Achmad Rizal[1,2,*], **Risanuri Hidayat**[1], **Hanung Adi Nugroho**[1]

J. Electrical Systems 12-3 (2016): 556-564

Journal of Electrical Systems

Regular paper

Multi-scale Grey-Level Difference for Lung Sound Classification

Lung sounds have information to seek abnormalities in the lung. With digital signal processing, the information in the lung sounds is extracted as the features in lung sound classification. In this paper, texture analysis was used to measure the complexity of lung sound as a feature in lung sound classification. Grey-Level Difference (GLD) method was performed on lung sounds with a number of different scales. Multi-scale GLD has produced accuracy up to 90.12% for five classes of data. Further, gradient entropy individually provided the highest accuracy up to 91.36% for the distance D = 20 and a scale of 1-10.

Keywords: Grey-level difference; texture analysis; multi-scale; classification; lung sound

Article history: Received 16 December 2016, Accepted 15 August 2016

1. Introduction

The lungs generate an audio signal in the form of lung sound. The lung sounds are used as information to determine the health of the lungs. Even though many devices such as X-rays and CT scans have been developed to diagnose lung diseases, lung sounds remain the first choice for the diagnosis of pulmonary disease [1]. Various feature extraction methods have been developed to automatically classify lung sounds. Some researchers have used a number of texture analysis techniques for lung sounds feature extraction. Here, lacunarity is one method often used for lung sounds analysis [2, 3]. Other frequently methods used are the statistical parameters such as skewness and kurtosis to distinguish among lung sound types [2, 4].

Some researchers revealed that the biological signal parameter measurements on a single scale are not sufficient in view of the complexity of biological signals [5]. Multi-scale measurement on the signal then is required to see the characteristics of the signal at some different scales. The techniques used by researchers to look at the characteristics of the signals on the multi-scale for example include multi-scale entropy (MSE) [6], multi-scale permutation entropy (MPE) [7] and multi-scale compression entropy (MCE) [8]. Overall, those methods are based on entropy measurement. Usage on multi-scale texture analysis for biomedical signal classification has not been done before. Therefore, feature extraction method based on multi-scale texture analysis is used for lung sound classification.

In this research, we modified grey-level difference (GLD) that is usually used for texture analysis [9]. Following this, we used the modified GLD on the multiscale scheme for lung sound features extraction. Feature extraction with GLD at various scales is expected to improve the accuracy and could indicate lung sound signal characteristics comprehensively.

[*] Corresponding author: Achmad Rizal, Department of Electrical Engineering and Information Technology, Universitas Gadjah Mada, 55281 Yogyakarta, Indonesia, E-mail: rizal.s3te14@mail.ugm.ac.id

[1] Department of Electrical Engineering and Information Technology, Universitas Gadjah Mada, 55281 Yogyakarta, Indonesia

[2] School of Electrical Engineering, Telkom University, 40257 Bandung, Indonesia

2. Notation

The notation used throughout the paper is stated below.

Indexes:

i, j	pixel index
τ	scale index
g	Absolute difference value of two pixels with d distance at θ direction

Constants:

θ	Direction at degree
$y_j^{(\tau)}$	Signal y on τ scale
$h(g\|\theta)$	Probability of pixel value g at θ direction
d	Pixel distance
GC	Gradient contrast
GSM	Gradient second moment
GE	Gradient entropy
GM	Gradient mean
IDM	Inverse-difference moment

3. Material and Methods

3.1. Lung Sound Data

Lung sound data used in this study has been divided into five data classes: normal bronchial, asthma, crackle, pleural rub and stridor. The same data was also used in previous studies [10]. The number of data in each class was in the range of 13-20. Normal lung sounds used were the bronchial sound. Normal bronchial was heard over trachea and larynx during the inspiration and expiration [11]. In particular circumstances, bronchial sound regarded as pathological sound [11].

