Classification Accuracy Comparison of Asthmatic Wheezing Sounds Recorded under Ideal and Real-world Conditions

MARIO MILICEVIC\textsuperscript{1}, IGOR MAZIC\textsuperscript{1}, MIRJANA BONKOVIC\textsuperscript{2}
\textsuperscript{1}Department of Electrical Engineering and Computing
University of Dubrovnik
Cira Carica 4, 20000 Dubrovnik
CROATIA
\textsuperscript{2}Faculty of Electrical Engineering, Mechanical Engineering and Naval Architecture
University of Split
Rudjera Boskovica 32, 21000 Split
CROATIA
mario.milicevic@unidu.hr, igor.mazic@unidu.hr, mirjana.bonkovic@fesb.hr

Abstract: - Asthma is the most common chronic disease among children. Diagnosis of asthma is often challenging, so the computerized lung sound analysis is important diagnostic aid. This research compares the efficiency of the classification algorithms applied both on signals available on the internet and signals recorded on children in real-life clinical settings. With an appropriate signal processing technique, resulting in MFCC features, it is possible to achieve high classification accuracy for signals recorded in suboptimal conditions.

Key-Words: - machine learning, classification, asthma, phonopneumogram, MFCC, SVM, k-NN

1 Introduction
The World Health Organization (WHO) estimates a number of 235 million people currently suffering from asthma [1], which is a public health problem not just for high-income countries; it occurs in all countries regardless of the level of development.

Asthma is considered to be the most common chronic disease among children in nearly all industrialized countries [2]. It is more prevalent in children with a family history of atopy, and symptoms and worsening thereof are frequently provoked by a wide range of triggers, which can include viral infections, indoor and outdoor allergens, exercise, tobacco smoke and poor air quality. A large number of infants and preschool children experience recurrent episodes of bronchial symptoms, especially wheezing and coughing, beginning at a few months of age. Since a clinical diagnosis of asthma can usually be made with certainty by the age of 5, early diagnosis, monitoring and treatment of respiratory symptoms are essential.

Lung auscultation is helpful in providing information concerning the patient’s respiratory function. The presence of wheezing in infants is used as an important parameter in assessing the predisposition to asthma [3]. History of repeated episodes of wheezing is a symptom universally accepted as the starting point for asthma diagnosis in children [4]. The required number of such episodes is generally unspecified, although an arbitrary number of three or more has been proposed. Typical symptom patterns are significant in establishing the diagnosis.

A wheeze can be described as an unintentional and continuous sound [5]. Acoustically, it is characterized by periodic waveforms with a dominant frequency usually over 100 Hz (or 400Hz [6]) and with duration of ≥100 ms. Wheezes are usually associated with airways that are obstructed due to various causes. Wheezes with a single peak or with the harmonics of a single basal peak are called monophonic wheezes, while those with variable peaks that differ in harmonics are called polyphonic wheezes [7].

Recently developed computer-based respiratory sound analysis methods serve as a powerful tool to diagnose the whole spectrum of disorders and abnormalities in the lungs, including asthma.

2 Problem Formulation
Phonopneumograms (acoustic breathing records) depend on anatomical and physiological parameters such as sex, age, type and stage of
disease. Digital methods of collecting, processing and analyzing phonopneumograms have been in wide use for more than 30 years. Measuring systems consist, among others, of transducers that are put on the chest or trachea which collect acoustic signals during breathing.

Several factors affect the results of auscultation signal analysis and make comparing the research centers more difficult [8]: age and corpulence of the patient, air volume changes in the lungs, location of sound capturing, breathing flow, position of the patient and characteristics of the measurement equipment.

Differences due to age are all the more visible with infants. Audible respiratory sounds in early childhood have acoustic characteristics recognizably different from those generally heard in adults. Therefore, Mazic et al. [9] propose to use more objective methods to automatically detect wheezing in asthmatics infants, during forced breathing.

Many methods were used by previous researchers during the past three decades to process the lung sounds for detecting wheezing [10]. Various types of extracted features have been used, such as the time-frequency spectrum, entropy, Mel Frequency Cepstral Coefficients (MFCC), power spectral density (PSD), standard deviation (SD), Peak Frequency (FP), skewness, kurtosis, etc. There is no consensus on which features are the best to be extracted, because the final system performance is a consequence of different signal processing and classification techniques applied.

Most authors used phonopneumograms to automatically detect wheezing from different internet databases (INT) or media [11, 12, 13], the purpose of which is first and foremost education, so the data on the measuring system, transducers, position of the measuring point, the age of the subject and breathing technique were mostly not published. A 2D spectrogram of such a signal [11] is shown on Fig. 1.

It is clear from the spectrogram that inspiration lasts almost as long as expiration in which wheezing is present. The signal to noise ratio (S/N), breathing regularity, the expiration/inspiration time ratio, as well as the presence of wheezing in all expirations point to the conclusion that this is a state of controlled forced breathing of an adult in laboratory measuring conditions. It is not possible to achieve these conditions with children under the age of 6.

In this paper, we compare the predictive abilities of models built from publically available phonopneumograms, to those built from phonopneumograms recorded in the Dubrovnik General Hospital (DGH) in realistic, suboptimal conditions with children aged from one to six.

