Artificial Neural Networks for Acoustic Lung Signals Classification

Alvaro D. Orjuela-Cañón[1], Diego F. Gómez-Cajas[1],
and Robinson Jiménez-Moreno[2]

[1] GIBIO - Electronics and Biomedical Faculty, Universidad Antonio Nariño,
Bogota D.C., Colombia
{alvorjuela,dfgomezc}@uan.edu.co

[2] GRITEL - Electronics Engineering Deparment, Universidad Autónoma de
Colombia,
Bogota D.C., Colombia
jimenez.robinson@fuac.edu.co

Abstract. A classification of acoustic lung signals for the respiratory disease diagnosis problem is studied in the present work. Models based on artificial neural networks, using Mel Frequency Cepstral Coefficients for training are employed in this task. Results show that neural networks are comparable, and in some cases better, with other classification techniques as Gaussian Mixture Models, that work on the same database.

Keywords: Artificial Neural Networks, Multilayer Perceptron, Mel Frequency Cepstral Coefficients, Acoustic Lung Signals.

1 Introduction

Chronic respiratory diseases (CRD) are a critical problem of health public in developing countries [1]. Diagnostic of this kind of diseases can be a challenge for the medical staff when there are have limited resources in rural regions far from big cities. In this way, new technologic tools can contribute to clinicians and physicians in diagnostic tasks, supporting with additional information, due to that diagnosis varies according to access to medical care of each patient [2].

In addition, for respiratory diseases diagnosis, traditional methods to assess lung functions are based on auscultation. Disadvantages of these methods has to be with the use of stethoscope, because it is a subjective process, depending on stethoscope characteristics and the capabilities of the physician [3]-[4].

Computer-aid decision support systems are commonly used in the biomedical fields due to information that can provide in diagnostic assignments [5]. This information is useful for the medical staff when need extra help. Most of these systems take advantage of previous stored data, and then are used in a procedure known as data mining [6], where artificial neural networks (ANN) are preferred due to flexibility in treatment for any kind of data.

Artificial neural networks are mathematical tools for modeling in high dimensional classification problems, establishing a non-linear relationship between

input variables and known outputs in a supervised learning [7]. Examples of ANN in respiratory diseases diagnosis can be seen in [8]-[10], where clinical and epidemiological variables are used to train the neural models.

For ANN training, datasets are necessary with parameters that represent the input vector. These features can be extracted from data of the patients involved, signal processing representation or image processing parameters. In the present case, lung acoustic signal is acquired and processed by Mel Frequency Cepstral Coefficients (MFCC) to obtain representative parameters of each signal, with this coefficients is built the dataset for ANN training. This signal processing technique has shown good results in the representation of acoustic signals of the respiratory system [11]-[13].

The present work studies the use of ANN as classifier of acoustic lung sounds related with respiratory diseases. Some experiments are presented, based on signal processing features and ANN training. Three different classes are defined, where two represent abnormal sounds and a third class is pointed out as normal. Results are compared with previous studies, which used Gaussian Mixed Models and Support Vector Machines for the classification [11]. Other studies were reviewed in [14], showing methodologies including neural networks, but without a use of MFCC technique.

Section two presents the used database, the signal processing and the classifier implemented, emphasizing in the training of the ANN, where the stage of pattern recognition is developed. Results of the methodology are presented in section three. Finally, a discussion and conclusions are showed in sections four and five.

2   Materials and Methods

In this section, first the used database is presented and signal processing developed over acoustic signals. Then, characteristics about neural network architecture and training are described to show the details of the methodology.

2.1   Database

RALE database [15], developed by the University of Manitoba, Winnipeg, Canada, was utilized in the present study. This repository is composed by recordings obtained from patients who exhibited normal breath, crackles, wheezes and other lung sounds. Database contains at least with 50 labeled recordings, and 24 non-labeled recordings for system tests. These signals were high-pass filtered at 7.5 Hz to suppress any DC offset by using a first-order Butterworth filter. Additionally, a second eighth-order low-pass Butterworth filter at 2.5 kHz was applied to avoid aliasing. All signals were sampled at 10 kHz.

From RALE database, signals from crackles, wheezes and normal recordings were used in the present work. These signals have labels necessary to training of the ANN, and in each of this kind of signals there are more than one register. Table 1 shows the number of signals for each class.

Signals from crackles represent discontinuous explosive adventitious lung sounds, and are obtained from patients with cardio-respiratory disorders.

Characteristics of these sounds are the time length, less than 20 ms, and frequency range, which is typically from 100 up to 2000 Hz [16].

The waveform of the wheeze signals is similar to a sine wave with fundamental frequency around 100-2000 Hz and. Time length for this kind of signals varies between 80 to 250 ms [16].

Table 1. Registers for each class

Class of lung signal	Number of registers
Crackles	4
Wheezes	4
Normal	5

2.2   Mel Frequency Cepstral Coefficients

Mel Frequency Cepstral Coefficients is a representation from speech signals based on perception of human being. MFCC use Discrete Fourier Transform (DFT) and the Discrete Cosine Transform. The main difference is that the bands are placed in a logarithmic way, according with the Mel scale. In this way, the speech is modeled by an answer more human, allowing a signal processing more efficient [17].

For computing of MFCC first the signal is segmented in frames and the DFT is calculated. Then, the spectrum is filtered using thirteen triangular windows corresponding to the Mel frequency scale. Logarithmic functions are applied for each energy computed in Mel frequency bands and Discrete Cosine Transform (DCT) is used for each log-energy. Finally, the MFCC correspond to the amplitude spectrum provided after DCT (Figure 1).
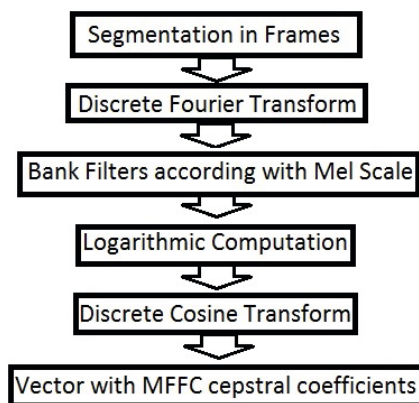

Fig. 1. MFCC process

In the present work, signal was divided in frames of 30 ms with a frame shift of 10 ms. For each frame, thirteen coefficients are computed to represent the acoustic lung signal due to performance obtained in previous studies [11], [13], [17].

2.3   Neural Network Training

Neural networks have the ability of learning complex non-linear patterns, based on input-output relationships, adjusting a set of free parameters known as synaptic weights. The most common type of networks employed in classification tasks are the Multilayer Perceptrons (MLPs), which have only feed forward connections and are trained in a supervised way [7]. In the present work, MLPs are employed to learn the patterns that represent the waveform of acoustic signal.

MLP architecture consists of an input, a hidden layer and a output layer. Number of nodes in the input is given by the 13 coefficients generated by the MFCC algorithm. Number of units in the hidden layer was heuristically studied, trying since one to fifteen units. In the output layer were used three units, each one identifies each pattern. More layers are not used because two layers are enough for a typical problem of classification [7].

All data were normalized maintaining all data between [0-1] interval for avoiding the saturation of values in the synaptic weights of network. This contributes to avoid the saturation of the values in the synaptic weights, which are initialized in a random manner.

Hyperbolic tangent functions are used in each unit as activation functions. In this way, the units in the output have the value of one when the input belongs to this class, and -1 when does not belong. Training for MLP is developed in a supervised manner, where patterns and its labels are necessary. The resilient backpropagation algorithm was used for the training of the network, due to its fast and stable convergence [18]. Training set must represent the statistics of the data, and its construction is important for a good training, because the network must generalize its performance to new inputs.

In the present study, the database has a limited size, making difficult the division of the data for training and validation. Alternatives as bootstrapping and cross validation methods can be useful in these cases extracted from applications on digital signal processing [19] - [21]. An example of this kind of methods is the Leave One Out (LOO) method, where the performance of pattern classification models can be evaluated based on LOO error [22], [23]. The applied method consists in training the network with all data, leaving only an observation out. When training was finished, the validation is calculated using just with the observation not included in the training. Finally, there is many models as observations of database. For this reason the method is employed with databases with limited observations.

The LOO error is a statistical estimator of the behaviour when a learning algorithm is used, and it is very useful for model selection because is slightly biased, despite its empirical error. Also, when the algorithm is stable, LOO error is low [22], [23]. The LOO error can be calculated by using:

$$E_{LOO}(f_D) = \frac{1}{m} \sum_{i=1}^{m} \ell(f^i, z_i) \tag{1}$$

where m is the number of samples in the D set, composed by the z_i elements, and it is built the learning algorithm, f^i is the function obtained after training. These methods have been used in applications where models of regression or structures in time series are required.

As each neural networks is different due to initialization of the synaptic weights, for each architecture were implemented 100 different initializations to observe the statistical results due to these parameters.

2.4   Decision Criteria

As mentioned, each signal was divided in frames to extract the MFCC values, each frame was presented as an input vector to the ANN. The classification was determined according with the number of activations given by the frames. In this way, lung signal was classified according with the class with the most number of activations in each neuron of the output of the network.

3   Results

Results for different ANN architectures are presented in Table 2. In each case, the number of neurons in the hidden layer was modified and the performance of network was computed. According with LOO method, validation is obtained with the sample left out. This means that was trained one neural network for each acoustic lung signal, obtaining thirteen networks for each architecture.

Table 2. Results for different number of neurons in the hidden layer

Number of Neurons	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Crackles	0%	**75%**	75%	75%	75%	75%	75%	75%	75%	75%	75%	75%	75%	75%	75%
Wheezes	50%	**100%**	100%	50%	75%	50%	75%	75%	50%	75%	50%	50%	50%	50%	50%
Normal	100%	**80%**	80%	80%	80%	80%	80%	80%	80%	80%	80%	80%	80%	80%	80%
Total	54%	**85%**	85%	70%	77%	70%	77%	77%	70%	77%	70%	70%	70%	70%	70%

For each signal, the output of the network was computed by each frame and with the most class activated, the classification was determined.

Figure 2 shows box plots for different initializations. In this way, the dispersion as effect of random initial weight values is displayed, the red line point out the median of the results for 100 different trainings.

Best result was reached with a neural network with two or three neurons in the hidden layer. In this case, the performance of classification was 85%. For the
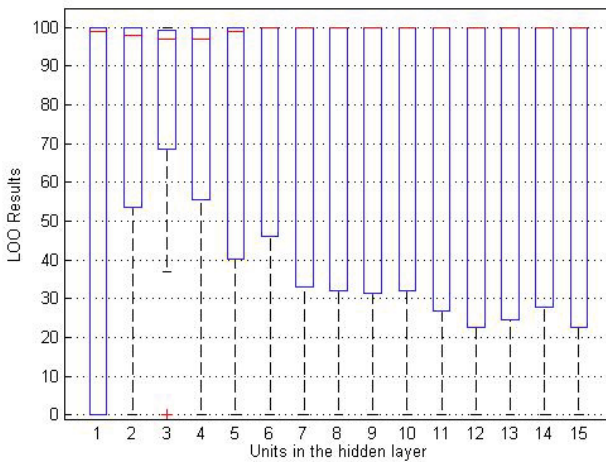
Fig. 2. Results with 100 initializations

crackles class is reached a 75% of accuracy, and for wheezes class all signals were correctly classified reaching an accuracy of 100%. The normal class obtained a performance of 80% of accuracy.

4   Discussion

As mentioned, the LOO method attempts to obtain results in terms of efficiency of the method in a general way. This means that there is not a only one network that solve the classification problem, there is a study of models based on neural networks for the classification methodology.

From Table 2 it is possible to see that the best performance for the classification is obtained when two or three neurons were used in the hidden layer. This can be observed in terms of dispersion (figure 2), where lower values are reached with two and three neurons, too. Models with less complexity are preferable, according with parsimonious principle, then neural networks with two neurons in the hidden layer are enough in this case. Models with more complexity are not necessary, because the network can learn the patterns in the signal representation. At same time, the results show that the wheezes class is better classified, where all signals were a correctly classified. In fact, results for two and three neurons in the hidden layer yield less dispersion when 100 trainings with different initialization were implemented (figure 2).

It is possible to see that when the number of neurons in the hidden layer was increased, the results decreased, this can be caused by overfitting in training, making the network be specialized in the subset training. This happens when models have higher complexity. This can be observed in high dispersion values for different initializations (figure 2).

As this classification is focused in the respiratory diseases problem, it is possible compute results in terms of sensitivity and specificity, taking results for

normal and abnormal signals (crackles and wheezes) these measures were obtained. In the sensitivity case, the method reaches an 87% and for specificity it is possible to reach an 80%.

Results obtained in the present work are comparable with results from previous studies [11],[13] where neural networks showed a better performance in the wheezes and normal cases. Despite of this, results for crackles class, GMM is better than ANN by 15% of accuracy. More databases and experiments are needed to conclude about the best classifier. It is important to note, than neural networks are not parametric estimation models, being an advantage for ANN over GMM models, which assumes that data have Gaussian distribution.

Complementary experiments with other databases would provide more information about the use of present methodology. Databases with more samples give the possibility of study other kind of validation methods, without the limitation here exposed in that sense.

5   Conclusions

It was presented an alternative for classification of acoustic lung signals for the respiratory disease diagnosis problem. Results suggest that neural networks can be implemented in aid support diagnosis system, when MFFC are used in the parametric of the signal.

Complex models are not necessary due to the neural network learns the patterns presented using an architecture of thirteen inputs, two neurons in the hidden layer and three neurons in the output layer.

