Comparison of the SVM classification results between original and DWT denoised respiratory signals considering to the transients noise

Igor Mazic, Anamaria Bjelopera and Luka Strazicic

Abstract— Asthma and COPD are the most common breathing diseases nowadays. Analysis and processing of breathing records are important diagnosis tools. This paper compares efficiency results of SVM classification of two classes: first the breathing noise and second the pause of signal samples recorded on subjects in real-life clinical conditions. In these conditions there is an appearance of noise and short impulse signals (transients) which are in this paper reduced by using DWT and different thresholding techniques to see if the denoised signal samples give better validation results. Classification results show that the best results are obtained when energy, Renyi entropy and standard deviation are taken as features for SVM classification. Testing of data revealed that original signal samples give better results of accuracy than the denoised signal samples.

Keywords**:** phonopneumogram, wavelet, classification, accuracy

I. Introduction

ANALYSIS of breathing sounds is a challenging task to scientists but also helpful in determination of type of breathing disease of a patient. In this paper we will focus on processing of acoustic signals recorded on patients with asthma and COPD (Chronic Obstructive Pulmonary Disease).

Bronchial asthma (asthma) is a reversible obstructive lung disease which results in recurrent attacks of losing breaths and wheezing. The symptoms can vary in severity and frequency from patient to patient and during the asthmatic attack, the lining of the bronchial tubes swells which results in narrowing of bronchi and reducing the flow of air into and out of lungs [1].

It is one of top five chronic diseases in the world and according to data of World Health Organization (WHO), 235 million people currently suffer from it. There were almost 383 000 deaths of asthma in 2015 [1]. Also, it is one of the most chronic diseases among children in almost all industrialized countries [2]. It can be provoked by different triggers including: viral infections, indoor and outdoor allergens, exercise, tobacco smoke and poor air quality [3]. Early diagnosis and treatment are essential for treating asthma.

Although there are diagnostic devices for determination of degree of obstruction in breathing during asthmatic attack, the patient must be cooperative and there still exists the problem of objective indicator of the respiratory system state.

Respiratory sounds are divided into two groups:normal and abnormal sounds. Normal sounds are heard from the top of the different parts of the chest wall in healthy subjects and abnormal respiratory sounds are heard from different parts of the chest wall of subjects with different respiratory diseases secretion [4].

Lung auscultation is helpful information about patient's respiratory function and the presence of wheezing is used as an important parameter to determine the predisposition to asthma [5]. Wheezes are the most known auscultation symptoms which can be shown in the beginning of asthmatic attack. Typical symptom patterns are significant in establishing the diagnosis. Other abnormal respiratory sounds are crackles which are mostly produced as a result of airway opening and airway secretion [6].

Important information about asthma is also collected on trachea and larynx. In our paper signals recorded on trachea will be processed (trachea connects the larynx to the lungs). Our problem is the identification which part of recorded acoustic signal is a pause and which is respiratory noise (inspiration/expiration). We will focus first on reducing noise and transients in recorded sound on trachea of asthma ill patients and healthy subjects by using discrete wavelet transform (DWT) with different types of wavelets and then thresholding of wavelet coefficients. Second, in order to see if the reducing of transients and noise will pollute better results of classification of recorded acoustical signals, we will use SVM (Support Vector Machine) tool for classification of two classes in recorded signal:

- respiratory noise (inspiration/expiration),
- pause of the signal.

In second section the problem will be identified and solution will be proposed. Third section describes the modeling and classification of breathing sounds and results of validation. Fourth section will bring the conclusion.

Igor Mazic is with the Department of Electrical Engineering and Computing, University of Dubrovnik, Cira Carica 4, 20000 Dubrovnik, Croatia (e-mail: igor.mazic@unidu.hr).

Anamaria Bjelopera is with the Department of Electrical Engineering and Computing, University of Dubrovnik, Cira Carica 4, 20000 Dubrovnik, Croatia (e-mail: anamaria.bjelopera@unidu.hr).

Luka Strazicic is a student at the Department of Electrical Engineering and Computing, University of Dubrovnik, Cira Carica 4, 20000 Dubrovnik, Croatia .(e-mail: luka.strazicic@unidu.hr).

II. Problem Identification and Proposed Solution

In medicine in order to identify the state of respiratory system, acoustic breathing records (phonopneumograms) based on anatomical and physiological parameters i.e. age, sex, type and stage of disease are used [7]. Measuring equipment is consisted of transducers that are put on chest or trachea and which collect acoustic signals during breathing [2]. Analysis of recorded acoustic signals is not an easy work and there are many factors that affect the results of auscultation signal analysis [8]: air volume changes in the lungs, corpulence of the patient and age, location of sound capturing, breathing flow, position of the patient and also the measurement equipment features.

A normal respiratory sound is the sound produced by the lungs of healthy people during inspiration and expiration. More than 75% of acoustic power during the inspiration is between 100 and 250 Hz, while 99% is less than 600 Hz [9]. Abnormal respiratory sounds that do not occur during breathing of healthy person and that are found in asthma are wheezes and crackels. Wheezes are mostly found in the expiration, but they can also be found during inspiration. Crackles are usually found during inspiration and rarely during expiration [10]. In normal breathing inspiration/expiration time ratio is from 1:1.5 to 1:2. During the asthmatic attack there is appearance of obstructions in breathing so the time of expiration is prolonged and it includes wheezing.