Phonopneumograms recorded with children contain not only muscular and cardiovascular sounds, but also many physiological and non-physiological artifacts, such as sounds which are results of forced breathing, or stridor and wheezing which do not originate in the bronchia, but are the consequence of infections in the upper part of the respiratory system, which is a common occurrence with children of that age. All these signals are superimposed in the transducer, which only adds to the difficulty of their recognition and classification.

Additionally, to ensure sufficient acoustic power for the respiratory sound, the children were usually encouraged to perform forced breathing, which often resulted in specific physiological artifacts such

Fig. 1 2D spectrogram of the signal from INT dataset
as inspiratory stridor, which sounds similar to asthmatic wheezes.

It can be expected that recording lung sounds in a noisy hospital versus laboratories under controlled conditions can demand more rigorous preprocessing techniques to combat the noise present in the acoustic signal. Also, there is the possibility that the classification algorithms are less effective.

3 Modelling and classification
The measuring equipment consists of a transducer, 4m long microphone cable, preamplifier, stable 5V source, and a personal computer with an integrated audio card [14]. The resolution of the AD/DA converter is 16 bits, signal-to-noise ratio (SNR) is 90 dB, and the total harmonic distortion at 1 VRMS was 0.01%. Sampling rate is 8 KHz.

Fig. 2 shows a 3D spectrogram of a three-year-old’s breathing. During forced breathing an inspiratory stridor was also recorded. To reduce the impact of cardiovascular and muscular noise, phonopneumograms were first filtered with the Yule–Walker 50th-order high pass filter, with the lower cut-off frequency of 100 Hz. Then, STFT is performed with 50% overlapping Hamming window using 256 samples which corresponds to 32 ms.

Fig. 3 shows a 2D spectrogram of the same phonopneumogram, where the acoustic power of the inspiratory stridor and wheezing within the same order of magnitude can be seen, while the 3D spectrogram shows that the inspiratory stridor appears at lower frequencies than the wheezing, which does not always appear to be the case. Researchers show that by using MFCC as features
[15, 16, 17], wheezing detection can achieve an accuracy higher than 95%. There is no standard number of MFCCs for recognizing the lung sounds. For both dataset sources (INT, DGH) we experiment to obtain the optimal number of MFCCs resulting in the maximum classification reliability. For the INT signals we also investigate previous works that are directly related to this topic. Finally, we used 15 MFCCs for the INT signals, and 12 MFCCs for our DGH signals.

For comparison purposes, we also used standard statistical features computed from spectral components calculated using FFT: Renyi entropy, Kurtosis, Spectral Flatness (SF), Skewness, Mean Crossing Irregularity (MCI), Standard Deviation (SD) and f50/f90 ratio [18].

The models are built based on several known machine learning methods: Support Vector Machine (SVM) [19], k-nearest neighbor algorithm (k-NN) [20], Neural Network (NN) [21], Random Forests (RF) [22], Logistic Regression (LR) [23], Naive Bayes (NB) [24].

For the aforementioned machine learning methods and both datasets, an optimization of important parameters which influence the model build was made. For some algorithms, this is a very important step, such as the SVM with RBF kernel, for which it is important to set the C and gamma parameters. This is why a grid search using cross-validation was used for finding optimal values in a parameter space.

**Fig. 4** Results of the grid search procedure

Fig. 4. shows the results of the grid search procedure used to determine the optimal gamma and C parameters for the SVM algorithm (RBF kernel). One can see that it is very important to pay attention to the choice of values for the mentioned parameters for the SVM algorithm. Despite this, many papers comparing the SVM algorithm to other algorithms can be found which have no clear methodology of the parameter optimization.

**Fig. 5** Classification accuracy for INT/DGH signals represented with 7 statistical features

Fig. 5. shows the results of the classification for both datasets based on 7 statistical features. The best results (accuracy 93.62% for INT signals, or 91.77% for DGH signals) were achieved with the Neural network algorithm with 2 hidden layers. The SVM and k-NN algorithms, pointed out in numerous articles as a good choice for pulmonary acoustic signals classification, scored somewhat lesser results here (accuracy between 80% and 85%). Looking at all 6 algorithms, one can conclude that the classification is more successful with internet data.

On the other hand, **Fig. 6**. shows the classification results for both datasets based on MFCC features. The results are, as expected, much better, and the most successful classifiers are SVM and k-NN, which had the accuracy of 99%. At the same time, all classifiers have achieved somewhat better results for DGH signals.

**Fig. 6** Classification accuracy for INT/DGH signals represented with MFCC features
Additionally, Fig. 7. shows the performance of binary classifiers for DGH signals (MFCC features) in the form of Receiver Operating Characteristic (ROC) curves. By analyzing the shown curves, as well as the Area Under the Curve (AUC), it can be seen that the best results are achieved with the SVM and k-NN algorithms.

4 Conclusion
The results of wheeze detection based on signals available on the internet were compared to the signals recorded in the Dubrovnik General Hospital. The signals were recorded using equipment comprised of standard components, with children in realistic conditions, including the effects of ambient sounds, cardiovascular and muscular noise etc. Different pattern recognition methods were used to classify both datasets of respiratory sounds into normal and wheeze classes. The experiments show that, by properly filtering and preprocessing the entry data, and using MFCC features, signals recorded in suboptimal conditions can achieve good results, very similar to those collected from the internet. That is an important prerequisite for the construction of a low cost automated system for monitoring asthma, based on a mobile device and the appropriate application, in which raw data from a transducer is processed and analyzed.

References:


