

RESEARCH ARTICLE

Non-Invasive Assessment of Lung Water Content Using Chest Patch RF Sensors: A Computer Study Using NIH Patients CT Scan Database and AI Classification Algorithms

CHRISTOPHER LEONG, (Member, IEEE), YUANZHANG XIAO^{ID}, (Member, IEEE),
ZHENGQING YUN^{ID}, (Senior Member, IEEE), AND MAGDY F. ISKANDER^{ID}, (Life Fellow, IEEE)

Hawaii Advanced Wireless Technologies Institute, University of Hawai'i at Mānoa, Honolulu, HI 96822, USA

Corresponding author: Magdy F. Iskander (magdy@hawaii.edu)

This work was supported in part by the U.S. National Science Foundation under Award 1822213.

ABSTRACT Chronic heart failure, pulmonary hypertension, acute respiratory distress syndrome (ARDS), coronavirus disease (COVID), and kidney failure are leading causes of death in the U.S. and across the globe. The cornerstone for managing these diseases is assessing patients' volume fluid status in lungs. Available methods for measuring fluid accumulation in lungs are either expensive and invasive, thus unsuitable for continuous monitoring, or inaccurate and unreliable. With the recent COVID-19 epidemic, the development of a non-invasive, affordable, and accurate method for assessing lung water content in patients became utmost priority for controlling these widespread respiratory related diseases. In this paper, we propose a novel approach for non-invasive assessment of lung water content in patients. The assessment includes quantitative baseline assessment of fluid accumulation in lungs (normal, moderate edema, edema), as well as continuous monitoring of changes in lung water content. The proposed method is based on using a pair of chest patch radio frequency (RF) sensors and measuring the scattering parameters (S-parameters) of a 915-MHz signal transmitted into the body. To conduct an extensive computational study and validate our results, we utilize a National Institute of Health (NIH) database of computerized tomography (CT) scans of lungs in a diverse population of patients. An automatic workflow is proposed to convert CT scan images to three-dimensional lung objects in High-Frequency Simulation Software and obtain the S-parameters of the lungs at different water levels. Then a personalized machine learning model is developed to assess lung water status based on patient attributes and S-parameter measurements. Decision trees are chosen as our models for the superior accuracy and interpretability. Important patient attributes are identified for lung water assessment. A "cluster-then-predict" approach is adopted, where we cluster the patients based on their ages and fat thickness and train a decision tree for each cluster, resulting in simpler and more interpretable decision trees with improved accuracy. The developed machine learning models achieve areas under the receiver operating characteristic curve of 0.719 and 0.756 for 115 male and 119 female patients, respectively. These results suggest that the proposed "Chest Patch" RF sensors and machine learning models present a promising approach for non-invasive monitoring of patients with respiratory diseases.

INDEX TERMS Biomedical measurements, biomedical monitoring, mobile measurements, RF sensors, artificial intelligence, machine learning.

The associate editor coordinating the review of this manuscript and approving it for publication was Ajit Khosla^{ID}.

I. INTRODUCTION

Chronic heart failure, pulmonary hypertension, acute respiratory disorder syndrome (ARDS), coronavirus disease (COVID), kidney failure, and their acute exacerbation are among the leading causes of hospitalization, health care costs,

and deaths in the United States and around the globe. For example, more than one million patients are hospitalized annually due to heart failure, which accounts for a total Medicare expenditure exceeding \$50 billion in 2013 [1]. Recently, the world was shocked with the wide spread of the COVID-19 pandemic which claimed about 3 million lives worldwide in a year; the death toll is still on rise as of the writing of this article. The cornerstone for managing these diseases is assessing patients' volume fluid status in lungs. Approximately 80% of the lung is made up of water, with gas-exchanging air spaces protected by various barriers and drains. In multiple disease states, through injury or pressure or both, these protective mechanisms fail, resulting in the abnormal accumulation of extravascular lung water. Hence, close monitoring of lung water status, respiratory rates and heart rates is key to proactively preventing the worsening of patient heart failure and treating acute exacerbation. For instance, the assessment of a patient with left ventricular systolic dysfunction and progressive dyspnea includes an evaluation of volume status and extravascular lung water – often assessed by measuring weight, jugular venous pressure, and presence of an S3 on cardiac exam, or peripheral edema. Similarly, patients with symptomatic hypotension are often assessed for their volume status, particularly if dehydration is a consideration and/or when complicating factors (e.g., renal insufficiency, peripheral neuropathy, and concomitant comorbid illnesses) confound the diagnostic exam. In summary, early detection of excess lung water is critical to provide timely fluid assessment and improve treatment for patients with chronic respiratory diseases.

However, existing modalities of monitoring lung water are either costly (e.g., chest X-ray and computerized tomography (CT) scan) and/or invasive (e.g., cardiac catheterization), making them unsuitable for continuous monitoring and early detection of excessive lung water. The widespread of COVID-19 made the development of non-invasive, accurate, and affordable methods for continuous assessment of lung water of utmost urgency and importance.

In this paper, we describe a novel approach for quantitatively assessing lung water content, in addition to measuring its changes when monitoring patients. The procedure is based on using a pair of chest patch radio frequency (RF) sensors and measuring the scattering parameters (S-parameters)¹ of a 915-MHz signal transmitted into the body. We aim to develop a machine learning model that assesses the lung water content (normal, moderate edema, edema) based on S-parameters and patient attributes. To build such a model, we use a large National Institutes of Health (NIH) data set containing CT scan images [2] of lungs in a diverse patient population. We then develop an automatic workflow to convert these images to three-dimensional (3D) lung objects in

¹S-parameter is the measure of the magnitude and phase of the received signal by the non-radiating sensor, as a result of the scattered RF signal inside the patient body.

High-Frequency Simulation Software (HFSS), and obtain the S-parameters of the lungs at different water levels.

Based on the data obtained from HFSS, we propose accurate and interpretable machine learning models for personalized assessment of lung water status. We first compare different classifiers and choose the decision tree classifier for its superior accuracy and interpretability. We then identify important features, namely S-parameters, genders, ages, and thickness of fat layers, for lung water assessment. Finally, we adopt a cluster-then-predict approach, where we cluster the patients into subgroups based on their genders, ages and fat thickness and train decision trees for each subgroup. This approach further improves the interpretability and accuracy of the decision trees for each subgroup. The final decision trees for each subgroup determine the lung water status using the magnitudes and the phases of the S-parameters only. On average, our personalized models achieve higher than 70% accuracy of lung water status assessment over a diverse group of 115 male and 119 female patients.