For analysis of breathing records it is important to distinguish pause and breathing noise of recorded signal so the main goal of our paper is classification of two classes in breathing records: respiratory noise (during inspiration/expiration) or the pause in recorded acoustic signals of different patients. First preprocessing of recorded breathing signals should be made, second feature extraction and modelling and third validation of classification.

In acquisition of respiratory sounds mostly two types of measuring sensors are used: accelerometer and microphone. Accelerometer is put on skin and microphone is built in the closed box. Also, accelerometer measures vibrations on skin, while the microphone measures the change of sound pressures made by vibration of skin during breathing. Measuring sensors are usually allocated on chest wall or on trachea. Spectrum of signal measured on trachea is different from that measured on chest wall because of the low-pass filter behaviour of skin and chest wall. The tissue of trachea is thinner and so the attenuation of higher spectral frequencies will be smaller than those higher frequencies measured on chest wall. Normal sound of breathing measured on trachea is in frequency band between 100 and 1500 Hz. Acoustical power measured on trachea is higher during expiration while on chest wall during the inspiration. We will analyse the recorded breathing on trachea.

Our measuring system consists of Thinklabs One Digital Stethoscope [11] connected to the notebook. In measuring of breathing sounds on trachea, except the noise present in the acoustic signal as a result of a noisy hospital, there is also an emergence of non-physiological artefacts in signal which are the result of restless subjects. Moving of the stethoscope during breathing caused appearance of very short repeating impulses of high intensity (transients) compared to normal breathing noise. This can reduce the results of validation of classification of acoustic signals so first preprocessing of signal is going to be made by using DWT and thresholding technique [12], [13], [14] on wavelet coefficients in order to reduce the transients and the noise in recorded breathing signal, so actually we will be denoising the signal. We use Wavelet Toolbox in MATLAB software for three steps:

- DWT- decomposition of acoustic signal,
- hard or soft thresholding of wavelet coefficients,
- IDWT - reconstruction of signal.

Wavelets are powerful tool for audio signal processing and it is possible to find the best wavelet family for each problem solution mostly by testing each one and finding the most suitable. Here in our testing we use five wavelet families: Haar, Daubechies (db6), Coiflet (coif5), Biorthogonal (bior3.9) and Symlet (sym5) family and for each one 6 levels of decomposition and then compare the results. We also use different thresholding techniques: hard thresholding which sets all coefficients values to zero if they are bellow given threshold and keeps the values if they are above threshold or soft thresholding which reduces the wavelet coefficients above threshold by the value of threshold and changes the signal energy. Also for soft and hard thresholding we found that 'sqtwolog' (universal threshold) and 'heursure' (mix of universal threshold and threshold defined by Stein's Unbiased Estimate of Risk) methods of finding threshold are most suitable for reducing transients and noise in signals.

Table I Standard deviation of used signals before and after wavelet denoising (Haar wavelet, 6 levels of decomposition, hard thresholding)

Signal	Original signal	Denoised signal
9T	0.0861754	0.0113438
12T	0.0222024	0.0039243
13T	0.1534788	0.0181377
14T	0.1613673	0.0183153
15T	0.1084039	0.0132604
16T	0.0412584	0.0084188
17T	0.0424603	0.0060601
18T	0.0097667	0.0019480
19T	0.0281140	0.0042319
22T	0.0092997	0.0025144
23T	0.0251523	0.0096047
27T	0.0059676	0.0031546
37T	0.0193358	0.0021933
38T	0.0171253	0.0030508
39T	0.0131973	0.0011727
44T	0.0103946	0.0051564

Calculated standard deviation of 16 different breathing signals recorded on trachea (T) used for training and validation before and after wavelet denoising can be seen in Table I in order to show how the noise and transients are reduced.

Results were calculated using MATLAB Wavelet Toolbox and best results were shown using Haar wavelet and hard denoising with 'sqtwolog' method. Also, 6 levels of decomposition were used. On Fig.1 original (red) and denoised 23T signal with Haar wavelet and hard 'sqtwolog' thresholding (purple) are shown.

After preprocessing with wavelet denoising we focus on feature extraction and classification of signals and then compare the results.The purpose of features extraction is to convert the signal waveform into a reduced number of parameters that will be used for further analysis and processing [15]. Various types of extracted features can be used: time-frequency spectrum, entropy, Mel Frequency Cepstral Coefficients (MFCC), power spectral density (PSD), standard deviation (SD), Peak Frequency (FP), skewness, kurtosis, etc. Which features are the best to be extracted, is a consequence of different signal processing and classification techniques applied on recorded acoustic signals [2].

