Investigation of Lung Sounds Features for Detection of Bronchitis and COPD Using Machine Learning Methods

Porieva H. S., Ivanko K. O., Semkiv C. I., Vaityshyn V. I.

National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”

E-mail: *porevanna@gmail.com, koondoo@gmail.com*

The study is dedicated to the issue of investigation of the lung sounds digital analysis processing methods, searching for new informative features of pathological respiratory sounds and using machine learning methods for classifying the state of the bronchopulmonary system. In particular, the use of various methods of lung sounds analysis is considered, namely: frequency, time-frequency, wavelet, and mel-frequency cepstral analysis. The application of signal processing methods to the problem of respiration signals analysis is considered in the paper. In order to investigate the possibilities of machine learning to the problem of classification of respiration signals, the dataset of lung sounds of 296 recordings, which represent 3 classes: norm, bronchitis, and chronic obstructive pulmonary disease, was used in this work. The purpose of this study is to identify and compare the informative features of the lung sounds obtained with different signal processing methods, as well as to choose the classification method that provides the highest accuracy in the identification of the bronchopulmonary system condition. To obtain frequency features, power spectrum density dependence on frequency was calculated for respiratory signals using fast Fourier transform method. The spectral measures, as well as ratios of spectrum powers in different frequency bands, were defined. To extract the spectral-temporal features of the lung sounds, the spectrograms of the analyzed signals were investigated. The mean time dependences of the power spectral density in the indicated frequency ranges were determined. The sum of magnitudes values of the power spectrum curve for each frequency band was used as the features obtained from the spectrogram. The ratios of the energies corresponding to detail levels of the wavelet decomposition to the total energy of the decomposed signal were used as the parameters of signal recognition based on wavelet analysis. The logarithmic (mel) filterbank energies, averaged over time frames, depending on channel index and time, as well as mel frequency cepstrum depending on cepstrum index and time, are proposed to use as features derived from mel-cepstral analysis. The supervised machine learning based on decision trees, discriminant analysis, support vector machines, logistic regression, k-nearest neighbors classifiers, and ensemble learning were applied to determine the best classification models for computerized disease screening. The accuracy of the different classifiers using these feature sets was determined and compared. Based on this, a combination of features and classifiers, which provides the highest accuracy of lung condition recognition, reaching 93%, is proposed.

Key words: lung sounds; bronchitis; chronic obstructive pulmonary disease; spectral wavelet decomposition; mel-frequency cepstral analysis; machine learning

DOI: 10.20535/RADAP.2021.84.78-87

Introduction

Respiratory diseases are a huge global health burden. It is estimated that in 2019, 235 million people had asthma, more than 200 million people had chronic obstructive pulmonary disease (COPD), 65 million patients had moderate to severe COPD, 1-6% of the adult population (over 100 million people) had respiratory problems during sleep, 8.7 million people get suffer from tuberculosis each year, millions of people live with pulmonary hypertension, and more than 50 million people struggle with occupational lung diseases, totaling more than 1 billion people with chronic respiratory disease [1, 2].

But 2020 year made significant adjustments to these statistics, making it much worse. The global catastrophe with the spread of COVID-19 posed new tasks and challenges to humanity. Virologists are searching for the development of a reliable vaccine every day, and humanity hopes to solve this problem as soon as possible. Unfortunately, many other problems arose in parallel. Patients who have suffered and been cured are needed further monitoring of their respiratory condition, as it is not known what exactly complications may occur after such a severe illness. Since the listed diseases are supplemented by the usual seasonal colds, which may be accompanied to one degree or another by lung disease, the problem of early diagnosis is very urgent in our time. The joint work of researchers, engineers and doctors to find a convenient and reliable tool for diagnosing and monitoring lung diseases is currently a promising and urgent task. Also, with such a flow of lung diseases, it is important to have an instrument that can quickly classify the state of the bronchopulmonary system with high accuracy according to certain categories [3]. Machine learning is increasingly being used for this purpose [4–7].

The vast majority of respiratory diseases are accompanied by various disorders of air movement through the respiratory system, which lead to the appearance of distinctive noises (sounds). Despite the development of technical diagnostic tools, auscultation, which is listening to the sounds of breathing, remains the most common non-invasive method of diagnosing respiratory diseases [8].

The sensitivity (threshold of audibility) of the human hearing organ and its ability to distinguish sounds by volume and frequency vary significantly from individual to individual. In addition, due to the peculiarities of the human organ of hearing at high volume, high-frequency sounds subjectively seem louder than low-frequency sounds. At the same time, "sound memory", talent and training of a doctor are extremely important for the auditory analysis of breathing sounds. For the average doctor, memorizing and analyzing all the nuances of such complex and highly informative signals, such as noises and wheezing, is a difficult task [9, 10].

Over the last 50-60 years, serious research work has been carried out to study the possibility of recording, visualization and classification of respiratory sounds based on instrumentalities and methods of electronic technology. The efforts of research centers around the world are coordinated by the International Lung Sounds Association [11].

The advantages of using electronic auscultation technology are obtaining high-quality audio signals regardless of the ability of the hearing organ of the doctor, the ability to repeatedly listen and compare the recorded signal with samples or later recordings, for example, in the recovery process. Due to the ability to create databases of breathing sounds, it is now possible to exchange relevant samples between research centers and learn from a large number of samples. Finally, such a system creates the preconditions for solving telemedicine problems, as the received signals can be accumulated and processed remotely, including the use of mobile communications [12–14].

A large number of numerically diagnostically valuable parameters can be obtained from the recorded lung sound signal by means of digital processing and analysis. The automated recognition of respiratory noise types can be applied to recorded respiratory sounds. Analysis of respiratory sounds by various methods provides a large number of parameters, which can be difficult for unambiguous perception by a doctor. Therefore, an important task is to classify lung sounds into certain categories. This problem can be solved by creating systems for identifying and classifying lung sound parameters that will help the doctor in the diagnosis process.

Such systems can be used for mass monitoring and screening of the population to detect respiratory pathologies without the use of the X-ray or computed tomography (CT) methods and thus reduce radiation exposure and congestion in CT labs.

Many technologies and mathematical approaches are currently used for digital analysis of lung sounds. Among these methods spectral analysis, spectral-time analysis, correlation analysis, and wavelet transform have gained wide popularity. Each of the methods has its own disadvantages and advantages. Unfortunately, a lot of the approaches, as a rule, are directed for a specific task: either for certain diseases, or for a database. Many of the methods require special signal preprocessing. Therefore, the search for a universal analysis method capable of giving high results for a wide range of diseases and data sets is an urgent task [15–18].