Section II provides a description of the chest patch RF sensor approach used in this computational study. A detailed description of the computational procedure of building a database of a diverse population from the NIH database is included in Section III. In Section IV, we propose a personalized machine learning model and present our results. Concluding remarks and comments on future work are included in Section V.

A two-page conference version of this work has been presented in [3]. The current work significantly expands our prior work [3] by describing the automatic workflow of building the database in detail (Section III) and including the development of the machine learning models (Section IV).

II. CHEST PATCH RF SENSORS FOR ASSESSING LUNG WATER CONTENT

The continuous non-invasive monitoring capability of the proposed chest patch RF sensors is achieved by using highly penetrating radio frequency and electromagnetic waves. At microwave frequencies, the dielectric properties of lung tissues are closely related to the water content in the lungs, as discovered and validated by our prior works through simulations [4], phantoms [5], and animal experiments [6]. Our prior efforts of clinical trials have demonstrated that this simple and noninvasive approach can accurately detect the change in lung water content [7]. Fig. 1 illustrates an example implementation of the chest patch RF sensor system used in this study. The RF system consists of an adhesive chest patch containing two electrocardiogram (EKG) lead sized RF sensors and measurement hardware and data analysis software components. The patch is placed in contact with the patient's chest, S-parameters are measured at 915 MHz, and multiple vital signs (e.g., heart rate and respiratory rate) are derived from a single measurement using digital signal processing algorithms.

The RF system was recently tested in clinical trials with heart failure and hemodialysis patients, which showed

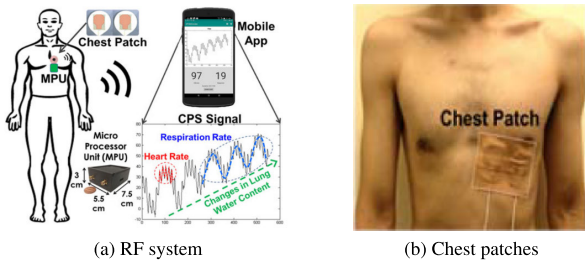


FIGURE 1. Illustration of the chest patch RF system and chest patches.

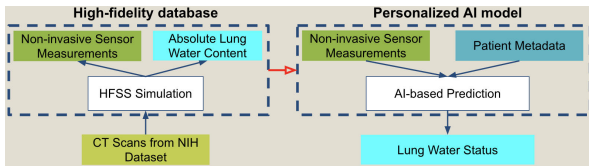


FIGURE 2. The proposed framework of artificial intelligence (AI) based lung water assessment.

excellent correlations with other available clinical monitoring methods [7]. Specifically, the heart and respiration rates measured by the RF system have correlation factors higher than 0.9 for all the patients. Comparisons with fluid removed during hemodialysis treatment showed correlation factors of 0.82 to 1, while comparisons with pulmonary capillary wedge pressure measurements for heart failure patients had correlation factors of 0.52 to 0.97.

In this work, we propose an artificial intelligent (AI) based lung water prediction to significantly expand the capability of the chest patch RF system (see Fig. 2 for illustration of the proposed framework). First, we aim to assess the lung water status (i.e., normal, moderate edema, and edema), in addition to the change of lung water. Second, we seek to build accurate personalized AI models for diverse patient populations. One key challenge to develop a personalized AI model is to build a database of a diverse patient population, because individuals are different both in the baseline for water content in normal lungs and in the changes of dielectric properties under different severities of edema. The high cost of collecting data from clinical trials makes the challenge even more difficult. To address this challenge, we use a large NIH data set of CT scan images and develop an automatic workflow to obtain high-fidelity data from 3D HFSS. This generates a data set that includes patients with varying ages, genders, and sizes, and the S-parameters of their lung tissues under various amounts of lung water. Compared to data collection from clinical trials, this approach is low-cost, less time-consuming, and risk-free. The following section provides detailed description of the procedure of building the database.

III. BUILDING THE DATABASE OF A DIVERSE PATIENT POPULATION

To address the challenges of the lack of data and the high cost of collecting data from clinical trials, we develop an automatic workflow to build the database from a large-scale NIH dataset, DeepLesion [2]. DeepLesion

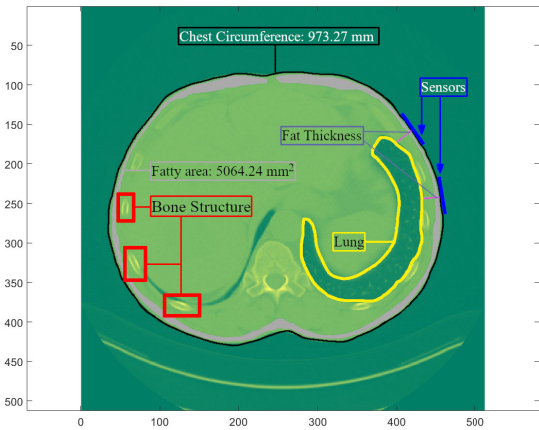


FIGURE 3. Axial cross-section at the height of the RF sensor in MATLAB. Dark gray structures are bones of the thoracic cage while those in yellow are lung tissues. The sensors are attached at the top-right of the body. We use the chest circumference and the fat thickness, which can be measured conveniently, as features for the machine learning models developed next.

contains 32,120 axial CT scan slices of 4,427 unique patients. We first select high-quality CT scans of the lungs. Then we use MATLAB for image processing. Finally, we use HFSS [8] to build 3D models of the lungs and obtain the S-parameters at lung water content of 20% (normal lung), 40% (moderate edema), and 60% (edema). Below we describe in detail our automatic workflow to build the database.

A. SELECTION OF CT SCANS

Since DeepLesion contains CT scans of different body parts (e.g., lungs, kidneys, pelvis), we first identify all CT scans of the lungs based on the metadata comma-separated values (CSV) file provided by DeepLesion. Among the CT scans of the lungs, our biomedical expert select the ones of good quality. For example, we discard the CT scans that show only the tip of the lungs and those taken during expiration of the respiratory cycle.