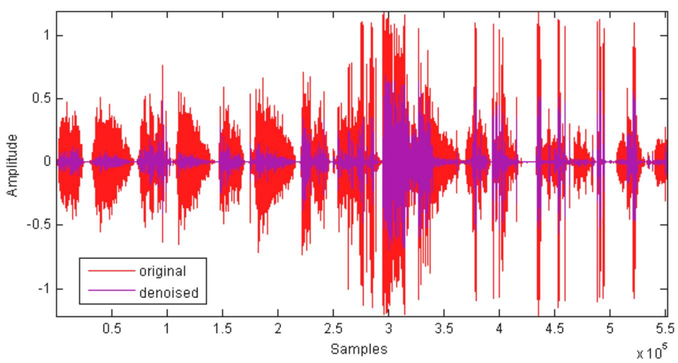

Fig. 1 Original and denoised 23T signal

Use of MFCC features has been very useful in acoustical signal analysis and classification [16]. There are many papers that describe helpful use of MFCC features in analysis and better accuracy of recognition of different sounds of acoustic signals [7], [15], [17], [18], [19], [20] and therefore they are used here for classification of respiratory sounds. Also, energy (ENE), Renyi entropy (E) (α=2) and standard deviation (SD) are taken into account.

For classification we take SVM [7], [21] which is nowadays considered to be a powerful tool for classification in processing of acoustic signals. It is a kernel-based learning algorithm designed for binary classification which is suitable for our case of classification of our two classes (respiratory sound and pause). Support vectors define a decision boundary in higher dimensional feature space that separates two classes and its parameter C is set to one. To improve classification results Radial basis kernel function (RBF) is used and its parameter γ varies.

In order to see if it is necessary to clean the signal of noise and transients to get better results of classification we do the validation of results in next chapter.

III. Modeling and Classification of Respiratory Sounds

A. Evaluation Measures

The efficiency of recognition of two classes: breathing noise and pause, was evaluated with following validation measures: overall accuracy, $\overline{\mathrm{TPR}}$ (True Positive Value or Sensitivity), $\overline{\mathrm{TNR}}$ (True Negative Value or Specificity) and overall reliability ($\overline{\mathrm{R}}=\overline{\mathrm{TPR}}\cdot\overline{\mathrm{TNR}}$). Overall accuracy is calculated as in following equation:

$$\overline{\mathrm{ACC}}=\frac{1}{\mathrm{N}}\sum_{i=1}^{\mathrm{N}}\frac{\mathrm{TP}_i+\mathrm{TN}_i}{\mathrm{TP}_i+\mathrm{TN}_i+\mathrm{FP}_i+\mathrm{FN}_i}. \quad (1)$$

$\overline{\mathrm{TPR}}$ was calculated as:

$$\overline{\mathrm{TPR}}=\frac{1}{\mathrm{N}}\sum_{i=1}^{\mathrm{N}}\frac{\mathrm{TP}_i}{\mathrm{TP}_i+\mathrm{FN}_i}, \quad (2)$$

$\overline{\mathrm{TNR}}$ as:

$$\overline{\mathrm{TNR}}=\frac{1}{\mathrm{N}}\sum_{i=1}^{\mathrm{N}}\frac{\mathrm{TN}_i}{\mathrm{TN}_i+\mathrm{FP}_i}, \quad (3)$$

and overall reliability as:

$$\overline{\mathrm{R}}=\overline{\mathrm{TPR}}\cdot\overline{\mathrm{TNR}}. \quad (4)$$

N is a number of experiments performed using the random subsampling validation method, and *i* is the index of iteration. The input data is divided into two subsets: breathing noise and pause.TP, FP, TN and FN are the numbers of the signal segments (samples) classified as true positive, false positive, true negative and false negative [7].

B. Ethics Statement

For our measured breathing records informed consent was obtained and measurements were carried out in Clinical Center for Pulmonary Diseases Jordanovac. The procedure was approved by the ethics committee of the University Hospital Centre Zagreb.

C. Results of Validation

From 10 phonopneumograms (of healthy and ill subjects) recorded on trachea taken for training and validation, 10 sequences of breathing noise and 10 sequences of pause were extracted (one sequence of breathing noise and one of pause for each phonopneumogram). Time window without overlapping with duration of 100 ms was moving across every sequence and four features were calculated: MFCC, energy, Renyi entropy and standard deviation. As a result, total of 149 samples (14.9 s) of respiratory noise and 119 samples (11.9 s) of pause were computed. 20% of samples were used for training and 80% for validation of model using the random subsampling method . The process was repeated N=20 times (each time with randomly selected samples) for $0.1<\gamma<3$, step=0.1 (C=1) and best averaged results with suitable gamma value are shown in following tables.

First Table II presents validation results for original signal samples (without denoising).