In recent years, mel-frequency cepstral analysis has become increasingly popular among tasks for processing human sounds signals, for example, voice, cardiac sound [19] or even some types of lung sounds [20]. Thus, the use of this method is promising for application to audio signals of the human body.

1 Materials and methods

For assessing the efficiency of the methods for lung diseases detection, a database CORA provided by the Institute of Hydromechanics was used in this research [21, 22]. The dataset of lung sounds consists of 296 recordings with sampling frequency of 3496 Hz and duration of 18 seconds. According to literary sources [23, 24], the main informative part of the lung sounds spectrum is in the range from 100 to 1500 Hz. Therefore, the sampling rate that was used when recording the signals is sufficient. If the technical characteristics of other recording devices are different, it is advisable to oversample the signals using bandpass filters for a given frequency range. In the used database, classes of recordings can be distinguished as: class 1- recordings of lung sounds in norm (112 signals), class 2 - lung sounds of patients suffering from bronchitis (84 signals), class 3 - lung sounds of patients with chronic obstructive pulmonary disease (100 signals). With this ratio of signals in classes, the sample can be considered balanced.

To extract the features from frequency domain, power spectrum density (PSD) dependence on frequency was calculated for respiratory signals using fast Fourier transform method. The signal mean value was subtracted from the investigated signals to avoid taking into account the contribution of the signal at zero frequency. Non-periodic symmetric Hamming window was applied in order to minimize the effect of spectral leakages. Four frequency bands were used for feature extraction: high frequencies (HF) – from 1000 to 1500 Hz, mid frequencies (MF) – from 500 to 1000 Hz, low frequencies (LF) – from 200 to 500 Hz, and very low frequencies (VLF) – from 100 to 200 Hz. We did not utilize the band from 0 to 100 Hz to characterize respiratory signals due to the significant influence of the heart sounds in this frequency band.

The following 8 spectral measures were calculated for each signal: normalized power in high, mid, low and very low frequencies ranges, ratios of spectrum powers in different frequency bands P_{VLF}/P_{HF}, P_{MF}/P_{HF}, P_{LF}/P_{HF}, P_{VLF}/P_{MF}, P_{VLF}/P_{LF}.

To extract the spectral-temporal features of the lung sounds in norm, bronchitis, and chronic obstructive pulmonary disease, the spectrograms of the analyzed signals were calculated (Fig. 1). The settings of the spectrogram were defined as follows: hamming window of duration 2 ms, 50% windows overlap (1 ms step) and 1 Hz step for frequency in range from 100 to 1500 Hz.

The submatrices were extracted from the spectrogram in order to define spectral-temporal features in HF, MF, LF, VLF ranges. The mean time dependences of PSD in mentioned frequency ranges, obtained by averaging all the values in the taken frequency range corresponding to the current time moment, were defined. As features for lungs condition recognition, derived from the spectrogram, the sum of magnitudes values of the obtained curve was used for each frequency band that gave us 5 features.

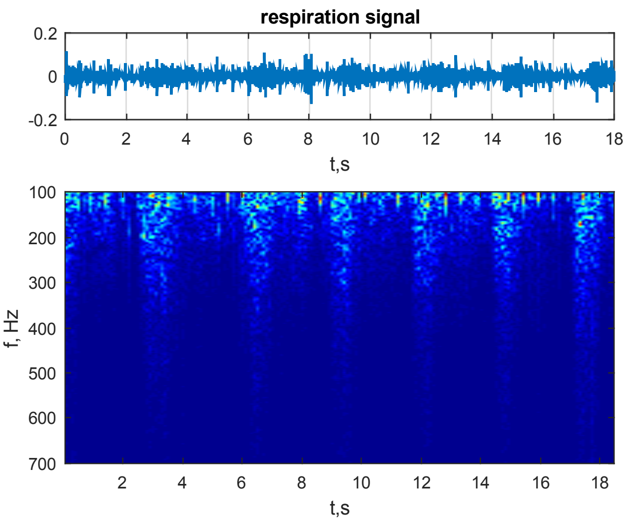


Fig. 1. Spectrogram calculated for a respiration signal of a patient suffering from bronchitis

To recognize the signs of pathological changes in respiratory signals, the features of *multilevel wavelet transform* (MWT) were also defined. The respiratory signal can be represented by MWT as a sum of an approximation component a_n and the detail components d_i:

$$S = a_n + \sum_{i=1}^{n} d_i \,,$$

where n – the number of decomposition levels.

The investigated signals were decomposed using symmetrical wavelet function of the 5-*th* order up to the 5-*th* level of wavelet transform (Fig. 2).

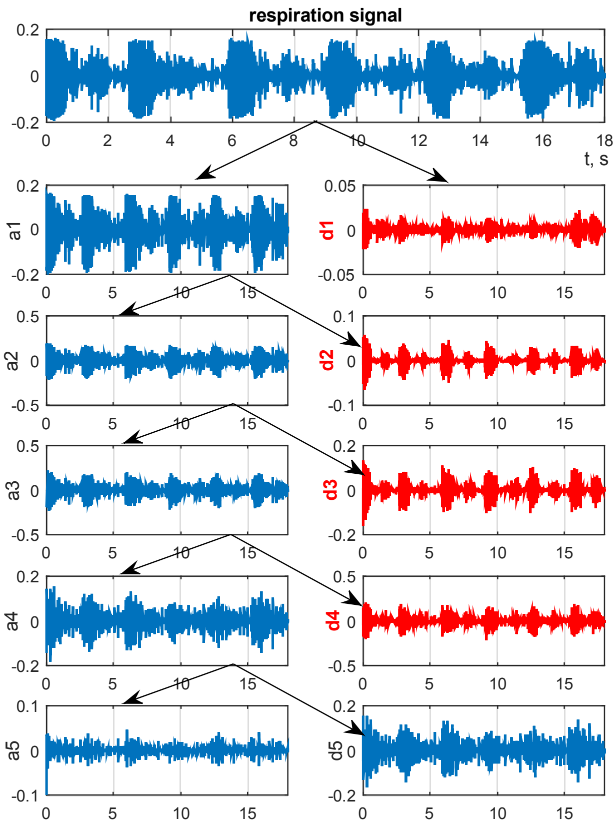


Fig. 2. Wavelet decomposition up to the 5-*th* level (with a "symmetric" wavelet of the 5-*th* order) performed for a respiration signal in norm. The details $d_1 - d_4$ representing the frequency regions of interest are shown in red color

The analyzed signals are sampled at 3496 Hz and have maximum informative frequency content till 1748 Hz. The series of wavelet-based highpass and lowpass filters repeatedly divide the input frequency range. Consequently, the detail component d_1 represent the most high-frequency part of the signal in the range from 874 to 1748 Hz. The detail component d_2 reflects frequency content from 437 to 874 Hz; d_3 corresponds to the subband from 218,5 to 437 Hz; and d_2 frequency content lies from 109,25 to 218,5 Hz.