In the asthma disease, the produced lung sounds were wheeze, an adventitious sound that is continuous, musical and has a dominant frequency > 100Hz [12]. Commonly, wheeze duration takes more than 100ms. Meanwhile, the pleural rub is the sound that arises due to the friction between the visceral pleura and parietal pleura [11]. Pleurisy is an example of diseases that produce pleural rub [11]. Crackle, in contrast, is an adventitious sound that is discontinuous, non-musical and very short duration [13]. One of diseases that produce this sound is pulmonary fibrosis. Stridor is a wheeze with high pitch sound that comes from the larynx or trachea [12]. It usually occurs in the obstruction or the inflammation of the trachea [14].

The data used has the *.wav format with a sampling frequency of 8000 Hz and 16-bit depth. The length of the data is one respiratory cycle and about 30000-34000 sample data length.

3.2. *Grey-Level Difference Matrix*

Grey-Level Difference Matrix (GLDM) is one of texture measurement techniques proposed by Weszka, *et al.* [9]. It has originally been intended to differentiate the aerial photograph of an area such as lakes, forests, road train, urban, wetlands and rural areas [9]. At GLDM, the absolute value of the difference between two pixels at a distance d is measured. For the horizontal direction, it can be calculated $y(i,j) = |x(i+d,j) - x(i,j)|$

where $x(i,j)$ is input image, d is distance between pixels and $y(i,j)$ is output image produced by GLD at 0° direction. On GLD, pixel difference is calculated in vertical and diagonal direction to measure the texture of an image. If θ is the direction of adjacent pixels with the distance d, the absolute value of its pixels can be expressed as $H(g|\theta)$, the probability of each pixel can be expressed as $h(g|\theta) = \frac{H(g|\theta)}{\sum_g H(g|\theta)}$.

The features of GLDM include gradient contrast (GC), gradient second moment (GSM), entropy gradient (GE), gradient mean (GM) and inverse difference moment (IDM) [9,15]. All features are expressed in equation (1)-(5) [16].

$$GC = \sum_g g^2 h(g|\theta) \quad (1)$$

$$GSM = \sum_g [h(g|\theta)]^2 \quad (2)$$

$$GE = -\sum_g h(g|\theta).logh(g|\theta) \quad (1)$$

$$GM = \sum_g h(g|\theta).g \quad (2)$$

$$IDM = \sum_g \frac{h(g|\theta)}{(g^2+1)} \quad (3)$$

The one-dimensional signal of length N can be considered as the image with 1xN dimension. Thus, GLDM can only be measured in 0° direction. Value at $H(g|\theta)$ is quantized, so $H(g|\theta)$ has N bin value.

3.3. Coarse-Grained Procedure

The multi-scale method used in this research adopted multi-scale as in [5]. The multiscale process used a procedure called the coarse-grained procedure [6]. Mathematically, coarse-grained procedure can be written as in equation (6) [17]:

$$y_j^{(\tau)} = \frac{1}{\tau}\sum_{i=(j-1)\tau+1}^{j\tau} x_i \;, \; 1 \le j \le \frac{N}{\tau} \quad (6)$$

In this study the scale used was in the range of 1-20. In scale τ=1, $y_j^{(1)} = x(i), i = 1,2,\dots,N$ if τ=2, then $y_i^{(2)} = \frac{[x(2i-1)+x(2i)]}{2}$, meanwhile if τ=3, then $y_i^{(3)} = \frac{[x(3i-2)+x(3i-1)+x(3i)]}{3}$ and so on. In the multi-scale GLD, GLD was calculated at each scale to obtain five features on each scale. In this study, we used 20 scales, so 100 features were obtained for each data [5].

3.4. Classifier and Validation

As a classifier, we used multilayer perceptron (MLP). It is the simplest form of artificial neural networks and consists of three layers: input layer, hidden layer, and output layer. Input layer contains neuron as many as numbers of the features that will be classified and output layer contain neurons as many as the number of classes. The number of neurons in the hidden layer is determined by trial and error [18].

Validation used the N-fold cross validation (N = 3) where the data was divided into three data sets. Two data sets were used as training data while one dataset was used as test data. The process was repeated so that each of the data sets was used as test data once. Total accuracy was the average of the thorough testing [19].