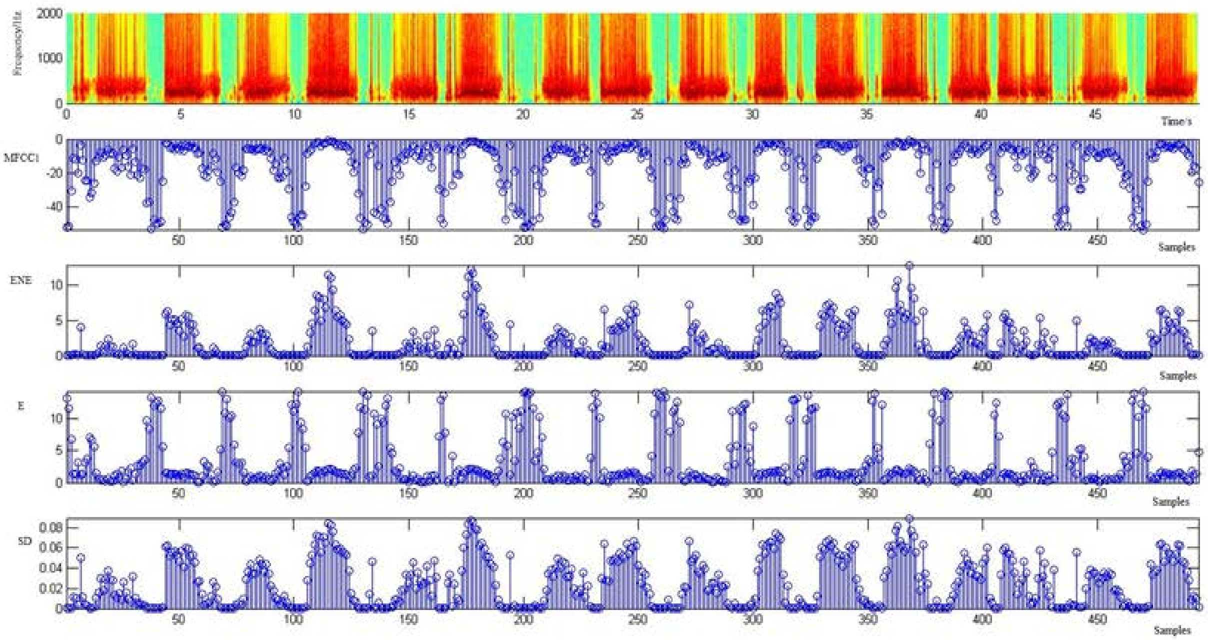

Fig. 2 ENE, E, SD and $MFCC_1$ features of 15T original signal samples

Table II Validation results expressed with evaluation measures achieved by using the appropriate features for original signal samples (without denoising)

Feature	$\overline{TPR}$ (%)	$\overline{TNR}$ (%)	$\overline{R}$	$\overline{ACC}$ (%)	γ
12 MFCC coeff.	51.3889	54.3689	0.2794	52.7293	0.6
1stMFCC coeff.	96.2598	86.8383	0.8359	91.9991	1.4
ENE, E, SD	98.1250	97.6699	0.9584	**97.9221**	0.9
1stMFCC coeff., E, ENE, SD	97.3228	97.1359	0.9454	97.2391	1.1

Table III Validation results expressed with evaluation measures achieved by using ENE, E and SD features for denoised signal samples ('sqtwolog' hard thresholding)

Wavelet	$\overline{TPR}$ (%)	$\overline{TNR}$ (%)	$\overline{R}$	$\overline{ACC}$ (%)	γ
bior 3.9	86.1328	82.0952	0.7071	84.31	2.6
haar	95.1575	91.9409	0.8749	**93.7028**	1.4
db6	72.1311	82.2330	0.5932	76.6937	0.5
sym5	78.8628	70.3944	0.5551	75.0339	0.2
coif5	88.3333	73.9320	0.6531	81.9397	0.6

Although in literature MFCC coefficients show the best results in classification of respiratory noises, the results of our model validation in Table II for original signal samples prove that for discrimination of respiratory noises and pause, energy, Renyi entropy and standard deviation (ENE, E and SD) give better results ($\overline{ACC}$ =97.92%) than MFCC (52.73%) and $MFCC_1$ (1^{st}MFCC coefficient) (92%). Moreover, by adding the $MFCC_1$ feature, accuracy of classification becomes reduced ($\overline{ACC}$ =97.23%).

Next we processed validation of samples of denoised original signal. Three types of wavelet denoising were taken for five types of wavelets: 'heursure' soft, 'sqtwolog' soft and 'sqtwolog' hard thresholding methods. The results were compared using evaluation measures. All calculation was made in MATLAB software. Results are shown in Tables III, V and VI.

In Table III haar wavelet shows the best result. Also, Table IV shows that taking $MFCC_1$ feature into account the accuracy of results reduces so it is better to just take ENE, E and SD.

Table IV Comparison of validation results for hard denoised signal samples using haar wavelet

Feature	$\overline{TPR}$ (%)	$\overline{TNR}$ (%)	$\overline{R}$	$\overline{ACC}$ (%)	γ
ENE, E, SD	95.1575	91.9409	0.8749	**93.7028**	1.4
1stMFCC coeff., E, ENE, SD	95.8730	90.5714	0.8683	93.4632	1.7

The best result of validation from denoised signal samples, occurs for: haar wavelet and 'sqtwolog' hard denoising, and features: ENE, SD and E resulting in $\overline{ACC}$ =93.70%, which is worse than result for original signal samples (without denoising) $\overline{ACC}$ =97.92%.