The approximation a_5 and the detail component d_5 together capture the most low frequency components of respiratory signals, which are below 109,25 Hz. We did not use these components for feature extraction of the lung sounds in order to avoid the heart sounds influence.

The spectral parameters originated from the details d_1–d_4 were calculated to distinguish between signals in norm, bronchitis, and chronic obstructive pulmonary disease. Power spectrum density dependence on frequency was calculated for each wavelet component a_5, d_5, d_4, d_3, d_2, and d_1, using fast Fourier transform.

Then the total energy of each component was defined as sum of all magnitudes in PSD. As parameters for signal recognition, we used the ratio of the energy of each detail d_4, d_3, d_2, and d_1 to the total energy of the decomposed signal, which can be defined as the sum of energies of all the components of wavelet decomposition:

$$R_{d_j} = \frac{P_{d_j}}{\left(P_{a_5} + \sum_{i=1}^{5} P_{d_i}\right)},$$

where $j=1,2\ldots4$. Thus, we got 4 spectral parameters R_{d_1}, R_{d_2}, R_{d_3}, R_{d_4}, which reflect the contribution of each detail component d_1-d_4 to the total signal energy.

Cepstral features of lung sounds were also used to distinguish normal and abnormal classes. The Mel scale correlates the perceived frequency of the sound by human hearing (pitch of the pure tone) with the actual measured frequency (Hz). This dependence is nonlinear and is described by the following equation: $M(f)=1127 * \ln(1+f/700)$.

To calculate the mel frequency cepstral coefficients (MFCC), the respiratory signal is divided into the frames. The duration of a frame affects the results of the analysis and should be chosen based on the assumption that the signal does not change its behavior significantly over the duration of the frame. To define MFCC, the next steps are applied to each frame. As the recorded respiratory signal is finite and not periodic, the effect of leakage occurs when applying the Fourier transform due to the gaps at the end points of the signal. In order to reduce this effect, each frame is multiplied by the Hamming window function. The discrete Fourier transform is applied to the result and the periodogram for each frame is calculated. Next, the set of mel filters is calculated. Triangular filters are multiplied by the periodogram and summed. The number of triangular filters also affects the results of the cepstral analysis. The energy of a set of filters is obtained, which is then logarithmized. Logarithm process is performed to smooth the primary spectrum and reduce the number of parasitic components in the cepstrum. Finally, using discrete cosine transform, mel cepstral coefficients are obtained. Filters overlap and the filter energies are fairly correlated. Discrete cosine transform decorrelates them.

Our anticipation was that cepstral analysis provides information about the features of lung sounds unapproachable to spectral or spectral-temporary analysis. The application of mel frequency cepstral analysis to the lung sounds investigation is justified, because the spectrum is projected on a mel-scale, allowing to select the most important sound frequencies. Moreover, the method is largely insensitive to changes in the phase of the studied signals.

As the features for recognition of lung diseases using machine learning, *log (mel) filterbank energies* depending on channel index and time as well as mel frequency cepstrum depending on cepstrum index and time were used (Fig. 3). However, for breath sounds recorded at a sufficiently high sampling rate, such data matrices contain thousands of values. Therefore, in this work, we used the corresponding values of these parameters averaged over time frames.

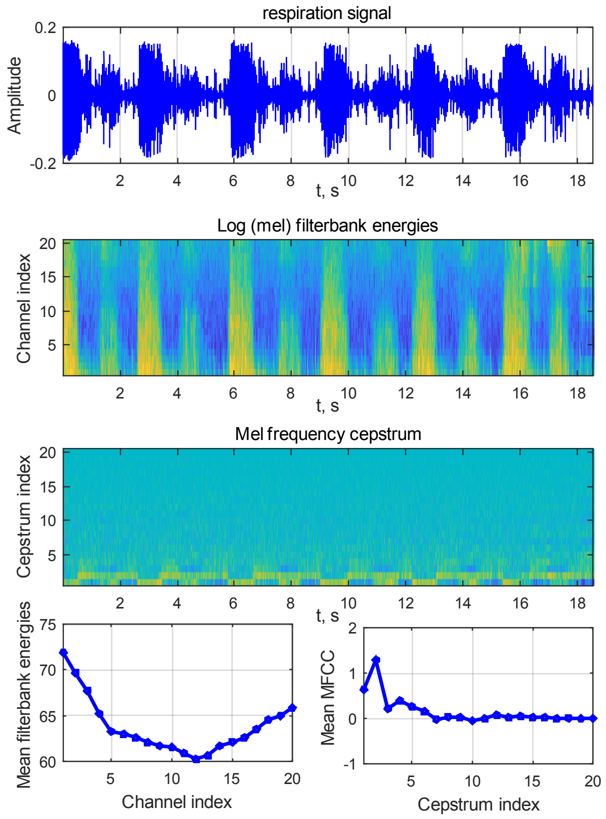


Fig. 3. Cepstral analysis of the respiration signal in norm

2 Application of Machine Learning Methods to Lung Sounds Classification

To determine the best classification models for computerized bronchitis and chronic obstructive pulmonary disease screening, we implemented supervised machine learning based on decision trees, discriminant analysis, support vector machines (SVM), logistic regression, k-nearest neighbors (KNN) classifiers, and ensemble learning.

The Decision Tree Classifier is the most common and widely used machine learning algorithm that performs regression and classification tasks. The classifier divides the data into smaller subsets based on different criteria, i.e. each subset has its own ordered category. With each step, the number of objects of a certain criterion decreases. The classification ends, when the network reaches a subset with only one object that contains the forecast or result of the decision trees. To choose the best approach, we used three options for decision trees: 1) coarse tree with maximum number of splits equal to 4; 2) medium tree with maximum number of splits equal to 20; 3) fine tree with maximum number of splits equal to 100.