In addition to accuracy, sensitivity (SE) and specificity (SP) were other parameters to be tested. Sensitivity was defined as the number of data with the positive conditions recognized correctly to have the condition (True Positive) and divided with all positive condition data (True Positive + False Negative) [12]. Specificity, meanwhile, is defined as the number of data with the negative conditions recognized correctly to have the negative condition (True Negative) and divided with all negative condition data (FP + TN). SE and SP mathematically can be formulated as follows:

$$SE = \frac{\text{True Positive}}{\text{False Negative+True Positive}} \tag{7}$$

$$SP = \frac{\text{True Negative}}{\text{False Positive+True Negative}} \tag{8}$$

4. Result and Discussion

4.1. Multi-scale GLD

At multi-scale GLD, due to scale (τ) used was 20; the number of features was 100. The distance between data sample being tested (D) was 1-5, 10, 20, 50, 100, and 200. Accuracy was generated for each D with the number of hidden neurons 0-75 MLP as shown in Fig. 1 and Fig. 2.

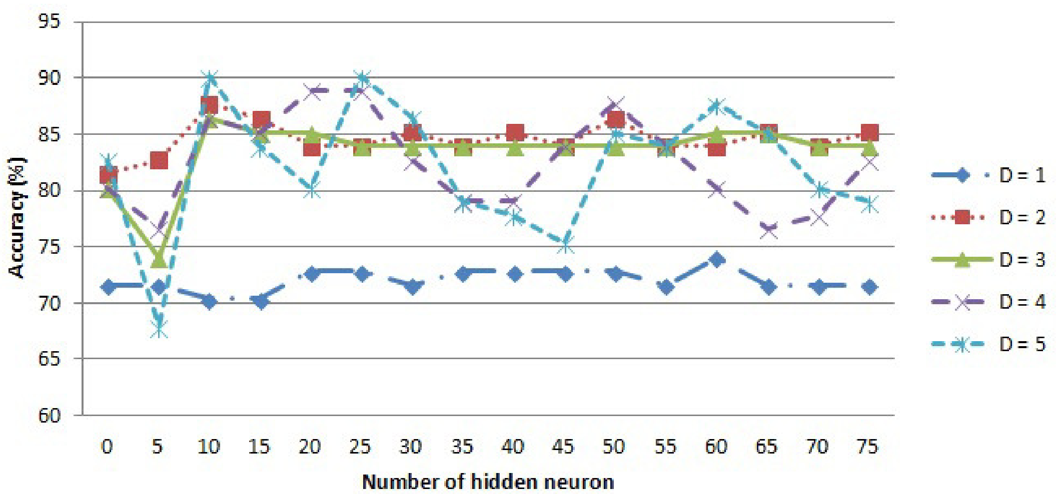

Fig 1. Accuracy for Multi-scale GLD with D =1-5 and scale τ = 20 with various numbers of hidden neurons on MLP

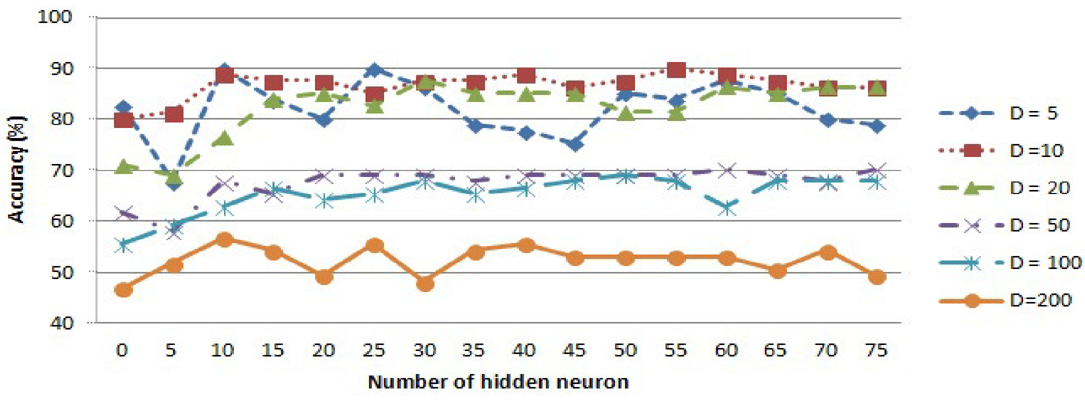

Fig 2. Accuracy for Multi-scale GLD with D =5, 10, 20, 50, 100, 200, and scale τ = 20 with various number of hidden neuron on MLP

