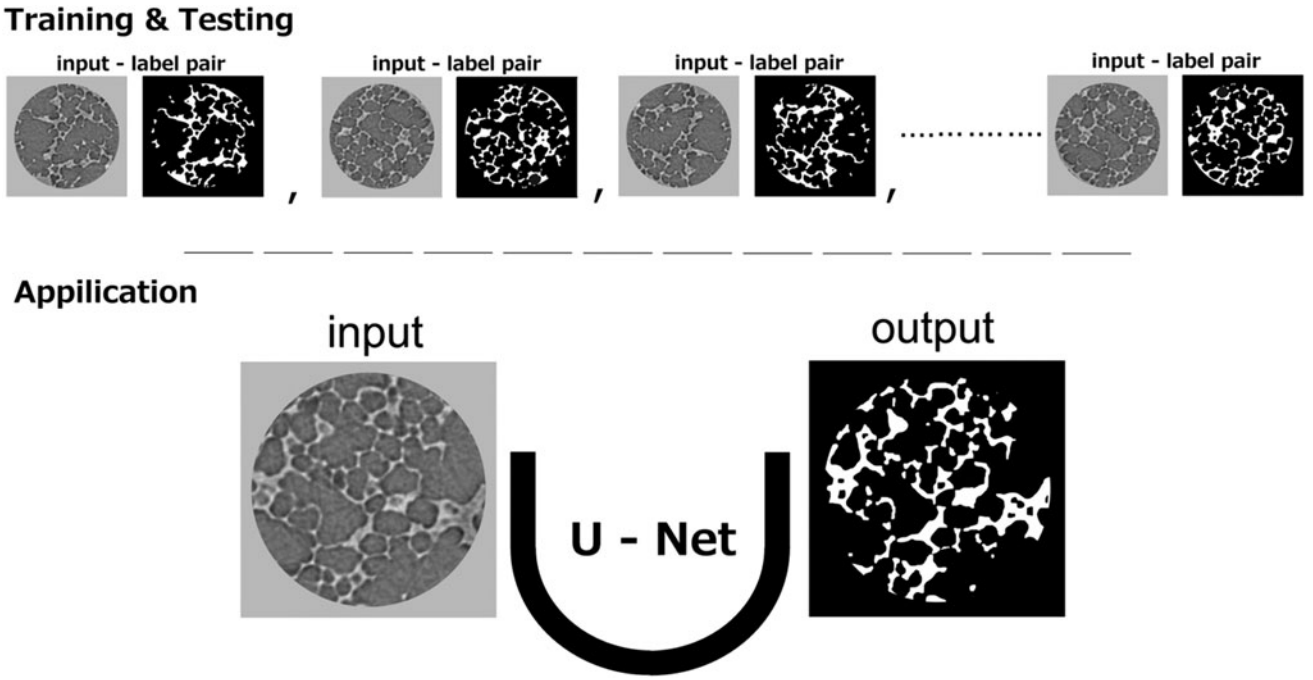

FIG. 2. ML approach. Top: 881 input label pairs, generated based on the techniques we used in the previous study (Tsuda et al, 2008), were used to turn and evaluate the parameters. Bottom: optimized parameters were applied to yield the best prediction.

Methods

An NN is a computer program that mimics the way in which a human brain learns. Today, NNs are being used in such diverse fields as tumor recognition and weather prediction. Similar to the human brain, an NN ''learns'' by being exposed to many different occurrences of a target event. A common example of the use of NN in image recognition is one that can distinguish digital images of cats from those of dogs. The NN knows nothing of the physical distinguishing features of cats and dogs, but it is trained to recognize particular patterns of pixels that generally occur in images of cats and not in images of dogs, and vice versa.

A vast amount of data are required to train the NN. These data are called ''learning material,'' and in the context of image recognition would consist of tagged, or labeled images. In the case of the cat/dog example, the learning material would be thousands of images of dogs and cats, each appropriately labeled dog or cat. The relationship between the input (e.g., an image of a dog) and the output (whether the image contains a dog or a cat) is determined mathematically using many relatively simple equations with unknown coefficients or parameters. The unknown parameters are optimized using the learning material. Using all the input data once, is termed an ''epoch.'' After each epoch, the error between the prediction and the label is calculated and another epoch is carried out if the desired accuracy has not been achieved.

