Computerized Diagnosis of Respiratory Disorders

SVM Based Classification of VAR Model Parameters of Respiratory Sounds

I. Sen; M. Saraclar; Y. P. Kahya
Electrical and Electronics Engineering Department, Bogazici University, Istanbul, Turkey

Keywords
Respiratory sounds, vector autoregressive (VAR) model, support vector machine (SVM) classifier, diagnostic classification

Summary
Introduction: This article is part of the Focus Theme of Methods of Information in Medicine on “Biosignal Interpretation: Advanced Methods for Studying Cardiovascular and Respiratory Systems”.

Objectives: This work proposes an algorithm for diagnostic classification of multi-channel respiratory sounds.

Methods: 14-channel respiratory sounds are modeled assuming a 250-point second order vector autoregressive (VAR) process, and the estimated model parameters are used to feed a support vector machine (SVM) classifier. Both a three-class classifier (healthy, bronchiectasis and interstitial pulmonary disease) and a binary classifier (healthy versus pathological) are considered.

Results: In the binary scheme, the sensitivity and specificity for both classes are 85% ± 8.2%. In the three-class classification scheme, the healthy recall (95% ± 5%) and the interstitial pulmonary disease recall and precision (100% ± 0% both) are rather high. However, bronchiectasis recall is very low (30% ± 15.3%), resulting in poor healthy and bronchiectasis precision rates (76% ± 8.7% and 75% ± 25%, respectively). The main reason behind these poor rates is that the bronchiectasis is confused with the healthy case.

Conclusions: The proposed method is promising, nevertheless, it should be improved such that other mathematical models, additional features, and/or other classifiers are to be experimented in future studies.

1. Introduction

The common practice in pulmonary clinics (after listening to the complaints and the medical history of the patient) is still stethoscope auscultation on the chest wall. Depending on the acoustic findings, the physician decides on the type and the severity of the disease and through successive auscultations s/he estimates the pathology location in the lungs. In those cases where the physician suspects a pathological condition that can be revealed via computed tomography (CT) imaging, s/he requires the patient to go through a CT scan, to be able to arrive at the final diagnosis. Stethoscope auscultation is subjective since the ensuing decision depends on the hearing ability and the experience of the physician, and does not allow patient follow-up and comparison since the data cannot be stored. On the other hand, CT imaging may be impractical and harmful for a group of patients (e.g., disabled or elder people and pregnant women), and may also be expensive for a group of clinical environments (e.g., small health centers, or hospitals in deprived regions of the country).

Computerized respiratory sounds analysis is more reliable and objective compared to stethoscope auscultation, moreover, it is more suitable for quantification of the data. Sounds acquired on the chest wall during respiration can be recorded, and analyzed mathematically in order to capture the details and characteristics that the physician cannot hear during a stethoscope auscultation, which has a very limited frequency bandwidth. Besides, computerized respiratory sounds analysis in the diagnosis of pulmonary disorders can be an alternative or complementary method which is more practical, cheaper, and less harmful than CT scanning. Acoustic data are invaluable in the diagnosis since most of the pathological conditions that occur in the lungs cause characteristic changes in the respiratory sounds. In the recent decades, computerized techniques have been increasingly employed for objective and quantitative analysis and comparison of respiratory sounds. Moreover, increasing number of studies has been carried out with multi-channel measurements with an aim to analyze the characteristics with respect to different auscultation locations simultaneously [1, 2].

Considering that the lungs are a multiple-input-multiple-output (MIMO) system, the aim in this work is to model the multi-channel respiratory sounds using an...
appropriate mathematical model, and to investigate whether the model parameters are useful for diagnosis. Accordingly, a multivariate mathematical model is fit on the 14-channel sound data acquired on the chest wall during respiration, and the estimated model parameters are used to feed a classifier.

Autoregressive (AR) processes are known to be successful in modeling respiratory sounds [2–5]. In this work, their multivariate versions, namely, vector autoregressive (VAR) processes, are considered. The search of the appropriate VAR model for respiratory sounds was conducted in a previous study [6], where the measure of appropriateness was the ability of model parameters to maximize the within-group similarities and between-group differences inherent in respiratory sounds, with regard to healthy and pathological subject groups. Accordingly, the optimal sample size and order of the model for respiratory sound data were proposed to be N = 250 and p = 2 respectively. In this study, those two selections are adopted to have the VAR model with the maximal discriminative ability, and it is investigated whether the model parameters are really useful in a diagnostic classification scheme. For the classification algorithm, support vector machine (SVM) classifier, a discriminant-based method that is widely used for a variety of applications, is adopted.