The highest accuracy was achieved in MLP configuration of 100-10-5 and 100-25-5, D = 5 with an accuracy of 90.12%. The smaller the distance D is, the more decrease the accuracy will be. In contrast, the larger the distance D, the more decrease the accuracy will be. It can be seen that for the value D = 1 and D = 200, and the accuracy < 75%. The distance D = 5 produced the highest accuracy = 90.12% with a mean = 82.18 ± 5.75%.

Table 1 presents SE and SP values for the highest accuracy conditions. Normal data, crackle, and stridor had the SE value of 100%; indicating that all data have been classified correctly. In the normal data and stridor, some data from other classes were recognized as normal or stridor so that the SP did not reach 100%. The best results were achieved by crackle where the values of the SE and SP reached 100%. Meanwhile, the worst SE value was found in asthma; showing that only a few asthma data were identified as asthma while the SP 100% indicated that there was no data in other class of data classified as asthma.

Table 1 SE and SP of each class data for multi-scale GLD in highest accuracy (D = 5, MLP 100-10-5)

Data Class	**Sensitivity (%)**	**Specificity (%)**
Normal	100	96.8
Asthma	61.54	100
Crackle	100	100
Friction rub	80	95.45
Stridor	100	95.08

4.2. Effect of Each GLD Parameter in Various Scales

The accuracy was measured using each of the GLD parameters for different scale with distance D = 5 and MLP parameters = N-10-5. It was intended to look the effect of each feature of GLD. GE on a scale of 1-10 yielded the highest accuracy of 91.36% as seen in Fig. 3.

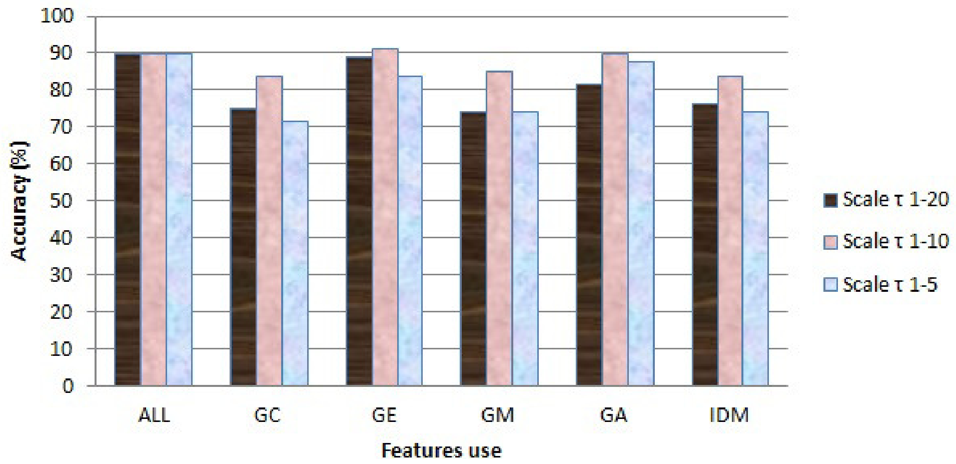

Fig 3. Accuracy of each features in different scale for distance D = 5,

GC = gradient contrast, GE = gradient entropy, GM = gradient mean, GA = gradient angular second moment, IDM = inverse-difference moment

4.3 Discussion

Feature extraction process consisted of two important parameters in the features generation. The first parameter was the scale τ in the process of coarse-grained procedure and the second one was the distance D at GLDM. Scale τ formed a series of new signals as in equation (6). Equation (6) was similar to the LPF followed by down-sampling by a factor of τ. The larger the scale τ, the fewer the number of data samples and the smaller the variance of the signal. This factor caused a bigger scale τ making accuracy obtained not higher - especially if only one feature was used in the classification process. It can be seen from Fig. 6 when only one parameter was used as the feature, the accuracy on a scale of 1-20 was always lower than the accuracy of scale 1-10. The selection of the amount of scale was made by trial and error.