B. IMAGE PROCESSING IN MATLAB

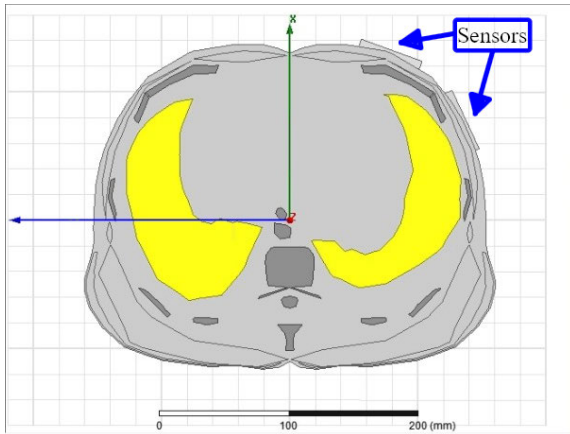
We import the CT scans into MATLAB as 512×512 images. We use the image processing functions in MATLAB to detect the edges of the chest, bone structures, and lungs. The polygons composed of these edges are written into a script to be used by HFSS. We also develop a MATLAB script to determine the exact locations of the two RF sensors. The two sensors are placed 8 cm apart on the cross-section outline with a clear view of the lungs. This avoids the blockage of RF signals by the bone structures such as the thoracic cage. Finally, we have a MATLAB script to calculate the thoracic circumference, the area of the fat layer, and the thickness of the fat layer, which will be used as part of the patient metadata in the database. See Fig. 3 for an illustration of the annotated MATLAB image after this step.

C. SIMULATION IN HFSS

HFSS reads the script, extract the polygons, and use them to build a cylindrical model of the lung. We use the pixel

TABLE 1. Dielectric properties used for the lung polygons, based on fractional volumes of blood, air, and lung tissue [9].

Lung Water Content	Relative Permittivity	Electrical Conductivity (S/m)
20%	20.604	0.417
40%	30.679	0.590
60%	40.510	0.755

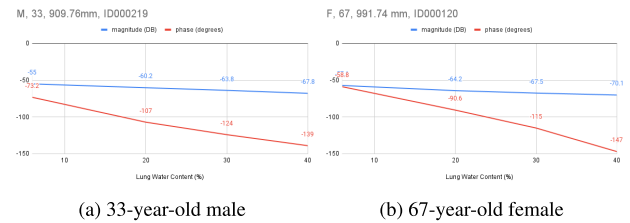
**FIGURE 4.** Illustration of the cylindrical lung model built from the axial CT scan in HFSS. The figure shows the axial cross-section of the lung model at the height of sensor attachment. Yellow polygons represent lungs, dark gray polygons represent bones, and gray polygons represent an averaged mixture of epithelial, muscle, and fat tissues.

to mm conversion provided by DeepLesion to ensure the correct sizes of our lung model. The horizontal locations of the RF sensors were determined by MATLAB, and the vertical location is the middle height of the cylinder. HFSS simulates a 915-MHz RF signal sent by one RF sensor, and its internal electromagnetic solver determines the S-parameters of the received signal at the other RF sensor. We perform the simulation at lung water content of 20% (normal lung), 40% (moderate edema), and 60% (edema). The dielectric properties of the bones and tissues are known and fixed, and those of the lung are determined according to the literature [9] (shown in Table 1). Magnitudes and phases of the S-parameters are taken for the lung water content of each patient. See Fig. 4 for illustration of the 3D lung model built in HFSS and Fig. 5 for S-parameters at different lung water levels for two representative patients.

D. DATABASE ENTRY

For each CT scan, we obtain three data samples from one unique patient. Each data sample contains the magnitude and the phase of the S-parameter at one of the three lung water levels. We also record the patient metadata of gender and age (available from DeepLesion), and thoracic circumference, area of the fat layer, and thickness of the fat layer (obtained by MATLAB). The lung water percentage is the label of each data sample.

Currently, we have data samples of 115 male patients and 119 female patients with ages from 20 to 86, which will be used in our AI-based lung water assessment described next. The sample size is similar to most clinical trials related to

**FIGURE 5.** Sample results of HFSS simulation for a 33-year-old male and a 67-year-old female. We obtain magnitudes and phases of S-parameters at different lung water levels.

pulmonary edema [10]. Therefore, our study serves as an indicator of how our AI model would perform in an actual clinical trial. We are also expanding our database as an ongoing work.

IV. AI-BASED LUNG WATER ASSESSMENT

We aim to assess the lung water status based on the patient attributes (i.e., gender, age, chest circumference, fat thickness) and S-parameters (i.e., magnitudes and phases). We pose the problem as a classification problem with three classes, namely *normal* (i.e., 20% lung water content), *moderate edema* (i.e., 40% lung water content), and *edema* (i.e., 60% lung water content).

We evaluate the AI models using three criteria:

- **Accuracy:** This is defined as the percentage of the samples that are correctly classified;
- **Receiver operating characteristic (ROC) curve:** This is the curve of true positive rates versus false positive rate, which provides a complete characterization of the performance of the classifier [11]. Since the ROC curve is defined for binary classification problem and since we have three classes of normal, moderate edema, and edema in our problem, we adopt a “one-vs-rest” approach when plotting the ROC curve [12]. Specifically, we will group normal and moderate edema into one class and investigate how well the classifier can distinguish between normal/moderate edema and edema; we will also group moderate edema and edema into one class and investigate how well the classifier can distinguish between normal and moderate edema/edema.
- **Area under the ROC curve (AUC) score:** This is the area under the ROC curve, which is a scalar summarizing the ROC curve [13].

For all the performance criteria, we perform stratified 10-fold cross validation [14], where we divide the data set into groups and use one group as the validation set and the other groups as the training set. All performance scores are the average scores over the 10 folds.

A. AI MODEL SELECTION

There are various models available to perform the classification task. Our first step is to select the model that is most suitable for our data based on accuracy and interpretability. We consider the following commonly-used classification models.