Most of the pulmonary diseases are grouped under two main categories, namely, obstructive and restrictive. As a starting point, bronchiectasis (of obstructive type) and interstitial pulmonary disease (of restrictive type) are selected for the classification problem. Both the three-class classifier (the two disorders and the healthy group) and the binary classifier (healthy versus pathological) schemes are considered.

The rest of the paper is organized as follows. The data acquisition system and the data used in this work are introduced in Section 2.1. Sections 2.2 and 2.3 summarize the methodology including VAR modeling and SVM classification. Results are presented in Section 3. Finally, Section 4 includes a brief discussion together with the conclusion.

2. Methods

2.1. Data Acquisition System and Data

The data used in this study have been selected from the database recorded via the 14-channel pulmonary sound data acquisition and processing system [7] that was designed and implemented in the Bogazici University Lung Acoustics Laboratory. The system is composed of 14 air-coupled electret microphones (SONY ECM-44 BPT) attached on the posterior chest wall (Figure 1), an analog amplifier-filter unit (with a gain of 100 and a pass band of 80 to 4000 Hz), a Fleisch type pneumotachograph (Validyne CD379) to measure the flow-cycle simultaneously for synchronization, a data acquisition card (NI DAQCard-6024E, 12-bit) for digitization, and a laptop computer to control the process (via an interface program implemented in LabVIEW) and store the data. The data are sampled at a rate of 9600 samples per second, and an acquisition session lasts for 15 seconds. During the recording sessions, subjects sit upright and wear a nose clip while breathing through the mouthpiece of the flowmeter.

Pulmonary sound data of 20 healthy and 20 pathological subjects (10 of them diagnosed with bronchiectasis (obstructive) and the remaining 10 with interstitial pulmonary disease (restrictive)) have been selected from the database to be used in this work, according to the consideration that the recordings were free from external noise and the data classes were proportional in size. Healthy subjects were non-smokers who had no history of any serious lung disease. This study has the approval of the Second Ethical Committee on Clinical Research of Istanbul (is in compliance with the Declaration of Helsinki). An informed consent is given by each subject.

As a pre-processing step, flow signal is automatically divided into respiratory sub-phases, namely, early/mid/late inspiration/expiration. The inspiration and expiration phases are determined by the flow signal portions above and below a threshold band allocated around the zero level, where the band is determined as a percentage of the maximum flow signal deviation. Sub-phases in turn are determined according to the area under the curve, i.e., the volume of air inhaled and exhaled. Early and late phases represent 30 percent of total air inhaled (or exhaled) during one inspiration (or expiration) period, while the mid phase represents the remaining 40 percent. Thereby, one complete inspiration-expiration cycle is composed of six respiratory sub-phases (flow-phases hereafter).
2.2 Vector Autoregressive Modeling

A K-dimensional vector autoregressive process of order p (K-dimensional VAR(p) process) can be expressed as [8]

\[ y_n = \mathbf{v} + \mathbf{A}_1 y_{n-1} + \ldots + \mathbf{A}_p y_{n-p} + \mathbf{u}_n, \]

where \( y_n = [y_{1n}, \ldots, y_{K_n}]^\top \) is a (K x 1) random vector, \( \mathbf{v} = [v_1, \ldots, v_K]^\top \) is a fixed (K x 1) vector of intercept terms allowing for the possibility of a nonzero mean \( E[y_n] \), \( \mathbf{A}_i \) are fixed (K x K) coefficient matrices and

\[ \mathbf{A}_i = \begin{bmatrix} a_{11,i} & \cdots & a_{1K,i} \\ \vdots & \ddots & \vdots \\ a_{K1,i} & \cdots & a_{KK,i} \end{bmatrix}, \quad i = 1, \ldots, p \]

and finally, \( \mathbf{u}_n = [u_{1n}, \ldots, u_{Kn}]^\top \) is a K-dimensional white noise, i.e., \( \mathbb{E} \{ \mathbf{u}_n \} = 0 \) and

\[ \mathbb{E} \{ \mathbf{u}_n \mathbf{u}_{n'}^\top \} = \Sigma_u, \quad \text{if } n = n', \]

\[ = 0, \quad \text{otherwise} \]