The k-Nearest Neighbor (kNN) algorithm is one of the simplest machine learning algorithms. To make a prediction for a new data sample, the kNN algorithm finds the training set closest to it, i.e. finds its "nearest neighbors". In the simplest case, the k-nearest neighbor algorithm considers only one nearest neighbor - the point of the training set closest to the point for which we want to get a forecast. When more than one neighbor is taken into account, the most common class is used to assign a label, i.e. the class that has gained the majority among the k-nearest neighbors is selected. In the case of a multi-class classification, as in our case with 3 classes ("norm", "bronchitis", and "chronic obstructive pulmonary disease"), the number of neighbors belonging to each class is counted and the most common class is predicted. The kNN classifier considered two important parameters: the number of neighbors and the measure of the distance between data points. The results are also affected by the size of the training data sample. To define the best options, we utilized and compared 6 customizations for kNN classifier: 1) fine KNN with number of neighbors equal to 1 and euclidean distance with equal weights; 2) medium KNN with 10 neighbors and euclidean distance with equal weights; 3) coarse KNN with 30 neighbors and euclidean distance with equal weights; 4) cosine KNN with 10 neighbors and cosine distance with equal weights; 5) cubic KNN with 10 neighbors and Minkowski (cubic) distance with equal weights; 6) weighted KNN with 10 neighbors and euclidean distance with squared inverse distance weights.

Classification method of discriminant analysis assumes that different classes generate data based on different Gaussian distributions, which are estimated by the fitting function to train a classifier. To predict the classes of new data, the trained classifier finds the class with the smallest misclassification cost. We used linear and quadratic discriminant machine learning methods.

The Support Vector Machine (SVM) method is a powerful machine learning method that has shown good results in many biomedical applications. Using a set of training data, SVM method finds the hyperplane that best separates the two classes of training data. Such a hyperplane is a boundary that has the maximum distance from different classes of training data. Solution boundary maximizes the distance from the nearest data points of all classes. The nearest points to the decision surface are called support vectors. To define the best solution, we used 6 options for SVM classifier: 1) linear kernel function; 2) quadratic kernel function; 3) cubic kernel function; 4-6) Gaussian kernel function (with kernel scales 1,1; 4,5; and 18).

The implemented ensemble learning algorithms included bagging, subspace, and boosting algorithms. The number of learners was equal to 30.

Table 1 provides the results with the classification accuracy percentage (% of correctly identified cases). The given values correspond to the total percentage of correctly determined signals. The percentages of correctly determined cases for each class "norm", "bronchitis", and "chronic obstructive pulmonary disease" are also indicated separately.

Since the initial database doesn't contain very large number of signals, it is advisable to use cross-validation approach for accuracy estimation with a commonly used data splitting ratio of 80% training data and 20% testing data. To assess the accuracy of the machine learning algorithms in lung sounds classification, five-fold cross-validation approach was used. The data were divided into five folders: four folders repeatedly were used for training, and one was used as a testing folder. As each of the five folders once was used for testing, it contributed to average classification accuracy of the machine learning method.

After analyzing the data obtained as a result of machine learning, we concluded that for the majority of the classifiers, class "chronic obstructive pulmonary disease" was well recognized. The most often the wrong decisions were made when recognizing the class "bronchitis", which very quite was classified as "norm". Therefore, in the context of medical diagnostics, the overall accuracy of machine learning algorithm is not enough for its performance estimation. It is preferably to have low false negative rates, i.e. to obtain the small number of patients suffering with bronchitis or COPD, who are tested and recognized as healthy. The opposite case, when healthy person is mistakenly assigned to the class corresponding to the presence of the lung disease, is not so objectionable, because additional investigations can discard the false positive results.

The first feature vector presented in Table 1 contains *8 spectral features* yielded from the fast Fourier transform method. A preliminary analysis of the spectrum of signals from different classes showed that the spectra of signals of lung sounds in normal and pathological conditions significantly intersect in the frequency domain. Therefore, as we expected, machine learning using spectral characteristics did not show acceptable performance. Fine KNN appeared to show the best results – 75.7% of overall accuracy, 82% of correctly defined signals in norm, 71% of detected patients with bronchitis, and 72% of correctly identified COPD signals.

Table 1 Comparison of machine learning algorithms performance: total classification accuracy (%) and true positive rate for classes "norm", "bronchitis", and "chronic obstructive pulmonary disease" in parentheses