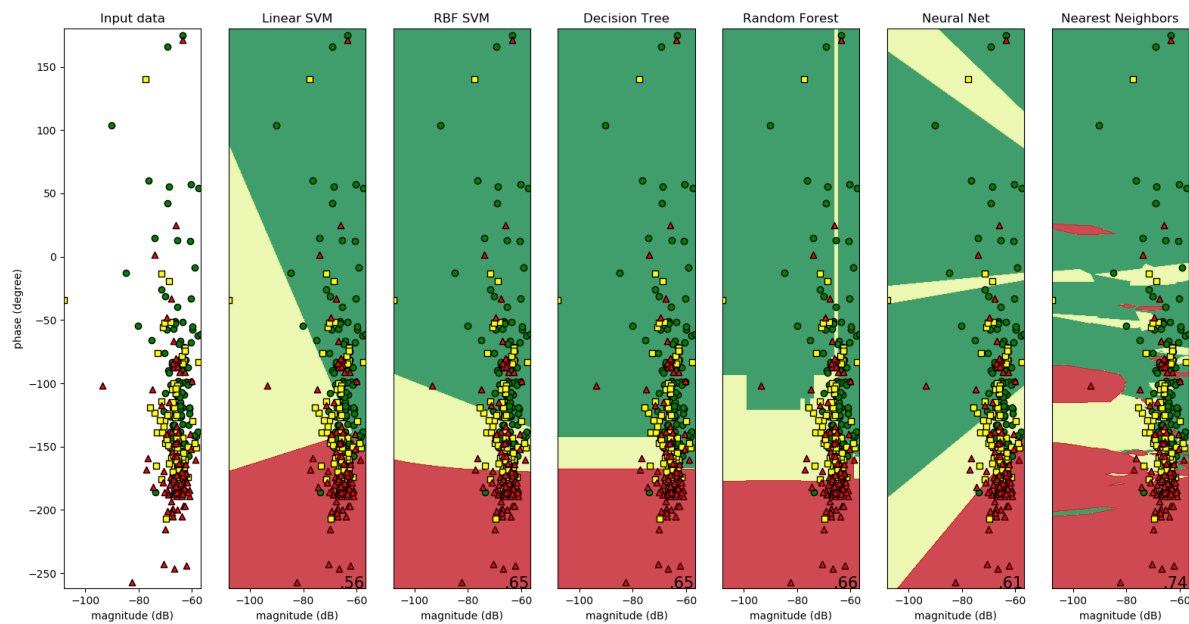


FIGURE 6. Illustration of data samples and decision regions of different machine learning models. In all the subplots, the x-axis and y-axis are the magnitude and the phase of the S-parameters, and the green circles, yellow squares, and red triangles are the data samples representing normal lung water (20%), moderate edema (40%), and edema (60%), respectively. Leftmost: the data samples; Right: decision regions of nearest neighbors, decision trees, random forests, and neural networks, respectively, where green, yellow, and red pixels indicate that the model classifies the corresponding samples as normal lung water, moderate edema, and edema. The accuracy is shown in the lower right corner of each subplot.

- Support vector machines (SVM) [15]: SVM maps the input vectors non-linearly to a high-dimension feature space, and then uses linear decision surfaces in the feature space to separate the data samples. We evaluate linear SVM (which removes the non-linear mapping) and radial basis function (RBF) SVM (which uses the radian basis function as the non-linear mapping).
- Decision trees [16]: A decision tree classifier uses “if-then” type decision rules on the features to predict the class of a data sample. A decision tree can be seen as a piece-wise constant approximation.
- Random forests [17]: A random forest classifier is the average of multiple decision tree classifiers.
- Neural networks [18]: A neural network is a network of neurons (i.e., simple processing units specified by an activation function). A neural network can implement highly non-linear classification rules by adjusting the connections between the neurons.
- Nearest neighbors [19]: A nearest neighbor classifier groups the training samples “closest in distance” to the new point, and predict the class from these samples. A nearest neighbor classifier can also be highly non-linear.

Fig. 6 illustrates the data samples for all the male patients, and decision regions and accuracy of the AI models above.

From the plot of the data samples, we can see that the data cannot be separated by straight lines. Therefore, linear models, such as linear SVM, are not accurate. Based on this observation, we focus on nonlinear models, namely

RBF SVM, decision trees, random forests, neural networks, and nearest neighbors. Among all the non-linear models, nearest neighbors have the highest accuracy. However, the decision regions of the nearest neighbors are too complex to explain. We observe that decision trees have same or higher accuracy than RBF SVM and neural networks. Compared to random forests, decision trees have slightly lower accuracy, but much simpler decision regions. Overall, we found that decision trees have the best accuracy and interpretability trade-off.

Another advantage of decision trees is that the decision rules have the if-then structure that resembles natural language and the way humans think (see Fig. 7 for an illustration of the decision tree). For example, the decision tree in Fig. 6 stipulates the following decision rule: If the phase is between -167.5 and -133.5 and the magnitude is above -64.7 dB, the lung is normal; if the phase is between -167.5 and -133.5 and the magnitude is below -64.7 dB, the lung has moderate edema; if the phase is between -177.5 and -172.5 , the lung has moderate edema; otherwise, the lung has edema. We can see that decision trees provide human-friendly explanations, which is especially desirable in clinical settings.

B. FEATURE IMPORTANCE AND FEATURE SELECTION

Now that we have decided to use decision trees as our model, we proceed to perform feature selection. Our goal is to select the most important features, instead of using all the features, for the model. Feature selection will improve the model generalizability and interpretability [20].

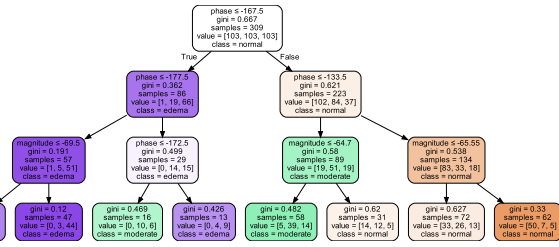


FIGURE 7. Illustration of the decision rules stipulated by the decision tree classifier in Fig. 6: If the phase is between -167.5 and -133.5 and the magnitude is above -64.7 dB, the lung is normal; if the phase is between -167.5 and -133.5 and the magnitude is below -64.7 dB, the lung has moderate edema; if the phase is between -177.5 and -172.5 , the lung has moderate edema; otherwise, the lung has edema.