Table V Validation results expressed with evaluation measures achieved by using ENE, E and SD features for denoised signal samples (‘sqtwolog’ soft thresholding)

Wavelet	$\overline{TPR}$ (%)	$\overline{TNR}$ (%)	$\overline{R}$	$\overline{ACC}$ (%)	γ
bior 3.9	81.28039	77.7261	0.631761	83.72247	0.2
haar	82.46032	86.06061	0.709658	84.04444	0.3
db6	86.25	76.97115	0.663876	82.09052	1.3
sym5	85.1153	83.59223	0.711498	**84.44**	1.5
coif5	87.57813	74.18269	0.649678	81.57328	2

Table VI Validation results expressed with evaluation measures achieved by using ENE, E and SD features for denoised signal samples (‘heursure’ soft thresholding)

Wavelet	$\overline{TPR}$ (%)	$\overline{TNR}$ (%)	$\overline{R}$	$\overline{ACC}$ (%)	γ
bior 3.9	82.72	84.95098	0.702715	**83.72247**	0.3
haar	85.44	72.40196	0.618602	79.5815	0.3
db6	70.23077	77.3301	0.543095	73.3691	0.8
sym5	69.4186	79.95146	0.555012	74.09483	1.8
coif5	76.21111	84.19498	0.641659	79.7593	0.4

D. Results of Testing

6 phonopneumograms were used for testing of model. None of the signal samples from testing set was used for training/validation. Set for training is consisted of signal samples from healthy subjects: 38T, 39T, 9T, 19T, 17T and from subjects with obstructions: 44T, 18T, 12T, 13T, 14T. Set for testing is consisted of signal samples for healthy subjects: 15T, 16T, 22T and for subjects with obstructions: 27T, 37T, 23T.

Table VII Comparison of testing results for ENE, E and SD features, γ =0.9, C=1

Signal	Samples (time)	FP	FN
15T	495 (49.5 s)	19	0
16T	470 (47 s)	11	0
22T	512 (51.2 s)	22	0
23T	344 (34.4 s)	1	3
27T	516 (51.6 s)	32	18
37T	530 (53 s)	27	18

Table VII shows result of testing with 2867 samples, with ENE, E and SD features, γ =0.9 and C=1, which gave total error number of 151 and $\overline{ACC}$ =94,73% (it was 97,92% in validation).

For testing with denoised signal samples (haar, hard ‘sqtwolog’ denoising) with ENE, E and SD features, γ =1.4 and C=1, errors cannot be counted and also spectrogram is very different from the original one which is shown on Fig. 6 for 27T denoised signal. Results of testing are shown of Fig. 3, 4, 5 and 6 which show spectrograms of signals and SVM classification results where the value one is assigned to respiratory noise and the value of zero to pause.

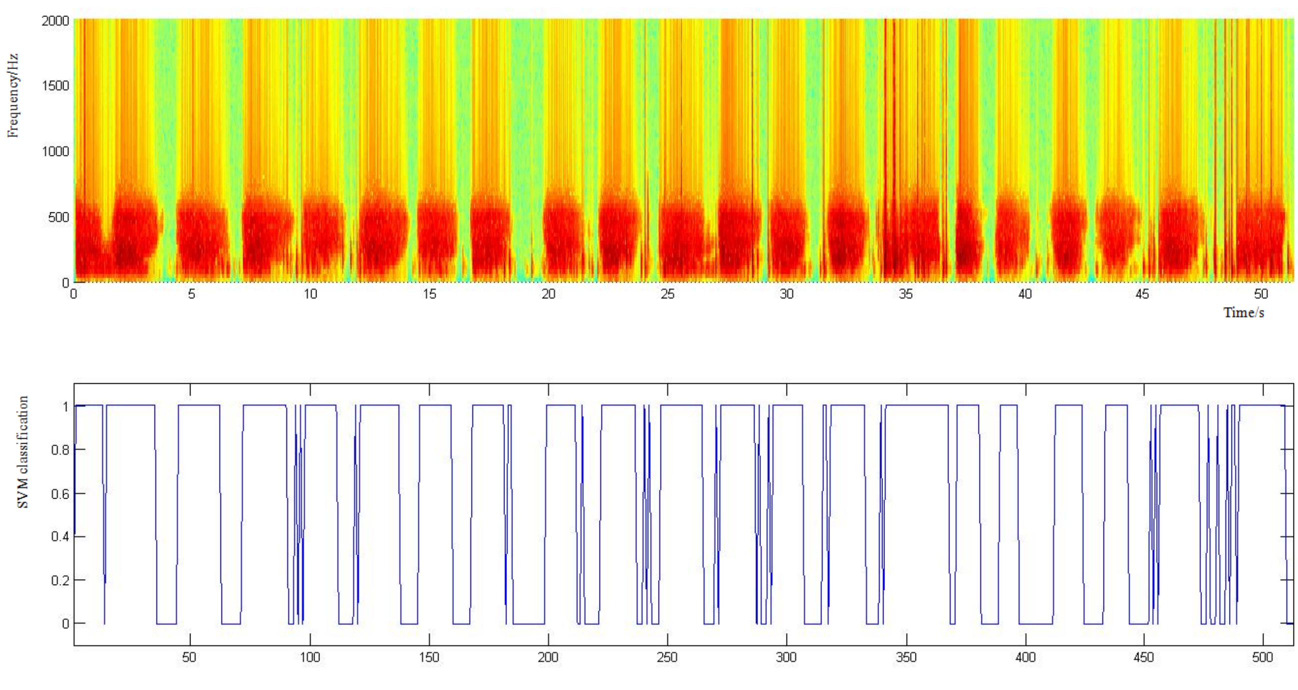

Fig. 3 Results of classification of respiratory sound and pause for 22T original signal samples: 2D spectrogram (first graph) and SVM classification result (second graph)