#	Machine learning method	Spectral features-8	Spectrogram features -5	Wavelet features -4	Cepstral features -20	logFBEs - 20	Combined (MFCC, logFBEs) -40	Combined (logFBEs, wavelet) -24
1	Coarse Tree (max 4 splits)	67.6 (66, 64, 72)	66.6 (77, 51, 68)	86.1 (92, 62, 100)	75.7 (82, 62, 80)	73 (79, 61, 77)	74.7 (83, 60, 78)	85.5 (87, 67, 100)
2	Medium Tree (max 20 splits)	71.6 (78, 61, 74)	72.6 (72, 63, 81)	82.8 (80, 65, 100)	80.1 (79, 70, 90)	80.1 (83, 67, 88)	79.4 (82, 69, 85)	85.5 (84, 70, 100)
3	Fine Tree (max 100 splits)	72.6 (72, 71, 74)	70.9 (69,61,82)	83.1 (80, 67, 100)	80.1 (79, 70, 90)	81.1 (84, 69, 88)	79.4 (82, 69, 85)	85.5 (84, 70, 100)
4	Linear Discriminant	67.2 (69, 57, 74)	65.5 (54, 50, 91)	84.1 (88, 60, 100)	-	84.1 (80, 73, 98)	-	88.5 (89, 74, 100)
5	Quadratic Discriminant	-	60.5 (92, 51, 33)	80.1 (88, 51, 96)	-	92.9 (90, 93, 96)	-	93.2 (89, 90, 100)
6	Linear SVM linear kernel function	68.9 (63, 63, 80)	72 (64, 54, 96)	82.8 (88, 56, 100)	85.1 (88, 71, 94)	79.7 (79, 65, 93)	86.1 (89, 71, 95)	85.8 (88, 67, 100)
7	Quadratic SVM (quadratic kernel function)	65.9 (71, 55, 69)	74.3 (76, 52, 91)	83.8 (86, 62, 100)	91.6 (96, 80, 96)	87.2 (92, 73, 94)	91.9 (98, 80, 95)	86.8 (87, 71, 100)
8	Cubic SVM (cubic kernel function)	73.6 (79, 70, 70)	64.9 (62, 56, 76)	80.4 (81, 56, 100)	90.2 (98, 75, 94)	88.2 (96, 76, 89)	89.2 (96, 74, 94)	89.5 (92, 74, 100)
9	Fine Gaussian SVM (kernel scale 1,1)	73.6 (76, 69, 75)	70.3 (62, 61, 88)	86.8 (85, 74, 100)	61.8 (66, 52, 65)	82.8 (87, 68, 91)	81.4 (89,49 100)	88.9 (86, 80, 100)
10	Medium Gaussian SVM (kernel scale 4,5)	70.3 (65, 62, 83)	63.9 (45, 49, 98)	86.1 (90, 64, 100)	88.2 (91, 81, 91)	76 (80, 64, 81)	90.9 (96, 82, 93)	85.8 (85, 70, 100)
11	Coarse Gaussian SVM (kernel scale 18)	66.2 (61, 55, 82)	52 (93, 38, 18)	84.1 (92, 55, 100)	79.7 (95, 63, 77)	64.9 (79, 57, 55)	75.7 (91, 54, 77)	80.4 (88, 54, 94)
12	Fine KNN(number of neighbors – 1, euclidean distance, equal weight)	75.7 (82, 71, 72)	80.7 92, 62, 84)	90.5 (96, 73, 100)	92.2 (100,83,91)	87.2 (98, 65, 93)	92.9 (100,86,91)	90.9 (95, 75, 100)
13	Medium KNN(number of neighbors –10, euclidean distance, equal weight)	66.2 (71, 57, 68)	73.3 (81, 57, 78)	86.1 (88, 68, 100)	80.4 (91, 67, 80)	75 (81, 56, 84)	80.4 (94, 65, 78)	83.1 (90, 54, 100)
14	Coarse KNN(number of neighbors –30, euclidean distance, equal weight)	61.8 (66, 52, 65)	55.7 (53, 40, 72)	83.4 (93, 51, 100)	65.9 (99, 23, 65)	64.2 (75, 61, 55)	63.5 (89, 39, 55)	67.2 (81, 58, 59)
15	Cosine KNN(number of neighbors –10, cosine distance, equal weight)	66.2 (64, 63, 71)	73 (81, 50, 83)	85.1 (85, 68, 100)	84.1 (84, 80, 88)	76.7 (67, 75, 89)	80.7 (73, 77, 92)	84.1 (84, 65, 100)
16	Cubic KNN(number of neighbors – 10, Minkowski (cubic) distance, equal weight)	65.2 (70, 58, 66)	74.3 (82, 56, 81)	86.1 (88, 58, 100)	81.1 (92, 67, 81)	77 (83, 60, 85)	81.4 (92, 68, 81)	82.8 (88, 55, 100)
17	Weighted KNN (number of neighbors – 10, euclidean distance, squared inverse distance weight)	74.7 (86, 63, 72)	81.4 (94, 61, 85)	90.9 (97, 71, 100)	86.1 (100, 69, 85)	83.4 (96, 60, 89)	87.5 (100, 71, 87)	87.5 95, 63, 100)
18	Boosted Trees (maximum number of splits - 20, number of learners - 30)	74.7 (84, 65, 72)	76.7 (88, 55, 82)	87.2 89, 69, 100)	75.3 94, 58, 69)	85.8 94, 68, 92)	87.5 96, 76, 88)	67.9 (92, 45, 60)
19	Bagged Trees(maximum number of splits - 295, number of learners - 30)	75.3 (84, 67, 73)	83.4 96, 65, 84)	88.5 (95, 67, 100)	88.5 (95, 75, 93)	86.5 (96, 67, 92)	87.8 (93, 76, 92)	89.5 (95, 70, 100)
20	Subspace Discriminant (maximum number of splits - 20, number of learners - 30)	68.6 (71, 58, 75)	65.9 80, 51, 62)	84.5 (88, 61, 100)	81.1 (85, 73, 84)	81.8 (78, 70, 96)	86.5 (88, 73, 96)	86.1 (87, 69, 100)
21	Subspace KNN (maximum number of splits - 20, number of learners - 30)	75 (84, 63, 75)	73.3 (90, 58, 67)	87.8 (92, 68, 100)	82.4 (94, 70, 80)	87.5 (98, 68, 92)	89.2 (100,73,91)	87.2 (96, 70, 91)

Using *5 spectrogram features* allowed us to obtain higher total accuracy of lung sounds classification comparing to the PSD based features – 83.4% for ensemble learning algorithm of bagged trees, which performed the best. However, despite the high accuracy of recognition of healthy control signals (96%), the spectral-temporal characteristics gave a very poor result for the recognition of bronchitis – 65%, and 84% of COPD signals were identified correctly.

Analyzing the machine learning results obtained for *4 features derived from wavelet transform*, we can see an interesting regularity: we still have problems with "bronchitis" class recognition, but majority of machine learning methods unmistakably recognized class "chronic obstructive pulmonary disease" (Fig. 4).

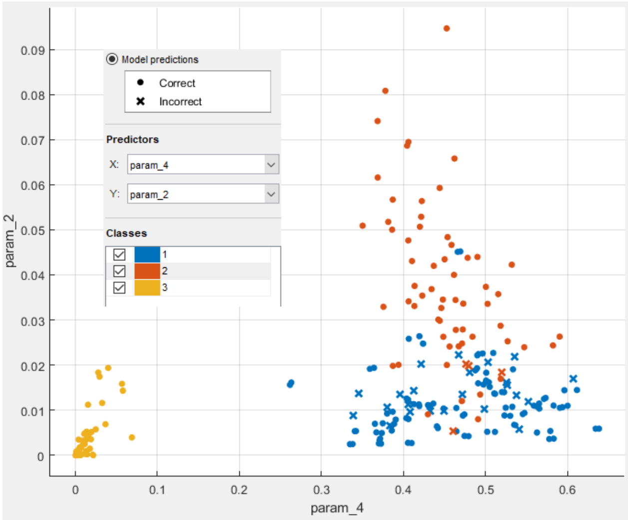
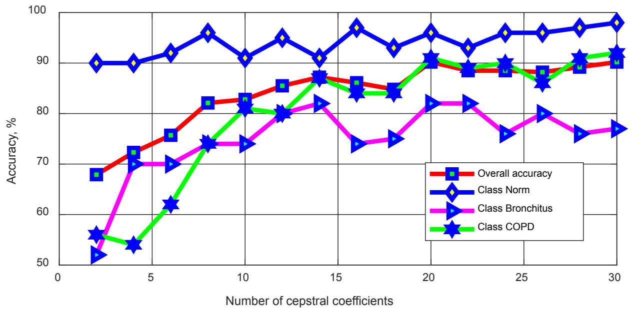