(3)

where \( \Sigma_u \) can be either diagonal or not. The process \( y_n \) is assumed to be stable and stationary, the latter assumption allowing \( \mathbf{v} \) and \( \mathbf{A}_i \) to be accepted as fixed (time invariant). Since the respiratory sound data are inherently non-stationary, segmentation into shorter intervals is performed before fitting the model, to meet the stationarity assumption. In this study, the sample size and order for the VAR model are selected to be 250 and 2 respectively, following the previous work [6] (please see Section 1), and K = 14 since the respiratory sounds are recorded via 14 channels.

Each full inspiration-expiration cycle is divided into six flow-phases and labeled as explained in Section 2.1. Let \( P_{s,j} \) denote one such flow-phase, where \( i = 1, 2, \ldots, (FCN) \), \( j = 1, 2, \ldots, 6 \) (flow-phase index, from early inspiration to late expiration) and \( s = 1, 2, \ldots, N \) (subject index, \( N = 40 \)). (FCN) is the total number of full cycles (one inspiration followed by one expiration will be called a full cycle hereafter) for subject \( s \), and \( i \) is the full cycle index.

For each of \( P_{s,j} \), the data are automatically divided into 250-point 50% overlapping segments. Let \( (WN)_{s,j} \) denote the maximum number of such segments fitting into \( P_{s,j} \). For each segment, a VAR(2) model is fitted and the parameter matrices \( \mathbf{A}_i \) and \( \mathbf{A}_1 \) are estimated, where they are (14 x 14) matrices since \( K = 14 \) (Equation 2). These are concatenated to form the (14 x 28) matrices \( \mathbf{A} = [\mathbf{A}_1, \mathbf{A}_1] \) which can be denoted as \( \mathbf{A}_{klij} \), where \( k = 1, 2, \ldots, (WN)_{s,j} \). Each such matrix is converted into a row vector \( \mathbf{a}_{klij} \in \mathbb{R}^d \), where \( d = 14 \times 28 = 392 \), and \( \mathcal{A}_j \) is the set of these row vectors, \( \mathcal{A}_j = \{ \mathbf{a}_{klij} : 1 \leq k \leq (WN)_{s,j}, 1 \leq l \leq (FCN) \} \), defined per flow-phase per subject, \( j = 1, 2, \ldots, 6 \) and \( s = 1, 2, \ldots, 40 \). These sets are the inputs of the classification algorithm.

2.3 Support Vector Machines

For binary classification problems, support vector machine (SVM) algorithm proposes a method to find the optimal separating hyperplane. By employing a nonlinear kernel function in an SVM classifier, the original feature space can be mapped through a nonlinear transformation to a new space where the optimal separating hyperplane is to be found. For more information, the reader is invited to refer to [9, 10].

In this work, LIBSVM [11] is used to train and test the classifier. For multi-class problems, LIBSVM follows one-against-one approach, which can be summarized as considering two of the classes at each time to estimate a label for the instance, then deciding on the true label by majority voting. As the kernel functions, linear, quadratic, and radial basis functions [11] are experimented.

To assess the classification performance, 40-fold cross validation scheme is adopted. At each fold, the data of one subject are omitted from the data set, the classifier is trained on the remaining data, and the omitted subset is used to validate the classifier performance at that fold. The classification is performed for the six flow-phases separately, since the VAR parameters are expected to reveal different characteristics due to the assumption that the six flow-phases have different underlying physical mechanisms.

Following the notation of Section 2.2, \( \mathcal{A}_j \) is the validation set for flow-phase \( j \) at fold \( s \), \( j = 1, 2, \ldots, 6 \) and \( s = 1, 2, \ldots, 40 \). The classification performances can be evaluated from the following three perspectives:

- Perspective 1 (On the basis of segment classification): Each segment, i.e., each of the 250-point segments of the sound data that the VAR model is fitted on, is assigned a class label and successes are calculated at this level. (Algorithmically, the percentages of correctly and incorrectly classified instances in \( \mathcal{A}_j \) are calculated for all \( j \) and \( s \).)