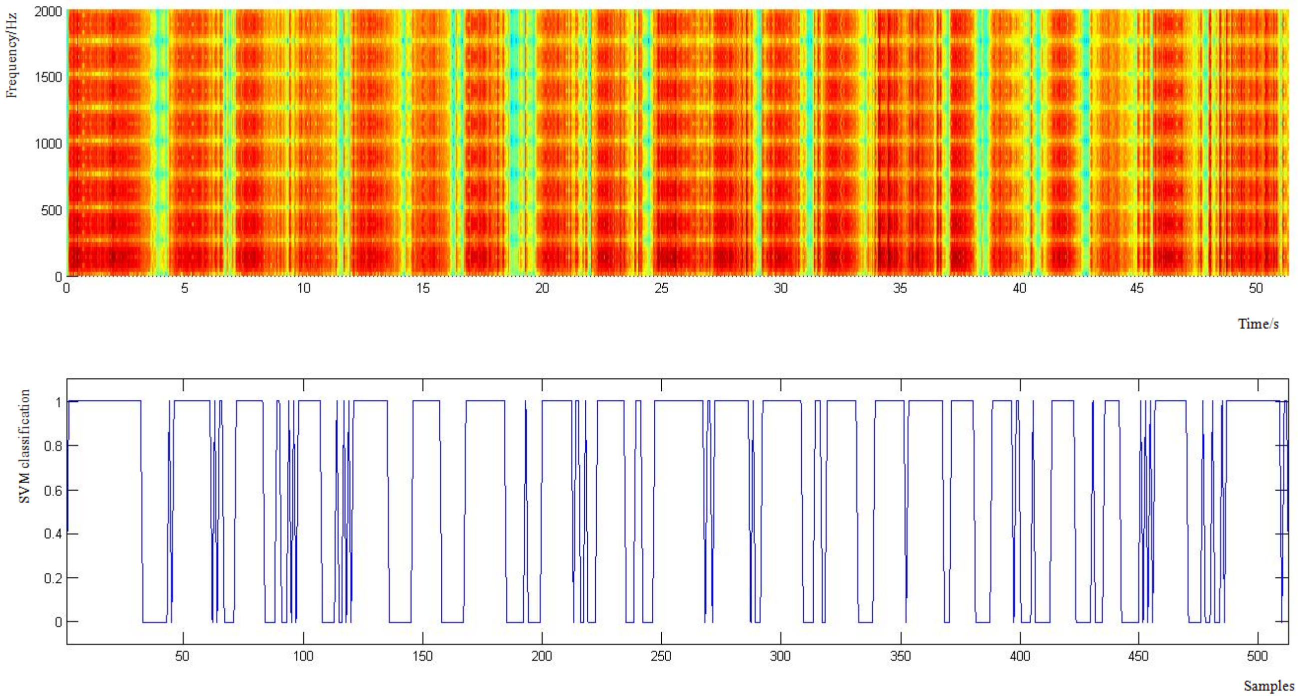

Fig. 4 Results of classification of respiratory sound and pause for 22T denoised signal samples:
2D spectrogram (first graph) and SVM classification result (second graph)