Fig. 4. Machine learning results obtained for 4 features derived from wavelet transform (trained model using weighted KNN algorithm)

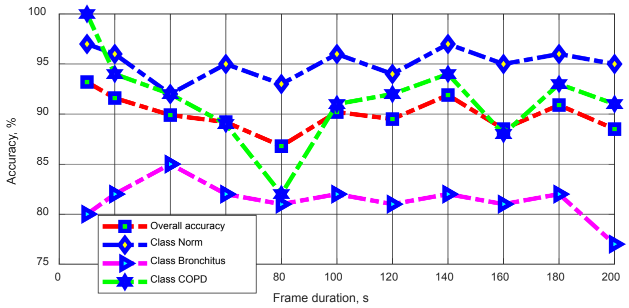
The best achieved total accuracy of classification was about 91% with using weighted KNN algorithm, trained using 10 nearest neighbors and euclidean distance with squared inverse distance weight. True positive rates for classes "norm", "bronchitis", and "chronic obstructive pulmonary disease" reached 97%, 71%, and 100% respectively.

It is obvious that the selected wavelet parameters, which reflect the contribution of each frequency-related detail component $d_1 - d_4$ to the total signal energy, perfectly convey the features of lung sounds in COPD compared to the norm. Therefore, it makes sense to combine wavelet features, for example, with cepstral coefficients, in order to increase the recognition accuracy when classifying lung sounds.

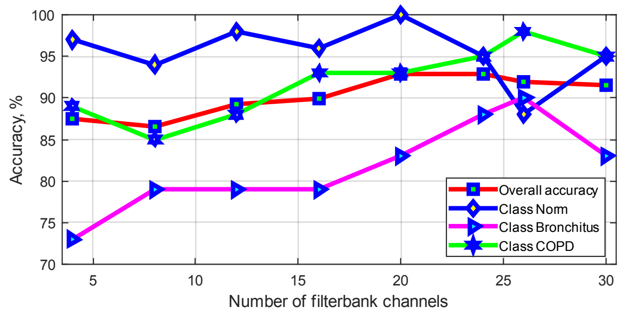
The results of *cepstral analysis* significantly depend on a number of parameters, among which are frame duration, frame shift, number of filterbank channels, number of cepstral coefficients, as well as lower and upper frequency limits. To find the optimal set of parameters for calculating the cepstral characteristics of breath sounds, we changed each of the parameters with a fixed set of other parameters and performed machine learning (Fig. 5). We selected the parameter values that provided the highest accuracy in classifying signals into 3 classes: norm, bronchitis and chronic obstructive pulmonary disease. Especially we paid attention to the possibility of distinguishing signals in norm and bronchitis, because the accuracy of detecting bronchitis was the lowest compared to other classes.



(a)



(b)



(c)

Fig. 5. The results of machine learning performance depending on cepstral parameters: number of cepstral coefficients (a), frame duration (b), and number of filterbank channels (c)

Finally, we selected the following set of parameters: frame duration T_f=20 ms, frame shift T_{sh}=10 ms, number of filterbank channels N_{ch}=20, number of cepstral coefficients N_{mfcc}=20, lower frequency limit F_L=100 Hz, and upper frequency limit F_U=1500 Hz.

Mel frequency cepstrum depending on cepstrum index averaged over time frames gave us 20 cepstral features. The highest achieved model accuracy based on these features was about 92% using KNN classifier with 1 nearest neighbor and euclidean distance with equal weight. In this case, absolutely all signals of the class "norm" were classified without errors, bronchitis was correctly identified in 83% cases and COPD was identified correctly in 91% cases.

Log (mel) filterbank energies depending on channel index averaged over time frames produced 20 energy features. Their using for models machine learning demonstrated promising results. The highest total accuracy of classification was 93% using the model build with quadratic discriminant method. True positive rates for classes "norm", "bronchitis", and "chronic obstructive pulmonary disease" were 90%, 93%, and 96% correspondingly. It should be noted that using these features allowed us to significantly increase the recognition accuracy of signals from the class of bronchitis. As it was mentioned above, bronchitis class was recognized worse than "norm" and COPD when using all other features, although the correct identification of bronchitis from the point of view of diagnosis is more important than the overall classification accuracy.

We also considered combining two types of discussed above cepstral features to build the models for classification. This variant contained 40 features. Moreover, we combined wavelet derived features with log (mel) filterbank energies depending on channel index averaged over time frames, which yielded 24 features. The total classification accuracy in these cases it turned out to be very close in its values – near 93%. The difference was in the redistribution of the accuracy value of identifying different classes: norm, bronchitis and COPD. Therefore, the doctor can choose the right model depending on the prevalence of the patient's morbidity.

Conclusion

In this study the signal processing methods for analysis of lung sounds and the possibility of using machine learning approach to perform diagnosis of bronchitis and chronic obstructive pulmonary disease, are investigated.

The best results were obtained for features of lung sounds derived from log (mel) filterbank energies depending on channel index averaged over time frames. The highest total accuracy of lung condition recognition reached 93% using the model based on quadratic discriminant method. True positive rates for classes "norm", "bronchitis", and "chronic obstructive pulmonary disease" in this case reached 90%, 93%, and 96% correspondingly. This result is also one of the best from the point of view of balancing the values of the correctly identified classes. This is especially important for recognizing the class of bronchitis, which was very poorly detected by most other methods and models.

Also combining of cepstral and wavelet features demonstrated the promising results. The most accurate models for classifying lung sounds were obtained using KNN classifier variations, as well as quadratic discriminant method. The proposed solutions will be useful for monitoring pulmonary state in patients suffering from bronchitis and COPD, as well as for routine scheduled medical examinations.