Table 1 Binary classification confusion matrices with performance rates and their standard errors. C: True label, Ĉ: Predicted label, H: Healthy, P: Pathological. Values are in percent.

<table>
<thead>
<tr>
<th>Quadratic</th>
<th>Linear</th>
<th>RBF</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Ĉ = H</td>
<td>Ĉ = P</td>
</tr>
<tr>
<td>Perspective 1</td>
<td>C = H</td>
<td>80.9 ± 4.9</td>
</tr>
<tr>
<td>C = P</td>
<td>35.2 ± 6.3</td>
<td>64.8 ± 6.3</td>
</tr>
<tr>
<td>Perspective 2</td>
<td>C = H</td>
<td>87.5 ± 5.9</td>
</tr>
<tr>
<td>C = P</td>
<td>28.3 ± 8.0</td>
<td>71.7 ± 8.0</td>
</tr>
<tr>
<td>Perspective 3</td>
<td>C = H</td>
<td>90 ± 6.9</td>
</tr>
<tr>
<td>C = P</td>
<td>20 ± 9.2</td>
<td>80 ± 9.2</td>
</tr>
</tbody>
</table>

Perspective 2 (On the basis of flow-phases): All labeled segments produce one common decision per flow-phase per subject. (Algorithmically, one label is estimated for every j and s by majority voting over the labels estimated for $A_j$.)

Perspective 3 (On the basis of subjects): All labeled flow-phases produce one common decision per subject, through majority voting.

The third perspective yields the estimated class labels of subjects, from where one can only calculate the percentages of correctly and incorrectly classified subjects. However, for the first two perspectives, one can calculate the average or percentage values either over the folds to observe the performance variation with respect to the flow-phases, or, over the flow-phases to observe the performance variation with respect to the folds. In this work, the latter approach is adopted.

### 3. Results

The performance results of the binary classification (healthy versus pathological) with the three different SVM kernel functions are presented in Table 1 from the three perspectives introduced above. All three functions yield comparable results, while the radial basis function yields the most balanced confusion matrices. For clarity, only the radial basis function is chosen to report the 3-class classification (healthy, bronchiectasis and interstitial pulmonary disease) results in Table 2. The reason that some of the rows in the table do not sum up to 100% in terms of the success rates is that there are indecisive cases due to majority voting. The indecisive cases in binary classification scheme are labeled as pathological since it is clinically meaningful as well.

Sensitivity (recall) and specificity are defined as TP/(TP + FN) and TN/(TN + FP), respectively (T/F: True/False, P/N: Positives/Negatives). Precision is defined on the other hand as TP/(TP + FP). Sensitivity, specificity, precision and recall rates are calculated using the classification results from the third perspective since it is the subject-based one, i.e., the most meaningful one from the clinical point of view. Sensitivity and specificity are 85% ± 8.2% for both classes in healthy versus pathological classification. The precision and recall (sensitivity) rates for the 3-class classification scheme are given in Table 3. The recall rate of the healthy class (95% ± 5%) and both the recall and precision rates of the interstitial pulmonary disease class (100% ± 0% both) are rather satisfactory. However, the recall rate of bronchiectasis is very low (30% ± 15.3%), resulting in poor precision rates for the healthy and bronchiectasis classes (76% ± 8.7% and 75% ± 25%, respectively). The main reason behind these poor rates is that the bronchiectasis is confused with the healthy case (Table 2).

### 4. Discussion and Conclusion

In this work, 14-channel respiratory sounds are modeled assuming a 250-point VAR(2) process, and the estimated model parameters are used to feed an SVM classifier. Both a 3-class classifier (healthy, bronchiectasis and interstitial pulmonary disease) and a binary classifier (healthy versus pathological) are considered. In the binary scheme, the sensitivity and specificity for both classes are 85% ± 8.2%. In the 3-class classification scheme, bronchiectasis has a low recall rate (30% ± 15.3%), which stems from the fact that it is confused with the healthy class. This in turn results in a low healthy precision (76% ± 8.7%), which is apt to reduce the reliability of the proposed algorithm in clinical applications. The reason that the bronchiectasis has a poor precision rate (75% ± 25%) is also that the bronchiectasis is confused with the healthy case, rather than that the healthy (or interstitial pulmonary disease) is confused with bronchiectasis. Except for the low rates due to the confusion, the success rates are rather high, which renders the method promising.