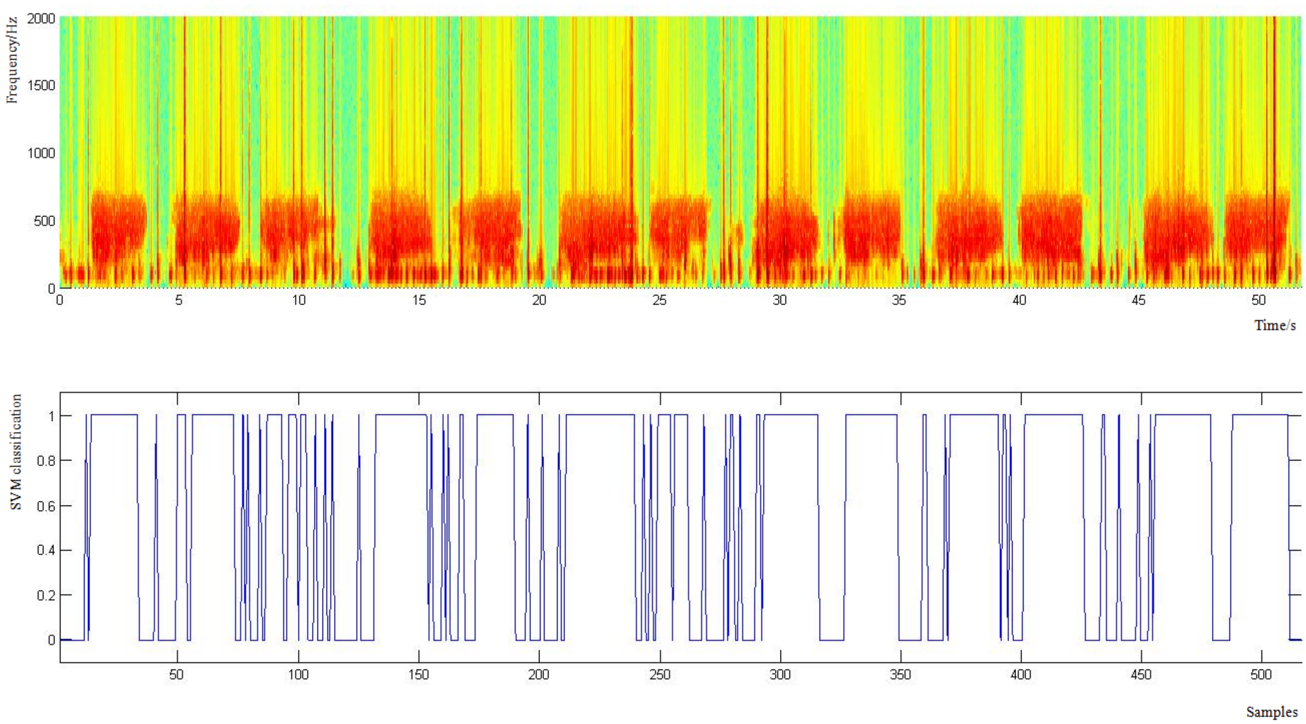

Fig. 5 Results of classification of respiratory sound and pause for 27T original signal samples:
2D spectrogram (first graph) and SVM classification result (second graph)

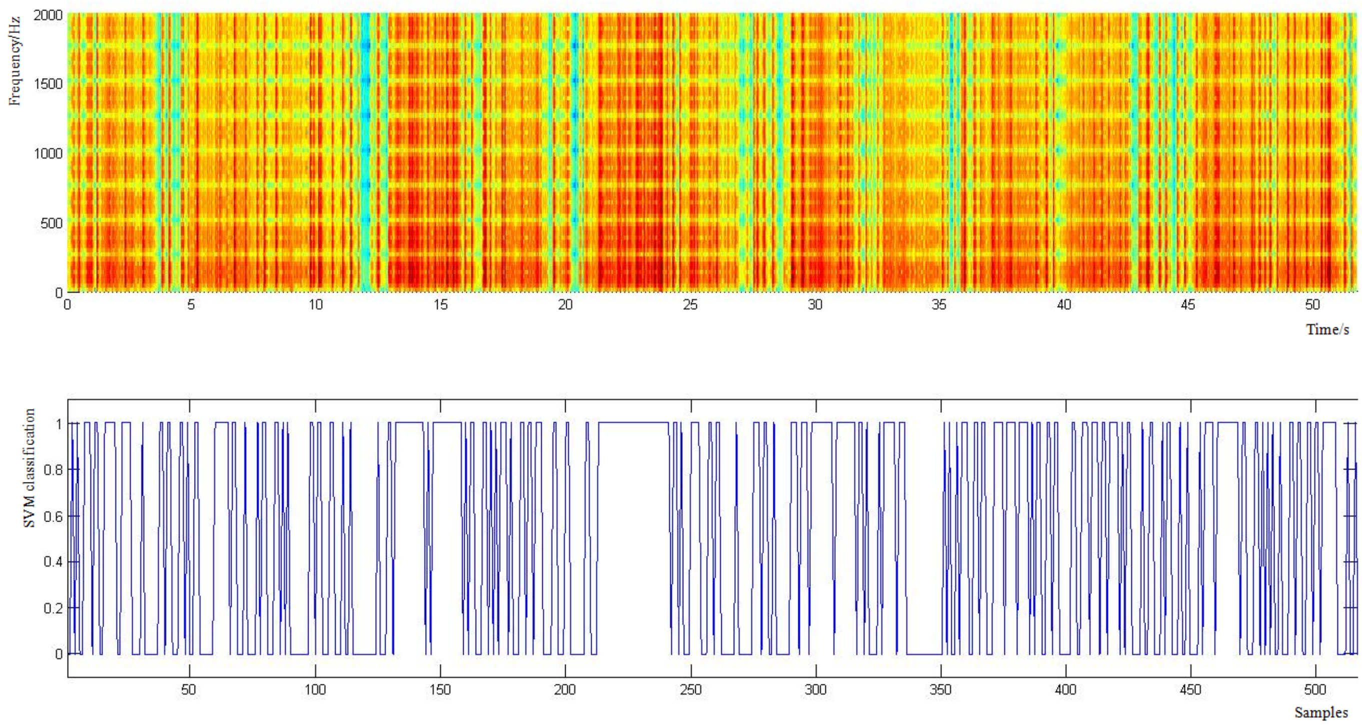

Fig. 6 Results of classification of respiratory sound and pause for 27T denoised signal samples: 2D spectrogram (first graph) and SVM classification result (second graph)