References

[1] Lung disease & information. *European Lung Foundation*.

[2] The top 10 causes of death. (2020). *World Health Organization*.

[3] Kumar S. S. (2020). Emerging Technologies and Sensors That Can Be Used During the COVID-19 Pandemic. *International Conference on UK-China Emerging Technologies (UCET)*, Glasgow, UK, pp. 1-. DOI: 10.1109/UCET51115.2020.9205424.

[4] GÜLER H. C. et al. (2020). Classification of Abnormal Respiratory Sounds Using Machine Learning Techniques. *2020 Medical Technologies Congress (TIPTEKNO)*, Antalya, pp. 1-4. DOI: 10.1109/TIPTEKNO50054.2020.9299294.

[5] Chekhovych M. G., Poreva A. S., Timofeyev V. I. and Henaff P. (2018). Using of the machine learning methods to identify bronchopulmonary system diseases with the use of lung sounds. *Visnyk NTUU KPI, Seriia Radiotekhnika Radioaparatobuduvannia*, No. 73, pp. 55–62. DOI: 10.20535/RADAP.2018.73.55-62.

[6] Serato J. H. L. and Reyes R. (2018). Automated lung auscultation identification for mobile health systems using machine learning. *2018 IEEE International Conference on Applied System Invention (ICASI)*, Chiba, pp. 287-290. DOI: 10.1109/ICASI.2018.8394589.

[7] Paraschiv E.-A. and Rotaru C.-M. (2020). Machine Learning Approaches based on Wearable Devices for Respiratory Diseases Diagnosis. *2020 International Conference on e-Health and Bioengineering (EHB)*, Iasi, Romania, pp. 1-4. DOI: 10.1109/EHB50910.2020.9280098.

[8] Sarkar M., Madabhavi I., Niranjan N., Dogra M. (2015). Auscultation of the respiratory system. *Annals of Thoracic Medicine*, Vol. 10, Iss. 3, pp. 158–168. DOI: 10.4103/1817-1737.160831.

[9] Gottlieb E. R., Aliotta J. M., Tammaro D. (2018). Comparison of analogue and electronic stethoscopes for pulmonary auscultation by internal medicine residents. *Postgraduate Medical Journal*, Vol. 94, Iss. 1118. DOI: 10.1136/postgradmedj-2018-136052.

[10] Ellington L. E., Gilman R. H., Tielsch J. M., et al, (2012). Computerised lung sound analysis to improve the specificity of paediatric pneumonia diagnosis in resource-poor settings: protocol and methods for an observational study. *BMJ Open*, Vol. 2, Iss. 1. DOI: 10.1136/bmjopen-2011-000506.

[11] Pasterkamp H., Brand P. L. P., Everard M., Garcia-Marcos L., Melbye H., Priftis K. N. (2016). Towards the standardisation of lung sound nomenclature. *European Respiratory Journal*, Vol. 47, Iss. 3, pp. 724-732. DOI: 10.1183/13993003.01132-2015.

[12] Arts L., Lim E.H.T., van de Ven P.M., et al. (2020). The diagnostic accuracy of lung auscultation in adult patients with acute pulmonary pathologies: a meta-analysis. *Scientific Reports*, Vol. 10, Article number: 7347. DOI: 10.1038/s41598-020-64405-6.

[13] Aras S., Öztürk M., Gangal A. (2016). Endpoint detection of lung sounds for electronic auscultation. *2016 39th International Conference on Telecommunications and Signal Processing (TSP)*, Vienna, Austria, pp. 405-408. DOI: 10.1109/TSP.2016.7760907.

[14] Tolnai J., Kapus K., Draskóczy M., Bari F., Peták F., Novák Z. (2019). Teleauscultation: an innovative initiative to categorize and analyse lung sounds. *European Respiratory Journal*, Vol. 54, PA749. DOI: 10.1183/13993003.congress-2019.PA749.

[15] Andrès E., Gass R., Charloux A., Brandt C. and Hentzler A. (2018). Respiratory sound analysis in the era of evidence-based medicine and the world of medicine. *Journal of Medicen and Life*, Vol. 11, No. 2, pp. 89-106.

[16] Bandyopadhyaya I., Islam M. A., Bhattacharyya P., Saha G. (2017). A novel spectrogram based approach towards automatic lung sound cycle extraction. *2017 IEEE Calcutta Conference (CALCON)*, Kolkata, pp. 448-451. DOI: 10.1109/CALCON.2017.8280773.

[17] Oliynik V. (2018). Time-Domain Fragment-Similarity Processing of the Auscultation Records for Detection of their Periodic Components. *2018 IEEE 38th International Conference on Electronics and Nanotechnology (ELNANO)*, Kyiv, Ukraine, pp. 340-345. DOI: 10.1109/ELNANO.2018.8477549.

[18] Pingale T. H. and Patil H. T. (2017). Analysis of Cough Sound for Pneumonia Detection Using Wavelet Transform and Statistical Parameters. *2017 International Conference on Computing, Communication, Control and Automation (ICCUBEA)*, Pune, India, pp. 1-6. DOI: 10.1109/ICCUBEA.2017.8463900.

[19] Rahmandani M., Nugroho H. A. and Setiawan N. A. (2018). Cardiac Sound Classification Using Mel-Frequency Cepstral Coefficients (MFCC) and Artificial Neural Network (ANN). *2018 3rd International Conference on Information Technology, Information System and Electrical Engineering (ICITISEE)*, Yogyakarta, Indonesia, pp. 22-26. DOI: 10.1109/ICITISEE.2018.8721007.

[20] Sengupta N., Sahidullah Md, Saha G. (2016). Lung sound classification using cepstral-based statistical features. *Computers in Biology and Medicine*, Vol. 75, pp. 118-129. DOI: 10.1016/j.compbiomed.2016.05.013.

[21] Grinchenko V. T., Makarenkov A. P., Makarenkova A. A. (2010). Computer auscultation is a new method of objectifying the characteristics of breath sounds. *Clinical Informatics and Telemedicine*, No. 7, p. 31-37. [In Russian].

[22] Grinchenko B. T., Vinogtadnyi G. P., Makarenkova A. A. (2006). Acoustic Sensor. [AKUSTYChNYI SENSOR]. *Patent Of Ukraine [Patent Ukrainy]*, No. 14732, Bul. № 5, 15.05.2006.