In our case of segmenting images of the lung parenchyma, we used the data and results of our previous article as a gold standard since they have been previously evaluated stereologically by experts (Tschanz et al, 2003). Roughly 70% of the data set was used as training material and the remaining data were used to test the accuracy of the NN. In this case, each input-label pair comprised an original gray-scale image (the input) and the corresponding segmented image (the labeled image), the latter being created using the classical approach.

CNN (Convolution neural network) is a subset of NNs; here we developed CNN-based U-Net to work with images. A more thorough description of the technical details of ML is given in the Supplementary Data.

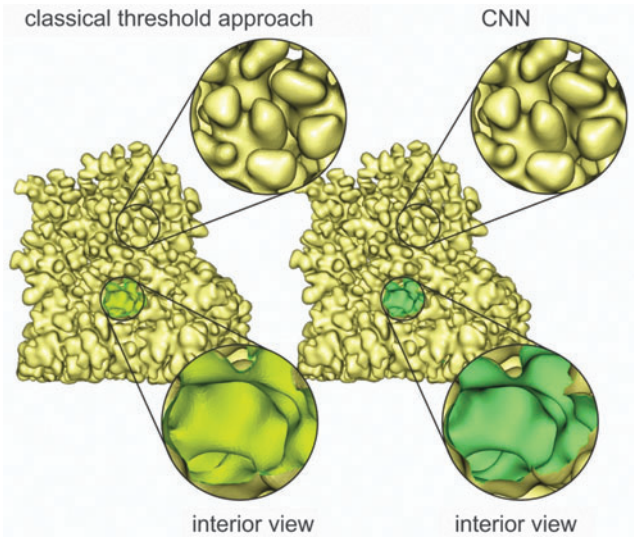

FIG. 3. Qualitative comparison in 3D structures between one segmented by the classical threshold approach and one segmented by the ML approach. They are qualitatively identical. 3D, three dimensional; CNN, Convolution neural network.

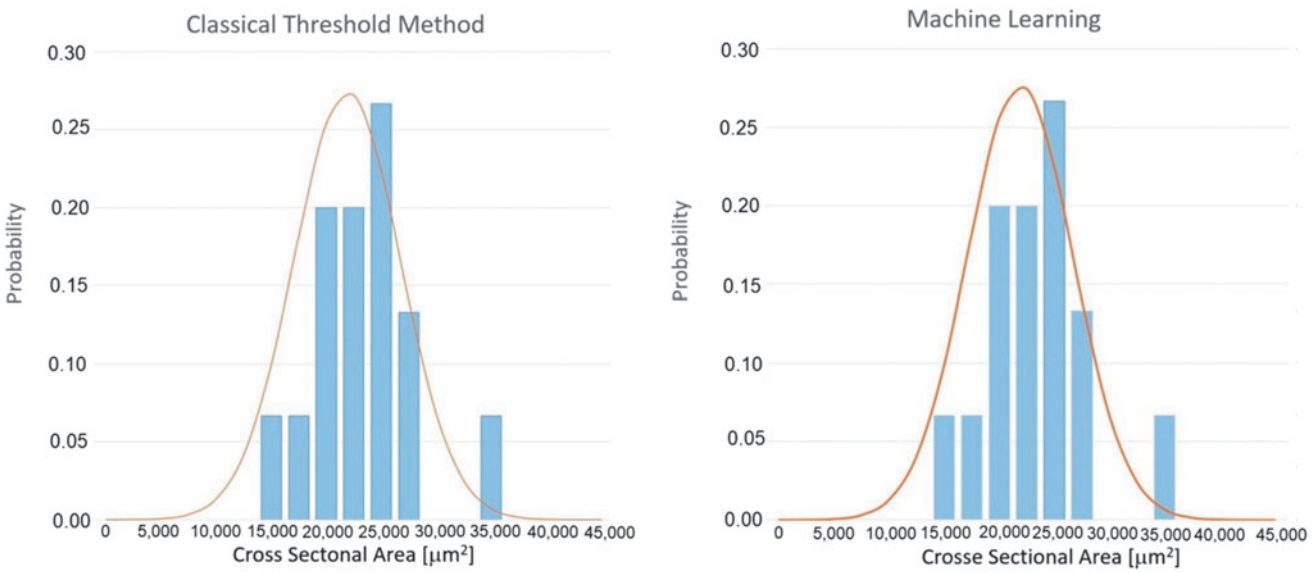

FIG. 4. Quantitative comparison in 3D structures between one segmented by the classical threshold approach and one segmented by the ML approach. A statistical analysis (Student's *t*-test) showed that the two methods are essentially identical (with $p = 0.001$), suggesting that the ML approach indeed yields the same level of segmentation by the classical threshold approach.