The next parameter was the distance D. The smaller the distance D; the smaller the signal variation as well as when the distance D was too large. Fig. 1 and Fig. 2 show that the distance D = 1 and distance D = 200 produced the lowest accuracy. On the average distance, distance D = 2 and D = 10 provided a high accuracy, though not the highest. The highest accuracy was achieved for a distance D = 5; indicating that the selection of D should be made by trial and error.

In general, a multiscale scheme in this research used the coarse-grained procedure as it has been done in [6] as expressed in equation (6). Equation (6) was similar to FIR filters with equation (9) [20]:

$$y_j^{(\tau)} = \frac{1}{\tau}\sum_{k=0}^{\tau-1} x(j-k) \;\;,\;\; 1 \le j \le N \qquad (9)$$

N is the length of signal and τ is scale. Valencia *et al.* stated that the frequency response of the filter was poor due to the appearance of side lobe in the stop band [17,20]. The standard deviation of the signal on the scale (τ) would decline sharply [21]; this made the signal flatter at a high scale. This study used scale (τ) = 1-20, the value of GLD features only produced the highest accuracy of 90.12% for the use of all the features and accuracy of 91.36% for GE features on a scale of 1-10.

From the results obtained, Gradient entropy (GE) was the most dominant feature; enabling it to individually produce the higher accuracy compared to the accuracy of the overall features. The result suggested that one of the main parameters to distinguish the sound of the lung is the entropy. Compared with the sample entropy (SE) method, which is used in [22] to detect a wheeze, GE on scale 1-10 has a simpler computation and can distinguish more classes of data.

Other researches on lung sound analysis using entropy are presented in [2, 23, 24, 25]. Mondal, *et al.* used sample entropy, skewness and kurtosis to distinguish normal and abnormal lung sound [2]. In [23], normalized bi-spectral entropy and normalized bi-spectral squared entropy were used as the features of lung sound. In [24], Zang et al. used short-time Fourier transform (STFT) and calculated the entropy of the lung sound from STFT result. The method was used for wheeze detection and produced 80% of accuracy for normal lung sound and 84.4% for wheeze sound. Meanwhile, Aydore et al. used Renyi entropy for lung sound feature extraction [25]. Combined with other features, Renyi entropy produced 93% of accuracy.

Table 2. Comparison with other methods

No	Data set	Scale	Feature	Number of feature	Classifier	Accuracy	Reference
1.	120 data, 2 classes	1	Skewness, kurtosis, sample entropy, lacunarity	4	SVM, ELM	92.86 %	Mondal, *et al.* [2]
2.	140 data, 2 classes	7 , using wavelet	Mean absolute values, average power, standard deviation, Ratio of the absolute mean values of adjacent sub-band on DWT, skewness, kurtosis	46	MLP	90%	Hashemi, *et al.* [4]
3.	Normal, pneumonia and asthma	1	Peak locations of bi-spectrum, secondary and third maximum points of bi-spectrum, amplitude of peaks, Energy of slice spectrum for peak, entropies of bi-spectrum	14	empiric	N.A	Sheng-jun and Yi [23]
4.	7 normal, 7 asthma	1	Sampling entropy, histogram distortion	2	Euclidean distance	85.3-97.9%	Jin *et al.* 2008 [22]
5.	7 normal, 14 pathology, plus data from database 5 normal, 19 pathology	1	averaged instantaneous kurtosis, Discriminating function, sample entropy, histogram distortion,	4	SVM	97.7-98.8%	Jin *et al.* 2014 [27]
6..	7 COPD	1	Kurtosis, renyi entropy, mean crossing irregularity	1	Discriminant analysis	93.5-95.1%	Aydore [25]
7.	8 healthy and 8 extrinsic allergic alveolitis patients,	4	Multi-scale entropy	12	Statistical test	N.A	Charleston-Villalobos, et al. [26]
8.	18 bronchial, 13 asthma, 15 crackle, 15 pleural rub, 20 stridor	10	Gradient entropy	10	MLP	91.36%	Proposed method