[23] Gross V., Dittmar A., Penzel T., Schüttler F. and Von Wichert P.(1999). The Relationship between Normal Lung Sounds, Age, and Gender. *American Journal of Respiratory and Critical Care Medicine*, Vol. 162, Iss. 3, pp. 905-909. DOI: 10.1164/ajrccm.162.3.9905104.

[24] Mantji B. N., Oloo F. R. A. (2019). Electronic Stethoscope Design, Prototyping and Testing. *IEEE EUROCON 2019 -18th International Conference on Smart Technologies*, Novi Sad, Serbia, pp. 1-7. DOI: 10.1109/EUROCON.2019.8861992.

Дослідження особливостей звуків легенів для виявлення бронхіту та ХОЗЛ за допомогою методів машинного навчання

Порєва Г. С., Іванько К. О., Семків Х. І., Вайтишин В. І.

У статті показана актуальність розгляду питання дослідження методів цифрового аналізу і обробки звуків легень, пошуку нових інформативних ознак для розпізнавання патологічних звуків легень і застосування методів машинного навчання для класифікації стану бронхолегеневої системи. Зокрема, в даному дослідженні розглянуто застосування різних методів аналізу звуків легень, а саме: частотний, спектрально-часовий, вейвлет і мел-частотний кепстральний аналіз. З метою дослідження можливості застосування методів машинного навчання до проблеми класифікації дихальних сигналів у роботі використано набір даних звуків легень з 296 сигналів, які представляють 3 класи: норма, бронхіт та хронічне обструктивне захворювання легень (ХОЗЛ). Метою даного дослідження є порівняння інформативних ознак звуків легень, отриманих за допомогою різних методів обробки сигналів, а також вибір методу класифікації, що забезпечує найвищу точність ідентифікації стану бронхолегеневої системи. Для отримання частотних ознак розраховано залежність спектральної густини потужності від частоти для сигналів звуків легень з використанням методу швидкого перетворення Фур'є. Для кожного сигналу були розраховані спектральні показники та співвідношення потужностей спектру в різних діапазонах частот. Для виділення спектрально-часових особливостей звуків легень були проаналізовані спектрограми сигналів дихання. Визначено середні часові залежності спектральної густини потужності в досліджуваних діапазонах частот. В якості ознак, отриманих зі спектрограми використовувалася сума значень кривої спектральної густини потужності для набору частотних смуг. У якості параметрів для розпізнавання сигналів дихання на основі вейвлет-аналізу розраховано співвідношення енергій рівнів деталізації вейвлет-розкладу до повної енергії аналізованого сигналу. В якості ознак мел-кепстрального аналізу пропонується використовувати усереднені по часовим фреймам логарифмічні (мел) енергії банку фільтрів, а також усереднений по часовим фреймам мел-частотний кепстр. З метою отримання кращих моделей класифікації для комп'ютеризованого скринінгу захворювань легень було застосовано машинне навчання з учителем на основі дерев рішень, дискримінантного аналізу, методу опорних векторів, логістичної регресії, класифікаторів на основі методу k-найближчих сусідів та ансамблевого навчання. Визначено та порівняно точність класифікації сигналів дихання для низки класифікаторів, що використовують розглянуті набори ознак. Для побудови моделей, що забезпечують найвищу точність розпізнавання стану легень, пропонується найкраще поєднання інформативних ознак звуків легень та методів машинного навчання.

Ключові слова: звуки легень; бронхіт; хронічне обструктивне захворювання легень; спектральний аналіз; вейвлет-розклад; мел-частотний кепстральний аналіз; машинне навчання

Исследование особенностей звуков легких для выявления бронхита и

ХОБЛ с помощью методов машинного обучения

Порева А. С., Иванько Е. О., Семкив К. И., Вайтышин В. И.

В статье показана актуальность рассмотрения вопроса исследования методов цифрового анализа и обработки звуков легких, поиска новых информативных признаков патологических звуков легких и применения методов машинного обучения для классификации состояния бронхолегочной системы. В частности, в данном исследовании рассмотрено применение различных методов анализа звуков легких, а именно: частотного, частотно-временного, вейвлет и мел-частотного кепстрального анализа. С целью исследования возможности применения методов машинного обучения к проблеме классификации сигналов дыхания в работе использован набор данных звуков легких, состоящий из 296 сигналов, представляющих 3 класса: норма, бронхит и хроническая обструктивная болезнь легких (ХОБЛ). Целью данного исследования является сравнение информативных признаков звуков легких, полученных с помощью различных методов обработки сигналов, а также выбор методов классификации, обеспечивающих наиболее высокую точность идентификации состояния бронхолегочной системы. Для получения частотных признаков была рассчитана зависимость спектральной плотности мощности от частоты для сигналов звуков легких с использованием метода быстрого преобразования Фурье. Для каждого сигнала были рассчитаны спектральные показатели и отношения мощностей спектра в разных диапазонах частот. Для выделения спектрально-временных особенностей звуков легких были исследованы спектрограммы анализируемых сигналов. Определены средние временные зависимости спектральной плотности мощности в исследуемых диапазонах частот. В качестве признаков, полученных на основе спектрограммы, использовалась сумма значений спектральной плотности мощности для каждой полосы частот. В качестве параметров распознавания сигналов на основе вейвлет-анализа определены отношения энергий уровней детализации вейвлет-разложения к полной энергии исследуемого сигнала дыхания. В качестве признаков мел-кепстрального анализа предлагается использовать усредненные по временным фреймам логарифмические (мел) энергии банка фильтров, а также усредненный по временным фреймам мел-частотный кепстр. С целью построения лучших моделей классификации для компьютеризированного скрининга заболеваний лёгких было применено машинное обучение с учителем на основе деревьев решений, дискриминантного анализа, метода опорных векторов, логистической регрессии, классификаторов на основе метода k-ближайших соседей и ансамблевого обучения. Определена точность классификации сигналов дыхания для ряда классификаторов, использующих рассмотренные наборы признаков. Для построения моделей, обеспечивающих наиболее высокую точность распознавания состояния легких, предлагается лучшее сочетание информативных признаков звуков легких и методов машинного обучения.

Ключевые слова: звуки легких; бронхит; хроническая обструктивная болезнь легких; спектральный анализ; вейвлет-разложение; мел-частотный кепстральный анализ; машинное обучение