Results

Considering the results of our previous published study (Tsuda et al, 2008) as a gold standard, a total of 881 input-label pairs were generated based on the techniques we used in the previous study and these were utilized as learning material. In the training period, about 70% of the learning material (i.e., pairs of input label) was used to tune the 50,946,699 parameters of the system for optimization (Fig. 2, top). About 50 iterations were necessary to optimize the parameters with a tolerance of 0.001. In the testing period, the rest of the learning material was used for evaluation (Fig. 2, top). In the application, optimized parameters were applied to yield the best prediction (Fig. 2, bottom). Also, as an example, a ''fly through'' animation produced based on the segmented data by the ML technique is shown in the Supplementary Data (see Supplementary Video S1).

Using the newly developed model, based on the ML approach, a grayscale raw dataset was segmented and 3D volume rendering was performed by a well-established finite element method similar to our previous study (Tsuda et al, 2008). For comparison purposes, the same dataset was segmented by a classical threshold method (Tsuda et al, 2008) and 3D volume rendering was performed by the same FE method (Fig. 3). The obtained 3D structures segmented by the classical threshold approach (Fig. 3, left) and segmented by the ML approach (Fig. 3, right) are qualitatively identical.

To demonstrate quantitative similarity between the two approaches, the sizes of the alveolar opening of 30 samples were approximated using both methods and compared. The process of how we approximate the alveolar entrance area is described in detail in the Appendix. Briefly, 50 position vectors in 3D space in each alveolar entrance rim (i.e., the classical approach vs. the ML approach, 30 each) were measured. Then, from those 50 position vectors, the optimal two-dimensional alveolar entrance plane was approximated for each alveolar opening by matching the direction of the unit vector normal to the plane to the one calculating from 50 position vectors under the definition that a moment of inertia for the alveolus would be maximal.

Frequency distributions of the approximated alveolar entrance area (blue bars in Fig. 4) shows the Gaussian distribution (orange curve in Fig. 4) in both cases. Indeed, the distributions are practically identical (Fig. 4). A statistical analysis (Student's *t*-test) confirmed that both distributions are essentially identical (with $p = 0.001$), indicating that segmentation by the ML approach was successful.

To demonstrate the advantage of CNN-based U-Net over fully connected network (FCN) (refer to Supplementary Data), mean Intersection over Union (mIoU)* was computed and compared. Although the mIoU value of the U-Net (0.848) is comparable to the mIoU of FCN (0.846), the number of parameters to be determined in the current U-Net model is about 60 millions, which is a half of that in the case of FCN (130 millions). This shows that the U-Net can be trained faster and is more memory efficient compared to FCN.

Discussion

The principal aim of this work was to describe in an easy-to-understand manner the basic concepts of the highly mathematical ML approach to those who may not be particularly familiar with this method, and to explain why the utilization of ML for acinar image segmentation may have

*mIoU is a common evaluation metric for semantic image segmentation, which first computes the IoU for each semantic class and then computes the average over classes. … The predictions are accumulated in a confusion matrix, weighted by sample_weight, and the metric is then calculated from it. mIoU New semantic segmentation algorithms are typically assessed by the mIoU on the VOC2012 dataset. The IoU is calculated for each class at the pixel level as … The mIoU is then the mean value across all classes in the dataset.

major advantages over classical approaches. We used stereologically well-evaluated data as a gold standard (Tschanz et al, 2003; Tsuda et al, 2008) and reproduced our previously reported results with this new approach. Comparison of our results obtained with the new ML approach to the results obtained using the classical threshold approach (Figs. 3 and 4) shows a good degree of accuracy. Considering the potential advantages of this approach on productivity (see the Introduction) over the classical approach, it is clear that this ML approach is opening up new possibilities in image processing.

As discussed previously, the classical threshold approach of lung segmentation is not feasible[†] when a very large number of images have to be processed. This intrinsic problem of the classical approach is because the validity of the segmentation is evaluated on the outcome of the process, rather than on the process itself. To overcome the intrinsic problem that the classical method suffers, the evaluation should be done at a functional level. Thus, instead of solving the function (i.e., process) exactly and evaluating the output of the process, we optimize (and approximate) the process,[‡] and we evaluate the validity of optimization of the process. In this way, the productivity is not limited once the validity of optimization is proved.

Physiological consideration

Since the structure and function of various regions of the lungs are interdependent, to assess the condition of the lungs, it is important to obtain detailed images of the whole lungs, instead of visualizing only a small portion of the organ. For instance, it is critical to see the lungs as a whole because it is well known that the ventilation distribution (i.e., flow distribution in the lung) is largely determined by the balance of downstream mechanical load, rather than upstream flow conditions and local flow conditions (Otis et al, 1956; Tsuda et al, 1990).

The distribution of pressure, the distribution of tissue stiffness, and the distribution of airway obstruction in parallel pathways determine local acinar configuration. Visualization of the whole lungs requires a significant increase in the efficiency of image processing. We believe that the suggested ML approach, which utilizes computational power, instead of the traditional evaluation by human experts, would significantly increase the feasibility of whole lung imaging and would advance our ability to diagnose and treat lung diseases.

[†]It is an ill-defined topic to solve biological function exactly since the possibility of biological issues is infinite.

[‡]When ''process (i.e., segmentation function)'' is case-by-case, especially very complex, the optimization of the process would be more suitable, rather than an impossible search for the exact solution. The current situation is exactly such as case. Moreover, the optimization of the process is indeed similar to what we do everyday. For instance, to become ''a skilled morphologist,'' one needs to go through many years' hard training and experiments with a good teacher morphologist, although he or she must have been poor in skill at the beginning because of a lack of experience. In general, one with more experiences can approximate better than others with non- or less-experiences. Similarly, a better degree of optimization could be achieved with more learning material.

Summary

In this study, we have described the basic concept of the heavily mathematically based ML approach for image processing in an easy-to-understand manner and we have explained why this approach is advantageous over the classical approach for processing images of the lungs at high resolution. Thanks to recent increases in computer power and advances in analytical techniques, the computationally demanding optimization of the segmentation function in the ML approach has become possible.

Acknowledgments

The authors thank Dr. James P. Butler for a useful discussion on the Appendix. We acknowledge Dr. Johannes Schittny for his research and initial data related to animal experiment-referenced article (Tsuda et al, 2008).

Authors' Contributions

A.T. designed the research. B.A. and I.S. performed experiments. B.A., I.S., and N.F. analyzed data. A.T. designed all figures. A.T. process, analyzed, and created Figure 5. A.T. drafted article. B.A., I.S., F.S.H., N.F., and A.T. edited and revised the article and approved final version of article.

Author Disclosure Statement

The authors declare they have no conflicting financial interests.

Funding Information

This work was supported by grants from the Ministry of Education, Science and Technological Development of the Republic of Serbia through Contracts No. 451-03-68/2022-14/200107 and SGABU No. 952603 of the European Union.

Supplementary Material

Supplementary Data (see Online)
Supplementary Video S1 (see Online)

References

Haberthür D, Barré SF, Tschanz SA, et al. Visualization and stereological characterization of individual rat lung acini by high-resolution X-ray tomographic microscopy. J Appl Physiol 2013;115:1379–1387.

Haberthür D, Semmler-Behnke M, Takenaka S, et al. Multimodal imaging for the detection of sub-micron particles in the gas-exchange region of the mammalian lung. J Phys Conf Ser 2009;186:012040.

Haberthür D, Yao E, Barré SF, et al. Pulmonary acini exhibit complex changes during postnatal rat lung development. PLoS One 2021;16(11):e0257349.

Hamet P, Tremblay J. Artificial intelligence in medicine. Metabolism 2017;69S:S36–S40; doi: 10.1016/j.metabol.2017.01.011

Hussain R, Zeadally S. Autonomous cars: Research results, issues, and future challenges. IEEE Commun Surv Tutorials 2018;21(2):1275–1313.