IV. Conclusion

Results for classification of respiratory noise and pause for recorded breathing signals were calculated and analysed. The signals were recorded using standard measuring equipment in realistic conditions so it caused appearance of noise and transients in signals. DWT analysis and different methods of thresholding were used to denoise the signal and reduce the transients. According to calculation haar wavelet and ‘sqtwolog’ hard denoising gave the best denoising results. We used this denoised signals and made the validation with different features upon we concluded that the most suitable features for our analysis were: standard deviation, Renyi entropy and energy. Although in literature MFCC features give the best results of validation for acoustic signals, here in our paper this is not the case. Also, we got better accuracy results for original signal samples than for denoised signal samples. In our case the denoising of signal, does not give better results of classification because not only that the transients and the noise reduce, but also the whole signal energy changes obviously enough to not give better results of accuracy.

References

[1] World Health Organization-Asthma. Available: http://www.who.int, (1.06. 2018)

[2] Milicevic M., Mazic I., Bonkovic M., “Classification Accuracy Comparison of Asthmatic Wheezing Sounds Recorded under Ideal and Real-world Conditions”, In: *15th International Conference on Artificial Intelligence, Knowledge Engineering and Databases* (AIKED 2016), Venice , 2016

[3] Milicevic M., Mazic I., Bonkovic M., “Asthmatic Wheezes Detection- What contributes the Most to the role of MFCC in classifiers accuracy”, *International Journal of Biology and Biomedical Engineering* , Vol. 10**,** 2016, pp. 176–182

[4] Molaie M., et al., “A chaotic viewpoint on noise reduction from respiratory sounds”, *Biomed. Signal Process. Control* , Vol. 10 , 2014, pp. 245–249

[5] Martinez F.D., Wright A.L., Taussig L.M., Holberg C.J., Halonen M., Morgan W.J., “Asthma and wheezing in the first six years of life”. *N Engl J Med* 1995, Vol. 332, pp. 133–138

[6] Hadjileontiadis L.J., “Lung Sounds: An Advanced Signal Processing Perspective”, 1st ed., Morgan & Claypool Publishers, 2008

[7] Mazic I., Bonkovic M., Dzaja B., “Two-Level Coarse-to-Fine Classification Algorithm for Asthma Wheezing Recognition in Children's Respiratory Sounds”, *Biomedical Signal Processing and Control*,Vol. 21 , 2015, pp. 105–118

[8] Earis J.E., Cheetham, B.M., “Current methods for computerized respiratory sound analysis”, *Eur. Respiratory Rev*., 2000 , Vol. 10, pp. 586–590

[9] Pasterkamp H., Kraman S. S., Defrain P., Wodicka G., “Measurment of respiratory Acoustical Signals (Comparison of Sensors) ”, *Chest*, Vol. 104, 1993, pp. 1518–1525

[10] Forgacs P., “Crackels and wheezes”, *Lancet*, Vol. 2, 1967, pp. 203–205

[11] Measurement equipment. Available: https://www.thinklabs.com

[12] Donoho D.L., Johnstone I. M., “Ideal spatial adaptation by wavelet shrinkage”, *Biometrika*, Vol. 81, 1994, pp. 425–455

[13] Donoho, D.L., “De-noising by soft- thresholding, ”, *IEEE Trans. on Inf. Theory*, Vol.41, No. 3, 1995, pp. 613–627

[14] Luisier F., Blu T., Unser M., “A New SURE Approach to Image Denoising: Interscale Orthonormal Wavelet Thresholding, ” *IEEE Transactions on Image Processing*, Vol. 16, No. 3, 2007, pp. 593–606

[15] Bahoura M., Pelletier C., “ New parameters for respiratory sound classification”, In: C*anadian Conference on Electrical and Computer Engineering, 2003, IEEE CCECE* 2003, Montreal

[16] Davis, S., Mermelstein, P., “Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences”, *IEEE Trans. Acoustics, Speech, Signal Process*. Vol. 28, No. 4, 1980, pp. 357–366

[17] Sengupta N., Sahidullah M., Saha G., “ Lung sound classifcation using cepstral-based statistical features”, *Computers in Biology and Medicine*, Vol. 75, 2016, pp. 118–129

[18] Sahidullah M., Saha G., “Design, analysis and experimental evaluation of block based transformation in MFCC computation for speaker recognition”, *Speech Communication* , Vol. 54, No.4, 2012, pp. 543–565

[19] Kinnunen, T., Li, H., “An overview of text-independent speaker recognition: From features to supervectors”, *Speech Comm.* , Vol.52 , No.1, 2010, pp. 12–40

[20] Togneri R., Pullella D., “An Overview of Speaker Identification: Accuracy and Robustness Issues”, I*EEE Circuits And Systems Magazine*, Vol. 11, No. 2 , pp. 23–61

[21] Cristianini N., Shawe-Taylor J., “An Introduction to Support Vector Machines and Other Kernel-based Learning Methods”, Cambridge University Press, New York, NY, 2000