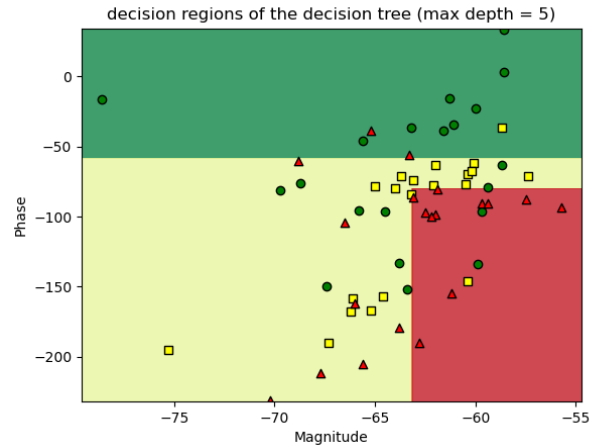


FIGURE 9. Illustration of decision regions of an *undesirable* classifier for females patients with age between 35 and 45. It is undesirable because the red region (i.e., edema) is on the right to the yellow region (i.e., moderate edema).

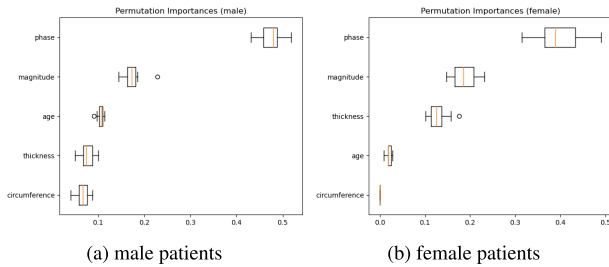


FIGURE 8. Illustration of the permutation importance of all the features for decision trees. For each feature, the box plot shows the median of the importance (the orange line), the first and the third quartiles (the box), 1.5 times the interquartile range (IQR) beyond the third quartiles (the whiskers extending from the box), and the outliers (the circles).

There are different methods to quantify feature importance. We use the permutation importance, a metric commonly used in practice, to measure the importance of each feature [20], [21]. Permutation importance measures the impact of a feature on the classification accuracy. Specifically, to evaluate a feature, we create data sets with random permutations of the feature to evaluate while keeping the other features the same, and see how much the classification accuracy drops. When permuted, a more important feature will result in a larger drop in the accuracy.

Fig. 8 shows the permutation importance of several decision trees with different tree depths. We can see that the phase is by far the important feature. The magnitude and the fat thickness are the next two important features, with the relative importance of these two varying across models. The circumference and the age are the least important.

C. THE CLUSTER-THEN-PREDICT APPROACH

Based on the important features identified, we adopt the cluster-then-predict approach to build the final AI models [22], [23]. Specifically, we first cluster the patients into subgroups, and then train a decision tree classifier for each subgroup of patients. The advantages of the cluster-then-predict approach are higher average accuracy due to the similarity of patients within one subgroup and better interpretability due to simpler classifiers for each subgroup.

We first determine the features that serve as the basis to cluster the patients. It is important to use features that are constant (i.e., age, circumference, and fat thickness) for

clustering. If we were to use time-varying features (e.g., magnitude and phase) for clustering, the subgroups would change quickly over time. Based on the feature importance shown in Fig. 8, we drop the circumference as a feature due to its low importance. We first try to cluster the patients based on the age, and then based on age and fat thickness.

1) CLUSTERING BASED ON AGE

We adopt a heuristic approach to determine the age brackets for the subgroups. Starting from the youngest patient, we gradually increase the size of the first age group until the accuracy drops a lot. Once the first age group is determined, we repeat the above process for the next age groups, until we cluster all the patients. For male patients, this approach results in seven age brackets, namely [20, 30], [31, 40], [41, 49], [50, 55], [56, 60], [61, 69], and [70, 99], and an average AUC score of 75.2%. For female patients, this approach results in some age brackets for which the model always predicts higher water content under higher magnitudes. Such results violate the physics, because the signal should attenuate more when there is more lung water. Please see Fig. 9 for an illustration. Note that we did not exhaust all the possible clustering, which is practically impossible. But the phenomenon of predicting higher water content under higher magnitudes seems persistent in all the trials. As a result, we conclude that it is challenging to obtain interpretable models when clustering female patients based on the age only.

2) CLUSTERING BASED ON AGE AND FAT THICKNESS

Since the fat thickness has been identified as an important feature, we cluster the patients based on both the age and the fat thickness. In particular, we divide the patients of the same gender into four subgroups by setting a threshold of the age and a threshold of the fat thickness. The median ages of male and female patients are 52 and 50, respectively. In our attempt to find the optimal thresholds of the age and the fat thickness, we choose the threshold of the age from {50, 55}

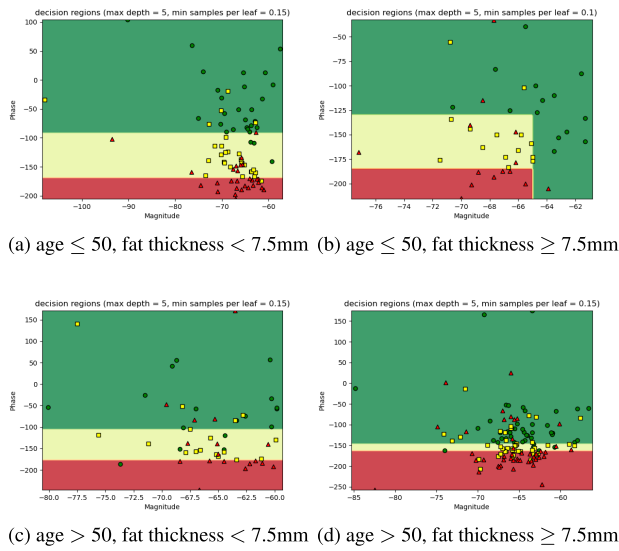


FIGURE 10. Decision regions of each subgroup of male patients. In all the subplots, the x-axis and y-axis are the magnitude and the phase of the S-parameters, and the green circles, yellow squares, and red triangles are the data samples representing normal water content (20%), moderate edema (40%), and edema (60%), respectively. The green, yellow, and red pixels indicate that the model classifies the corresponding samples as normal water content, moderate edema, and edema.