Another multi-scale approach for classification of lung sounds was used in [4, 26]. Hashemi, *et al.* used skewness, kurtosis and other features for classification wheeze on lung sounds [4]. Lung sounds were decomposed using a wavelet to order 7. Accuracy was achieved 90%, used 46 features [4]. Meanwhile, Charleston-Villalobos, *et al.* used multi-scale sample entropy to recognize the voice of the lung in patients with extrinsic allergic alveolitis (EAA) [26]. The paper used scale $\tau = 4$ and tolerance window r = 0.1, 0.15, 0.2 of standard deviation. This study did not explain the accuracy obtained for only using a statistical analysis [26]. Table 2 shows the comparison of the proposed method with other methods.

Even worse if compared with some previous researches, our proposed method offers some advantages. GE calculation was done on time series directly, not requiring a transformation into another signal domain. It will save some computation time. Our proposed method is able to classify more classes of data with relatively less number of features. Some works to do next is to try another entropy calculation on GLD to improve the accuracy of the system.

5. Conclusion

This paper presents GLD measurements on one-dimensional signals with a multi-scale scheme. By using multi-scale scheme it is expected that lung sound’s signal complexity can be measured on a different scale. Thus, the difference between the data classes will be more prominent. GE on a scale of 1-10 yielded 91.36% accuracy. The measurement of GLD in multi-scale provided more information compared to than the GLD on a single scale. Using another multi-scale scheme may be needed to improve the accuracy of GLD multi-scale system.

References

[1] H. Melbye, Auscultation of the lungs: still a useful examination,Tidsskr. Nor. Laegeforen., vol. 121, no. 4, pp. 451–454, 2001.

[2] A. Mondal, P. Bhattacharya, and G. Saha, Detection of lungs status using morphological complexities of respiratory sounds., Science World Journal, vol. 2014, pp. 1829–38, Jan. 2014.

[3] L. J. Hadjileontiadis, “A texture-based classification of crackles and squawks using lacunarity.,” IEEE Transaction on Biomedical Engineering, vol. 56, no. 3, pp. 718–32, Mar. 2009.

[4] A. Hashemi, H. Arabalibiek, and K. Agin, Classification of Wheeze Sounds Using Wavelets and Neural Networks, in International Conference on Biomedical Engineering and Technology (2011), vol. 11, pp. 127–131. 2011

[5] M. Costa, A. L. Goldberger, and C.-K. Peng, Multiscale entropy analysis of complex physiologic time series., Phys. Rev. Lett., vol. 89, no. 6, p. 068102, 2002.

[6] M. Costa, A. L. Goldberger, and C. K. Peng, Multiscale entropy analysis of biological signals, Phys. Rev. E - Stat. Nonlinear, Soft Matter Phys., vol. 71, pp. 1–18, 2005.

[7] W. Aziz and M. Arif, Multiscale Permutation Entropy of Physiological Time Series, in 9th International Multitopic Conference, IEEE INMIC 2005, 2005.

[8] A. Humeau-heurtier, M. Baumert, G. Mahé, and P. Abraham, Multiscale Compression Entropy of Microvascular Blood Flow Signals: Comparison of Results from Laser Speckle Contrast and Laser Doppler Flowmetry Data in Healthy Subjects, Entropy, vol. 16, pp. 5777–5795, 2014.

[9] J. S. Weszka, C. R. Dyer, and A. Rosenfeld, A Comparative Study of Texture Measures for Terrain Classification, Syst. Man Cybern. IEEE Trans., vol. SMC-6, no. 2, pp. 269–285, 1976.