References

1. Alman, A.: World Health Organization, Global status report on noncommunicable diseases, pp. 1–176 (2010-2011)
2. Jamison, D.T., Breman, J.G., Measham, A.R.: Disease Control Priorities in Developing Countries, 2nd edn. World Bank (2006)
3. Sovijärvi, A.R.A., Malmberg, L.P., Charbonneau, G., Vanderschoot, J., Dalmasso, F., Sacco, C., Rossi, M., Earis, J.E.: Characteristics of breath sounds and adventitious respiratory sounds. Eur. Respir. Rev. 10(77), 591–596 (2000)
4. Schreur, H.J.W., Vanderschoot, J., Zwinderman, A.H., Dijkman, J.H., Sterk, P.J.: Abnormal Lung Sounds in Patients with Function Asthma during Episodes with Normal Lung. 106(1), 91–99 (1994), doi:10.1378/chest.106(1).91 ISSN:0012-3692
5. Belle, A., Kon, M.A., Najarian, K.: Biomedical Informatics for Computer-Aided Decision Support Systems: A Survey. The Scientific World Journal (2013)
6. Kumar, D.S., Sathyadevi, G., Sivanesh, S.: Decision Support System for Medical Diagnosis Using Data Mining. International Journal of Computer Science 8(3(1)) (May 2011)

7. Haykin, S.: Neural Networks and Learning Machines, 3rd edn. Pearson -Prentice Hall (2008)
8. Er, O., Termutas, T., Tanrikulu, A.C.: Tuberculosis Disease Diagnosis Using Artificial Neural Networks. Journal of Medical Systems 34, 299–302 (2010)
9. Elveren, E., Yumusak, N.: Tuberculosis Disease Diagnosis Using Artificial Neural Network Trained with Genetic Algorithm. Journal of Medical Systems 35, 329–332 (2011)
10. Seixas, J.M., Mello, F.C., Santos, A.M., Pereira, B.B., Kristki, A.L.: Neural networks: an application for predicting smear negative pulmonary tuberculosis, Boston, USA, Birkhuser. Advances in Statistical Methods for the Health Sciences, pp. 279–292 (2006)
11. Mayorga, P., Druzgalski, C., Morelos, R.L., Gonzlez, O.H., VIdales, J.: Acoustic Based Assessment of Respiratory Diseases using GMM Classification. In: 32nd Annual International Conference of the IEEE EMBS (2010)
12. Abushakra, A., Faezipour, M.: Acoustic Signal Classification of Breathing Movements to Virtually Aid Breath Regulation. IEEE Journal of Biomedical and Health Informatics 17(2) (2013)
13. Mayorga, P., Druzgalski, C., Gonzlez, O.H., Zazueta, A., Criollo, M.A.: Expanded Quantitative Models for Assessment of Respiratory Diseases and Monitoring. In: Pan American Health Care Exchanges PAHCE Conference (2011)
14. Gurung, A., Scrafford, C.G., Tielsch, J.M., Levine, O.S., Checkley, W.: Computerized lung sound analysis as diagnostic aid for the detection of abnormal lung sounds: A systematic review and meta-analysis. Respiratory Medicine 105, 1396–1403 (2011)
15. Database RALE Univeristy of Manitoba, Canada, `http://www.rale.ca/`
16. Charbonneau, G., Ademovic, E., Cheetham, B.M.G., Malmberg, L.P., Vanderschoot, J., Sovijrvi, A.R.A.: Basic techniques for respiratory sound analysis. Eur. Respir. Rev. 10(77), 625–635 (2000) ISSN 0905–9180
17. Rabiner, L., Biing-Hwang, J.: Fundamentals of Speech Recognition. Prentice Hall (1993) ISBN 0 13 015157
18. Riedmiller, M.: Rprop: Description and Implementation Details. Technical report, University of Karlsruhe (1994)
19. Kohavi, R.: A study of cross-validation and bootstrap for accuracy estimation and model selection. In: Proceedings of the 14th International Joint Conference on Articial Intelligence, pp. 1137–1143 (1995)
20. Zoubir, A., Boashash, B.: The Bootstrap and its Application in Signal Processing. IEEE Signal Processing (January 1998)
21. Zoubir, A., Iskander, R.: Bootstrap Techniques for Signal Processing, 1st edn. Cambridge University Press, Cambridge (2004)
22. Elisseeff, A.: Leave-one-out error and stability of learning algorithms with applications. Nato Science Series Sub Series III Computer And Systems Sciences, vol. 190, pp. 111–130 (2003)
23. Zhang, T.: Leave-one-out Bounds for Kernel Methods. Neural Computation 15(6), 1397–1437 (2003)