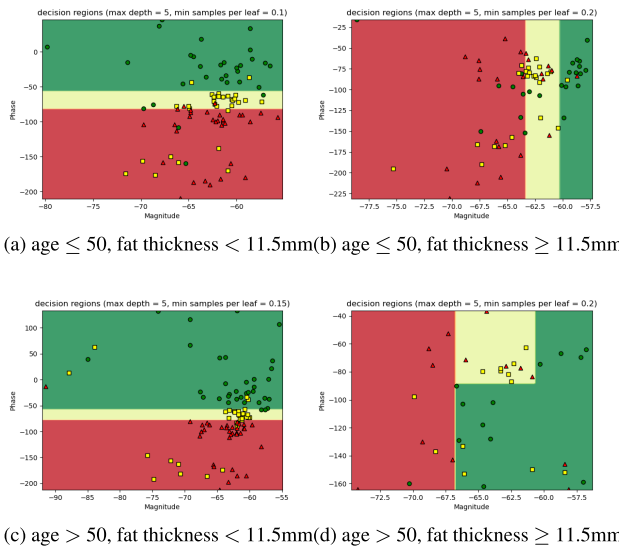


FIGURE 11. Decision regions of each subgroup of female patients.

and the threshold of fat thickness from {6, 6.5, . . . , 11.5, 12}. The optimal threshold is determined based on the average AUC score over the four subgroups.

We find that the optimal threshold of the age is 50 for both male and female patients, and that the optimal thresholds of the fat thickness are 7.5mm and 11.5mm for male and female patients, respectively. Clustering based on age and fat thickness has two advantages over clustering based on age only. First, for male patients, we only need four subgroups, as opposed to seven in clustering based on age, to achieve similar average AUC score of 71.9%. For female patients,

TABLE 2. Summary of model performance in terms of the AUC score.

	cluster	# of patients	subgroup AUC	average AUC
Male	1	28	78.6%	71.9%
	2	14	78.6%	
	3	17	68.6%	
	4	44	66.7%	
Female	1	31	81.7%	75.6%
	2	23	60.9%	
	3	34	84.3%	
	4	13	64.1%	

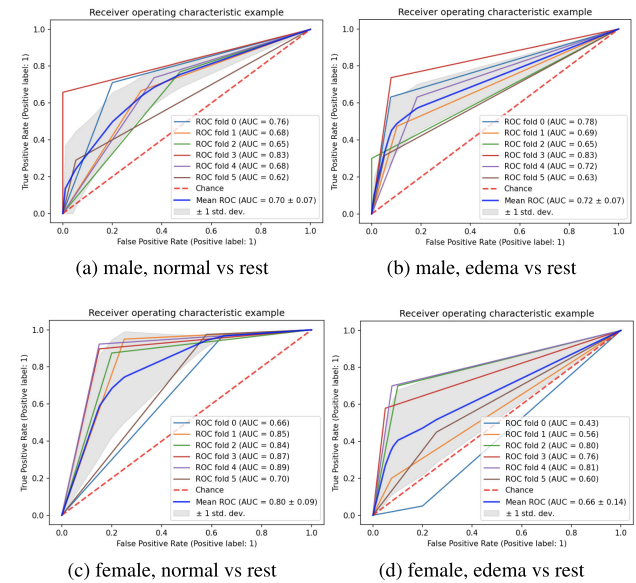


FIGURE 12. ROC curves of male and female patients. In all the subplots, the x-axis is the false positive rate and the y-axis is the true positive rate. The thick blue curves are the mean ROC curves, and thin curves are six ROC curves from the sixfold cross validations. The shaded regions indicate the standard deviation.

the predicted lung water content always decreases with the magnitude, which conforms to the physics. We summarize the results in Table 2, and illustrate the decision regions of the classifiers for male and female patients in Fig. 10 and Fig. 11, respectively.

Compared to clustering based on the age only, clustering based on the age and the fat thickness is better, because it leads to fewer subgroups (i.e., four subgroups as oppose to seven) with negligible sacrifice in the AUC score. For each subgroup, the decision trees determine the lung water status based on the magnitude and the phase of the S-parameters, resulting in simple and interpretable decision rules.

Finally, we show the ROC curves of male and female patients in Fig. 12.

D. MODEL INTERPRETATION

It is important to interpret the machine learning model [20]. For our study, we check if the model conforms with the following expert knowledge:

- Our prior works suggest that the phase is the most important feature [24].

- According to electromagnetics, the magnitude of the signal decreases when there is more water in the lung.
- We apply the two chest patches, acting as the transmitter and the receiver, on the chest, as opposed to one on the chest and one on the back. Therefore, the chest circumference has limited impact on the signal strength.
- In clustering based on age and fat thickness, the threshold of the fat thickness is lower for male patients (7.5mm) compared to female (11.5mm). This makes sense because males have lower body fat percentages on average.

Our results are consistent with the expert knowledge mentioned above. The importance of the phase and the limited impact of the circumference are observed in our study on feature importance (Fig. 10 and Fig. 11). The monotonic relationship between the magnitude and the lung water content is observed from the decision regions (Fig. 10 and Fig. 11).

V. CONCLUSION AND FUTURE WORK

In this paper, we propose a novel approach for lung water assessment by the chest patch RF sensors and measurement system. This provides a significant extension in the capabilities of our previous chest patch RF sensors system as it provides quantitative baseline assessment of lung water status (i.e., normal lung water content, moderate edema, and edema) for a diverse patient population, in addition to monitoring the change of lung water content. We use a large NIH database containing CT scan images of the lungs of a diverse population. Then an automatic workflow is proposed to convert these images to 3D lung objects in HFSS and obtain the S-parameters of the lungs at different water levels. This approach results in a database of a diverse patient population without expensive and time-consuming clinical trials. Using this database, we develop a personalized machine learning model to assess lung water status based on the patient attributes and S-parameter measurements. Our AI model adopts decision trees as the classifier for its superior accuracy and interpretability. Then we propose a “cluster-then-predict” approach, namely clustering the patients into subgroups and training a decision tree for each subgroup. This leads to even simpler and more interpretable decision rules with high accuracy. When the patients are clustered based on their ages only, the resulting decision trees for the male patients perform well, but those for the female patients are hard to interpret. Therefore, the patients of each gender are clustered based on both their ages and fat thickness. The final decision trees for each cluster determine lung water status using S-parameters only, which are easy to interpret. Overall, our models achieve areas under the receiver operating characteristic curve (AUC) of 0.719 and 0.756 over 115 male and 119 female patients of different ages (20 to 86) and body fat levels. These results demonstrated the potential of the proposed AI-based chest patch RF system in non-invasive monitoring of lung water content.