Huynh E, Hosny A, Guthier C, et al. Artificial intelligence in radiation oncology. Nat Rev Clin Oncol 2020;17(12):771–781; doi: 10.1038/s41571-020-0417-8

LeCun Y, Bengio Y, Hinton GE. Deep learning. Nature 2015; 521(7553):436–444.

Mead J. Mechanics properties of lungs. Physiol Rev 1961; 41(2):281–330.

Mead J, Takashima T, Leith D. Stress distribution in lungs: A model of pulmonary elasticity. J Appl Physiol 1970;28(5): 596–608.

Otis AB, McKerrow CB, Bartlett RA, et al. Mechanical factors in distribution of pulmonary ventilation. J Appl Physiol 1956; 8:427–443.

Schmidhuber J. Deep learning in neural networks: An overview. Neural Netw 2015;61:85–117.

Sznitman J, Sutter R, Altorfer DM, et al. Visualization of respiratory flows from 3D reconstructed alveolar airspaces using X-ray tomographic microscopy. J Vis 2010;13:337–345.

Tschanz SA, Makanya AN, Haenni B, et al. Effects of neonatal high-dose short-term glucocorticoid treatment on the lung: A morphologic and morphometric study in the rat. Pediatr Res 2003;53:72–80.

Tsuda A, Filipovic N, Haberthür D, et al. Finite element 3D reconstruction of the pulmonary acinus imaged by synchrotron X-ray tomography. J Appl Physiol 2008;105:964–976.

Tsuda A, Kamm RD, Fredberg JJ. Periodic flow at airway bifurcations Part II: Flow partitioning. J Appl Physiol 1990; 69(2):553-561.

Vasilescua DM, Gaoa Z, Saha PK, et al. Assessment of morphometry of pulmonary acini in mouse lungs by nondestructive imaging using multiscale microcomputed tomography. Proc Natl Acad Sci U S A 2012;109(42):17105–17110.

West JB. Respiratory Physiology, The Essentials. Chapter 2. (6th ed). Lippincott Williams & Wilkins: Philadelphia, PA, USA; 2000.

Xiao L, Sera T, Koshiyama K, et al. A semiautomatic segmentation algorithm for extracting the complete structure of acini from synchrotron micro-CT images. Comput Math methods Med 2013;2013:575086; doi: 10.1155/2013/575086

Received on July 28, 2022
in final form, October 13, 2022

Reviewed by
Carsten Ehrhardt
Jessica Oakes

Address correspondence to:
Nenad Filipovic, PhD
Department for Applied Mechanics
Faculty of Engineering
University of Kragujevac
Sestre Janjica 6
Kragujevac 34000
Serbia

E-mail: fica@kg.ac.rs

Akira Tsuda, PhD
Tsuda Lung Research
28 Keyes House Road
Shrewsbury, MA 01545
USA

E-mail: tsudalung@gmail.com

(*Appendix follows* →)

Appendix: Approximation of an Alveolar Entrance Area

The alveolar entrance area was approximated as follows (Fig. 5): our calculation is based on two simple procedures: (1) define the optimal plane (shown in black in Fig. 5) and (2) given that plane, compute the projected area. First, we approximate the alveolar entrance rim on the plane (shown as a black curve in Fig. 5), on which the data points measured on the actual alveolar entrance rim[§] (shown as a red curve in Fig. 5) are projected, by defining the unit vector $\vec{\nu}$ normal to the plane. Expressing a coordinate system (x,y,z), whose origin is at the center of mass of the given data (N) points, let $M \left[= \begin{pmatrix} M_{xx} & M_{xy} & M_{xz} \\ M_{yx} & M_{yy} & M_{yz} \\ M_{zx} & M_{zt} & M_{zz} \end{pmatrix} \right]$ be the second-order spatial moment matrix of the rim.[**]

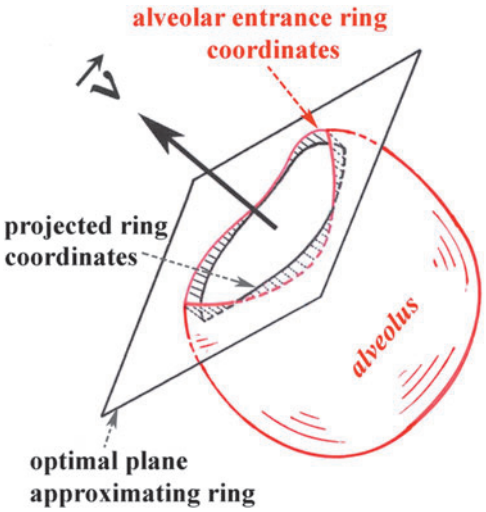

FIG. 5. A schematic view of how we obtain approximation of the alveolar entrance area.