[10] A. Rizal, R. Hidayat, and H. A. Nugroho, Determining Lung Sound Characterization Using Hjorth Descriptor, in International Conference on Control, Electronics, Renewable Energy, and Communications 2015 (ICCEREC 2015), 2015, pp. 20–23. 2015
[11] A. Bohadana, G. Izbicki, and S. S. Kraman, Fundamentals of lung auscultation., N. Engl. J. Med., vol. 370, no. 8, pp. 744–51, Feb. 2014.
[12] S. Reichert, R. Gass, C. Brandt, and E. Andrès, "Analysis of Respiratory Sounds : State of the Art," Clin. Med. Circ. Respir. Pulm. Med., vol. 2, pp. 45–58, 2008.
[13] B. Flietstra, N. Markuzon, A. Vyshedskiy, and R. Murphy, Automated Analysis of Crackles in Patients with Interstitial Pulmonary Fibrosis, Pulmonary Medicine, vol. 2011, 2011.
[14] D. Arnall, "Breath Sounds." [Online]. Available: http://jan.ucc.nau.edu/~daa/heartlung/breathsounds/contents.html. [Accessed in 01-May-2015].
[15] R. W. Conners and C. a Harlow, A theoretical comparison of texture algorithms., IEEE Trans. Pattern Anal. Mach. Intell., vol. 2, pp. 204–222, 1980.
[16] A. Goshtasby, Image Registration, Priciples, Tools and Methods, First. Springer-Verlag London Limited, 2012.
[17] A. Humeau-heurtier, The Multiscale Entropy Algorithm and Its Variants: A Review, Entropy, vol. 17, pp. 3110–3123, 2015.
[18] R. Palaniappan, Biological Signal Analysis. Ventus Publishing ApS, 2010.
[19] T. Andersen and T. Martinez, Cross validation and MLP architecture selection, IJCNN'99. Int. Jt. Conf. Neural Networks. Proc., vol. 3, pp. 1614–1619, 1999.
[20] J. F. Valencia, A. Porta, M. Vallverdu, F. Claria, R. Baranowski, E. Orlowska-Baranowska, and P. Caminal, Refined Multiscale Entropy : Application to 24-h Holter Recordings of Heart Period Variability in Healthy and Aortic Stenosis Subjects, IEEE J. Biomed. Eng., vol. 56, no. 9, pp. 2202–2213, 2009.
[21] V. V Nikulin and T. Brismar, Comment on ""Multiscale Entropy Analysis of Complex Physiologic Time Series,"' Phys. Rev. Lett., vol. 92, no. 8, p. 89803, 2004.
[22] F. Jin, F. Sattar, and D. Y. T. Goh, Automatic wheeze detection using histograms of sample entropy.,in Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society, 2008, vol. 2008, pp. 1890–3. 2008
[23] L. Sheng-jun and L. Yi, Feature Extraction of Lung Sounds Based on Bispectrum Analysis, in Third International Symposium on Information Processing Feature, 2010, pp. 393–397. 2010
[24] J. Zhang, W. Ser, J. Yu, and T. T. Zhang, A Novel Wheeze Detection Method for Wearable Monitoring System, in 2009 International Symposium on Itelligent Ubiquitous Computing and Education, 2009, vol. 639798, pp. 1–4. 2009
[25] S. Aydore, I. Sen, Y. P. Kahya, and M. K. Mihcak, Classification of respiratory signals by linear analysis., in Annual International Conference of the IEEE Engineering in Medicine and Biology Society., 2009, vol. 2009, pp. 2617–20. 2009
[26] S. Charleston-Villalobos, L. Albuerne-Sanchez, and R. Gonzalez-Camarena, Linear and Nonlinear Analysis of Base Lung Sound in Extrinsic Allergic Alveolitis Patients in Comparison to Healthy Subjects, Methods Inf Med, vol. 52, pp. 266–276, 2013.
[27] F. Jin, F. Sattar, and D. Y. T. Goh, New approaches for spectro-temporal feature extraction with applications to respiratory sound classification, Neurocomputing, vol. 123, pp. 362–371, Jan. 2014.