For future work, we intend to conduct a true clinical trial on a diverse population of patients in collaboration with medical

centers from across the U.S.. The presented computational study and the obtained promising results will significantly help in guiding these clinical studies. In particular, our results help focus the clinical studies on parameters identified as critically important in assessing the baseline value of lung water content in patients.

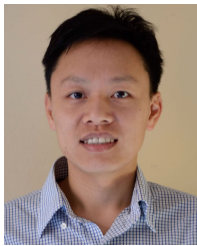
REFERENCES

- [1] G. Savarese, P. M. Becher, L. H. Lund, P. Seferovic, G. M. C. Rosano, and A. J. S. Coats, “Global burden of heart failure: A comprehensive and updated review of epidemiology,” *Cardiovascular Res.*, vol. 118, no. 17, pp. 3272–3287, Jan. 2023.
- [2] K. Yan, X. Wang, L. Lu, and R. M. Summers, “DeepLesion: Automated mining of large-scale lesion annotations and universal lesion detection with deep learning,” *J. Med. Imag.*, vol. 5, no. 3, Jul. 2018, Art. no. 036501.
- [3] C. Leong, Y. Xiao, Z. Yun, and M. F. Iskander, “Non-invasive stethoscope for continuous assessment of lung water: Towards AI-based data augmentation and prediction,” in *Proc. IEEE Int. Symp. Antennas Propag. USNC-URSI Radio Sci. Meeting (AP-S/URSI)*, Jul. 2022, pp. 341–342.
- [4] N. Celik, R. Gagarin, H.-S. Youn, and M. F. Iskander, “A noninvasive microwave sensor and signal processing technique for continuous monitoring of vital signs,” *IEEE Antennas Wireless Propag. Lett.*, vol. 10, pp. 286–289, 2011.
- [5] N. Celik, R. Gagarin, G. C. Huang, M. F. Iskander, and B. W. Berg, “Microwave stethoscope: Development and benchmarking of a vital signs sensor using computer-controlled phantoms and human studies,” *IEEE Trans. Biomed. Eng.*, vol. 61, no. 8, pp. 2341–2349, Aug. 2014.
- [6] M. F. Iskander and C. H. Durney, “Microwave methods of measuring changes in lung water,” *J. Microw. Power*, vol. 18, no. 3, pp. 265–275, Jan. 1983.
- [7] M. F. Iskander, T. B. Seto, R. G. Perron, E. Lim, and F. Qazi, “Cardio-pulmonary stethoscope: Clinical validation with heart failure and hemodialysis patients,” *IEEE Trans. Biomed. Eng.*, vol. 65, no. 5, pp. 1176–1180, May 2018.
- [8] ANSYS HFSS. (Version 14.0). ANSYS. Accessed: Jul. 2, 2023. [Online]. Available: <https://www.ansys.com/products/electronics/ansys-hfss>
- [9] P. C. Pedersen, C. C. Johnson, C. H. Durney, and D. G. Bragg, “Microwave reflection and transmission measurements for pulmonary diagnosis and monitoring,” *IEEE Trans. Biomed. Eng.*, vol. BME-25, no. 1, pp. 40–48, Jan. 1978.
- [10] J. Masip, M. Roque, B. Sanchez, R. Fernandez, M. Subirana, and J. A. Exposito, “Noninvasive ventilation in acute cardiogenic pulmonary edema: Systematic review and meta-analysis,” *Jama*, vol. 294, no. 24, pp. 3124–3130, 2005.
- [11] J. A. Hanley and B. J. McNeil, “The meaning and use of the area under a receiver operating characteristic (ROC) curve,” *Radiology*, vol. 143, no. 1, pp. 29–36, 1982.
- [12] D. J. Hand and R. J. Till, “A simple generalisation of the area under the ROC curve for multiple class classification problems,” *Machine Learn.*, vol. 45, no. 2, pp. 171–186, Nov. 2001.
- [13] A. P. Bradley, “The use of the area under the ROC curve in the evaluation of machine learning algorithms,” *Pattern Recognit.*, vol. 30, no. 7, pp. 1145–1159, 1997.
- [14] P. Refaeilzadeh, L. Tang, H. Liu, L. Angeles, and C. D. Scientist, “Cross-validation,” *Encyclopedia Database Syst.*, vol. 5, pp. 532–538, Jan. 2020.
- [15] C. Cortes and V. Vapnik, “Support-vector networks,” *Mach. Learn.*, vol. 20, pp. 273–297, Apr. 1995.
- [16] T. Hastie, R. Tibshirani, J. H. Friedman, and J. H. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, vol. 2. Berlin, Germany: Springer, 2009.
- [17] T. K. Ho, “The random subspace method for constructing decision forests,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 20, no. 8, pp. 832–844, Aug. 1998.
- [18] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016.
- [19] J. Goldberger, G. E. Hinton, S. Roweis, and R. R. Salakhutdinov, “Neighbourhood components analysis,” in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 17, 2004, pp. 1–8.
- [20] C. Molnar, *Interpretable Machine Learning*, 2nd ed. 2022. Accessed: Jul. 2, 2023. [Online]. Available: <https://christophm.github.io/interpretable-ml-book>

- [21] A. Fisher, C. Rudin, and F. Dominici, "All models are wrong, but many are useful: Learning a variable's importance by studying an entire class of prediction models simultaneously," *J. Mach. Learn. Res.*, vol. 20, no. 177, pp. 1–81, Jan. 2019.
- [22] R. Soni and K. J. Mathai, "An innovative 'cluster-then-predict' approach for improved sentiment prediction," in *Advanced Computing and Communication Technologies*. Berlin, Germany: Springer, 2016, pp. 131–140.
- [23] J. Galvez-Goicurla, J. Pagan, A. B. Gago-Veiga, J. M. Moya, and J. L. Ayala, "Cluster-then-classify methodology for the identification of pain episodes in chronic diseases," *IEEE J. Biomed. Health Informat.*, vol. 26, no. 5, pp. 2339–2350, May 2022.
- [24] M. F. Iskander, R. Maini, C. H. Durney, and D. G. Bragg, "A microwave method for measuring changes in lung water content: Numerical simulation," *IEEE Trans. Biomed. Eng.*, vol. BME-28, no. 12, pp. 797–804, Dec. 1981.