Then the (rotational) moment of inertia $I^{\vec{\nu}}$ about any direction specified by a unit vector $\vec{\nu}$ is given by $I^{\vec{\nu}} = \mathrm{tr}M - \vec{\nu}^T M \vec{\nu}$, where tr denotes the trace. This quantity is maximal when $\vec{\nu}$ is normal to the approximating plane. Since the trace is invariant under rotations, maximizing $I^{\vec{\nu}}$ is equivalent to minimizing $\vec{\nu}^T M \vec{\nu}$. This is simply the minimal eigenvalue of M, and for which $\vec{\nu}$ is the (normalized) eigenvector. Second, having solved for $\vec{\nu}$, which minimizes $\vec{\nu}^T M \vec{\nu}$, we now project the points onto the plane defined by $\vec{\nu}$, and compute the area by triangular summation. For points in the (x, y) plane, this is the formula $A = (1/2)\left| \sum_{n=1}^{N} x_n y_{n+1} - x_{n+1} y_n \right|$, where the wraparound is defined by $x_{N+1} = x_1,\ y_{N+1} = y_N$.

[§]Following Haberthür et al (2021), we defined the alveolar entrance rings at the disappearance of the alveolar wall.

[**]The components of the matrix M are defined as $M_{pq} = \sum_{n=1}^{N} p_n q_n \quad (p,\ q = x, y, or\ z)$. Namely they are,

$$M_{xx} = (x_1 * x_1) + (x_2 * x_2) + (x_3 * x_3) + (x_4 * x_4) \cdots\cdots\cdots + (x_n * x_n) + \cdots\cdots\cdots + (x_N * x_N)$$

$$M_{xy} = (x_1 * y_1) + (x_2 * y_2) + (x_3 * y_3) + (x_4 * y_4) \cdots\cdots\cdots + (x_n * y_n) + \cdots\cdots\cdots + (x_N * y_N)$$

$$M_{xz} = (x_1 * z_1) + (x_2 * z_2) + (x_3 * z_3) + (x_4 * z_4) \cdots\cdots\cdots + (x_n * z_n) + \cdots\cdots\cdots + (x_N * z_N)$$

$$M_{yx} = (y_1 * x_1) + (y_2 * x_2) + (y_3 * x_3) + (y_4 * x_4) \cdots\cdots\cdots + (y_n * x_n) + \cdots\cdots\cdots + (y_N * x_N)$$

$$M_{yy} = (y_1 * y_1) + (y_2 * y_2) + (y_3 * y_3) + (y_4 * y_4) \cdots\cdots\cdots + (y_n * y_n) + \cdots\cdots\cdots + (y_N * y_N)$$

$$M_{yz} = (y_1 * z_1) + (y_2 * z_2) + (y_3 * z_3) + (y_4 * z_4) \cdots\cdots\cdots + (y_n * z_n) + \cdots\cdots\cdots + (y_N * z_N)$$

$$M_{zx} = (z_1 * x_1) + (z_2 * x_2) + (z_3 * x_3) + (z_4 * x_4) \cdots\cdots\cdots (z_n * x_n) + \cdots\cdots\cdots + (z_N * x_N)$$

$$M_{zy} = (z_1 * y_1) + (z_2 * y_2) + (z_3 * y_3) + (z_4 * y_4) + \cdots\cdots\cdots (z_n * y_n) + \cdots\cdots\cdots (z_N * y_N)$$

$$M_{zz} = (z_1 * z_1) + (z_2 * z_2) + (z_3 * z_3) + (z_4 * z_4) + \cdots\cdots\cdots (z_n * z_n) + \cdots\cdots\cdots + (z_N * z_N)$$