The deviation errors in Table 1 and Table 2 are mainly based on the variation between the subject scores. However, one other reason is that the six flow-phases are not altered to the same extent even in the case of a pathological condition. This is the case especially if the alteration in the spectral characteristics is mainly due to the existence of adventitious sounds, because they may appear in different flow phases depending on the type of the disease.

![Table 2](https://example.com/table2.png)

<table>
<thead>
<tr>
<th>Perspective</th>
<th>$\hat{C} = H$</th>
<th>$\hat{C} = B$</th>
<th>$\hat{C} = I$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perspective 1</td>
<td>83.6 ± 4.1</td>
<td>11.2 ± 2.9</td>
<td>5.2 ± 2.0</td>
</tr>
<tr>
<td>$C = H$</td>
<td>54.3 ± 9.1</td>
<td>29.4 ± 8.3</td>
<td>16.3 ± 4.1</td>
</tr>
<tr>
<td>$C = B$</td>
<td>22.5 ± 3.4</td>
<td>11.9 ± 2.4</td>
<td>65.6 ± 3.6</td>
</tr>
<tr>
<td>Perspective 2</td>
<td>91.7 ± 4.3</td>
<td>5.8 ± 3.5</td>
<td>2.5 ± 1.8</td>
</tr>
<tr>
<td>$C = H$</td>
<td>63.3 ± 12.4</td>
<td>23.3 ± 12.5</td>
<td>10.0 ± 4.4</td>
</tr>
<tr>
<td>$C = B$</td>
<td>15.0 ± 4.6</td>
<td>5.0 ± 3.6</td>
<td>80.0 ± 5.4</td>
</tr>
<tr>
<td>Perspective 3</td>
<td>95 ± 5.0</td>
<td>5 ± 5.0</td>
<td>0 ± 0</td>
</tr>
<tr>
<td>$C = H$</td>
<td>60 ± 16.3</td>
<td>30 ± 15.3</td>
<td>0 ± 0</td>
</tr>
<tr>
<td>$C = B$</td>
<td>0 ± 0</td>
<td>0 ± 0</td>
<td>100 ± 0</td>
</tr>
</tbody>
</table>

![Table 3](https://example.com/table3.png)

<table>
<thead>
<tr>
<th>Class</th>
<th>Recall</th>
<th>Precision</th>
</tr>
</thead>
<tbody>
<tr>
<td>Healthy</td>
<td>95% ± 5%</td>
<td>76% ± 8.7%</td>
</tr>
<tr>
<td>Bronchiectasis</td>
<td>30% ± 15.3%</td>
<td>75% ± 25%</td>
</tr>
<tr>
<td>Interstitial pulm. dis.</td>
<td>100% ± 0%</td>
<td>100% ± 0%</td>
</tr>
</tbody>
</table>
Therefore, another reason that the third perspective is more meaningful is that it is based on combining the six flow-phase decisions to reach a stronger decision about the subject. It can further be observed in Table 1 and Table 2 that the success rates are improved in general as one proceeds from the more detailed perspective (the first) towards the more general (the third) by combining the decisions, the third case being the most meaningful clinically.

The success rates imply that the interstitial pulmonary disease class is distinctly located in the feature space, whereas the bronchiectasis class is scattered in the vicinity of the healthy class, possibly overlapping the healthy region. This seems to be the main reason behind the misclassification of bronchiectasis instances in the one-against-one comparisons of the multi-class SVM setup. The results are promising for a binary classification (healthy versus pathological) but additional features related to adventitious sounds are needed to improve the classification of pulmonary sounds belonging to bronchiectasis cases. Moreover, other classifier algorithms and setups may be considered to improve the performance of the method. If the lowest of the reported sensitivity, specificity, precision and recall rates can be raised above at least 90%, the results may be accepted to be successful and the proposed method may be usable in clinical environments.

Acknowledgments
This project is supported by Bogazici University Research Fund under Project Numbers 09A203D and 13A02P2. We thank Sibel Yurt, MD, from Yedikule Chest Disease and Thoracic Surgery Education and Research Hospital for her guidance and advices on data acquisition and diagnosis of pulmonary disorders of patients.

References