CHRISTOPHER LEONG (Member, IEEE) received the B.S. degree in biological engineering from the University of Hawai'i at Mānoa, Honolulu, HI, USA, in 2020, where he is currently pursuing the M.S. degree in electrical engineering.

In 2020, he worked as a Research Technician at Jun Innovations. Since 2021, he has been a Research Assistant with the Hawaii Advanced Wireless Technologies Institute, University of Hawai'i at Mānoa. His current research interests include the biomedical applications of electromagnetics for detection and monitoring of respiratory conditions.



YUANZHANG XIAO (Member, IEEE) received the B.E. and M.E. degrees in electronic engineering from Tsinghua University, in 2006 and 2009, respectively, and the Ph.D. degree in electrical engineering from the University of California at Los Angeles (UCLA), Los Angeles, CA, USA, in 2014. He was a Postdoctoral Fellow at Northwestern University, from 2015 to 2017. He has been an Assistant Professor with the University of Hawai'i at Mānoa, since 2017. His research

interests include reinforcement learning, game theory, and optimization, with applications in smart grids, wireless communications, wireless health, and socio-technological networks.

ZHENGQING YUN (Senior Member, IEEE) received the Ph.D. degree in electrical engineering from Chongqing University, Chongqing, China, in 1994.

He is currently an Associate Professor with the Hawaii Advanced Wireless Technologies Institute (previously Hawaii Center for Advanced Communications, HCAC), College of Engineering, University of Hawai'i at Mānoa. His current research interests include radio propagation in complex environments, such as urban, indoor, and mountainous areas, and ocean surface and atmospheric ducts.

Dr. Yun has served as the Technical Program Co-Chair for the IEEE Antenna and Propagation Society International Symposium, Honolulu, in 2007, and the Technical Program Chair for the IEEE International Conference on Wireless Information Technology and Systems, Honolulu, in 2010, and Maui, HI, USA, in 2012. He was an Associate Editor of the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY and the IEEE TRANSACTIONS ON ANTENNAS AND PROPAGATION. He is currently an Associate Editor of IEEE ACCESS.



MAGDY F. ISKANDER (Life Fellow, IEEE) is currently a Professor of electrical engineering and the Director of the Hawaii Advanced Wireless Technologies Institute (HAWT Institute), College of Engineering, University of Hawai'i at Mānoa, Honolulu, HI, USA. Prior to joining the University of Hawai'i at Mānoa, in 2002, he was a Professor of electrical and computer engineering at The University of Utah, from 1977 to 2001. He was the Engineering Clinic Endowed Chair Professor in 2000. From 1997 to 1999, he was the Program Director at the National Science Foundation, where he formulated a "Wireless Information Technology" Initiative in the Engineering Directorate. He has authored the textbook *Electromagnetic Fields and Waves* (Prentice Hall, 1992) and (Waveland Press, 2001); the *CAEME Software* (Second Edition, 2012, Vol. I and II, 1991–1994); and edited four books on *Microwave Processing of Materials* (Materials Research Society, 1990–1996). He has published more than 250 articles in technical journals and book chapters, 11 patents, and made numerous presentations in national and international conferences. His research is funded by the National Science Foundation, U.S. Army CERDEC, Office of Naval Research, and several corporate sponsors. As a result of the NSF Major Research Instrumentation Grant, he established wireless testbed and indoor antenna range, and the Microwave Network Analysis Laboratory and the RF Fabrication And Characterization Laboratory at the University of Hawai'i at Mānoa. His center HCAC (now HAWT Institute) had an ongoing ten-year grant, from 2005 to 2014, for partnership in the NSF Industry/University Cooperative Research Center in Telecommunications with The University of Arizona, Arizona State University, Rensselaer Polytechnic Institute (RPI), and The Ohio State University. In 2018, he established another partnership with the NSF IUCRC for Electromagnetic Compatibility with the Missouri University of Science and Technology, Rolla, MO, USA, and the University of Houston, Houston, TX, USA. After receiving NIH grants and conducting successful clinical trials, his group established MiWa Technologies, LLC, for the commercialization of their patented Cardio-Pulmonary Stethoscope Technology for monitoring vital signs and lung water measurements. His research interests include antenna design and propagation modeling for wireless communications and radar systems.

He was a member of the 1999 WTEC Panel on "Wireless Information Technology-Europe and Japan," and chaired two International Technology Panels on "Asian Telecommunication Technology" sponsored by NSF/DoD, in 2001 and 2003. He was the 2002 President of IEEE Antennas and Propagation Society, a Distinguished Lecturer of IEEE AP-S, from 1994 to 1997, and a fellow of the Applied Computational Electromagnetics Society. He received many teaching excellence and research awards, including the ASEE Curtis W. McGraw National Research Award, in 1985, the ASEE George Westinghouse National Education Award, in 1991, the Richard R. Stoddard Award from the IEEE EMC Society, in 1992, the University of Utah Distinguished Teaching Award, in 2000, the Northrop Grumman Excellence in Teaching Award, in 2010, the Board of Regents' Medal for Teaching Excellence, in 2010, the Hi Chang Chai Outstanding Teaching Award, in 2011, the University of Hawai'i Board of Regent Medal for Excellence in Research, in 2012, the IEEE AP-S Chen-To Tai Distinguished Educator Award, in 2012, and the IEEE MTT-S Distinguished Educator Award, in 2013. He organized and was the General Chair of nine IEEE conferences both in Utah (two) and Hawaii (seven). Since 1992, he has been a Founding Editor of the *Computer Applications in Engineering Education* (CAE) journal (Wiley). He edited two Special Issues of the IEEE TRANSACTIONS ON ANTENNAS AND PROPAGATION on "Wireless Communications Technology," in 2002 and 2006, and co-edited the Special Issue of IEICE TRANSACTIONS, Japan, in 2004.

• • •