Sleep snoring detection using multi-layer neural networks

Tan Loc Nguyen and Yonggwan Won[*]
School of Electronics and Computer Engineering, Chonnam National University, 77 Yongbong ro, Buk-gu, Gwangju 500-757, Korea

Abstract. Snoring detection is important for diagnosing obstructive sleep apnea syndrome (OSAS) and other respiratory sleep disorders. In general, audio signal processing such as snoring sound analysis uses the frequency characteristics of the signal. Recently, a correlational filter Multilayer Perceptron neural network (f-MLP) has been proposed, which has the first hidden layer of correlational filter operations in frequency domain. It demonstrated a superior classification performance for the pattern sets; of these, frequency information is the dominant feature for classification. The first hidden layer is implemented with the correlational filter operation; its output is the power spectrum of the filter output, while the other layers are the same as the ordinary multilayer Perceptron (o-MLP). By using the back-propagation learning algorithm for the correlational filter layer, f-MLP was able to self-adapt the filter coefficients to produce its output with more discrimination power for classification in the higher layer. In this research, this f-MLP was applied for sleep snoring signal detection. As a result, the f-MLP achieved an average detection rate of 96% for the test patterns, compared to the conventional multilayer neural network that demonstrates an 82% average detection rate.

Keywords: Sleep snoring, Fourier transform, correlational filter, classification, multilayer perceptron neural network

1. Introduction

In the last 20 years, snoring has been a popular topic in clinical medicine. About 50% of the adult population snores frequently. Snoring has been also reported as a risk factor for developing disorders such as ischemic brain infraction, systemic arterial hyper-arterial hypertension, coronary artery disease, and sleep disturbance [1]. Several studies have shown the relationship between snoring and OSAS, which is usually associated with loud and heavy snoring [1]. Snoring also disrupts sleep for partners of snoring. Investigation into sleep sounds also provides information about breathing abnormalities, OSAS, or other pathologies such as upper airway resistance syndrome. It also supports health assessment [2].

For several decades, many studies have been performed on snoring detection. The detector usually contains a pre-processing block and a following classifier. The pre-processing procedure is usually composed of one or more steps of filtering, segmentation, feature extraction, and normalization for the sleep sound to be in a suitable form for the classification process [3]. Based on the discriminative

---
* Address for correspondence: Yonggwan Won, School of Electronics and Computer Engineering, Chonnam National University, 77 Yongbong ro, Buk-gu, Gwangju 500-757, Korea. Tel.: +82-62-530-1804; Fax: +82-62-530-1759; E-mail: ykwon@jnu.ac.kr.

characteristics of pre-processed data, the classifier will be trained to discriminate the snoring state from the normal respiratory state. Researches ranging from laboratory-based studies to real environment studies have produced higher detection rates using several classification methods. These methods include robust linear regression with principal component analysis (PCA) on the sub-band energy distributions [1], AdaBoost classifier using acoustic features from time and spectral domains [3], hidden Markov model (HMM) [4], fuzzy mean clustering [5], wavelet-based spectral analysis [6], genetic algorithm, and a support vector machine (SVM) [7]. Most of these classification methods use information extracted from acoustic signals during sleep.

This paper introduces a new neural network approach for sleep snoring detection. The neural network is a multilayer feed-forward neural network with the first hidden layer of correlational filter operation [8, 9]. This correlational filter multilayer Perceptron neural network (f-MLP) is simultaneously trained for both the correlation filter coefficients and the weight parameters using the back-propagation algorithm. The correlational filter layer was implemented in the frequency domain, which included the transformation of the input patterns in the time domain into the frequency domain. The filter operation was applied to this and the power spectrum was computed as the output of this first layer. Then, the output of the filter layer was fed into the next higher layer of the classification network. The pattern data set for classification was collected from many over-night sleep breathing signals, and was pre-processed by normalization without any pre-process noise reduction. The classification performance was compared with the ordinary multilayer feed-forward neural network (o-MLP).

2. Literature reviews

2.1. Snoring detection

In the last 20 years, there has been much research about snoring detection. Jane and Sola-Soler provided a snoring detector containing 2 blocks: a segmentation subsystem and a 2 layer feed-forward multilayer neural network with a back-propagation learning algorithm [10]. Input data included snoring sounds from normal snorers, OSAS patients, and other sounds. This detector showed an acceptable performance and achieved a sensitivity of 82%. Even though they changed the dataset bias in [11], the accuracy rate was not significantly improved.

Later, in [12], Snider and Kain proposed a snoring classification method including a pre-processing unit, a spectral domain transformation, and a classifier with hidden Markov Models. Accuracy increased to 87% compared to previous approaches.

Moreover, in [1], M. Cavusoglu proposed another technique to detect snoring by using principle component analysis (PCA) as a feature extraction. The power spectrum of the data set was necessarily calculated to obtain the sub-band energy distribution before applying PCA. New features were then computed by projecting the feature vectors onto PCA's eigenvectors space. In this paper, a two dimensional pattern data obtained by PCA was used for classification with a robust linear regression method. The accuracy was found to be 97% when the system was trained using only simple snorers' data. It dropped to 90% when training data included both simple snorers and OSA patients.

Other classification methods were also investigated, such as an AdaBoost classifier using acoustic features from time and spectral domains [3], hidden Markov model (HMM) based method [4], and fuzzy mean clustering [5]. Their performances were acceptable. However, it is difficult to draw a fair

comparison between them because their data collection conditions are different. However, most of the methods used the characteristics of the signal in the frequency domain.

2.2. Correlational filter multilayer perceptron neural network (f-MLP)

Recently, we introduced a new type of multilayer feed-forward neural network that was able to successfully classify the patterns that has frequency-information-dominant characteristics hidden in time domain patterns, where the ordinary multilayer feed-forward neural networks normally fail [8, 9]. The neural network is composed of two operation stages: the first hidden layer of correlational filter operation and the ordinary feed-forward neural network operation. This correlational filter multilayer Perceptron neural network (f-MLP) is simultaneously trained for both the correlation filter coefficients and the weight parameters using the back-propagation algorithm.

The correlational filter layer is implemented in the frequency domain, which implies the transformation of the input patterns in the time domain into frequency, on which the filter operation applied and the power spectrum is computed as the output of this first layer. Note that the correlational operation in the time domain can be done by the point-wise multiplication in the frequency domain. Then, the power spectrum of the filter output is calculated and fed into the next higher layer of the network. Figure 1 depicts this neural network architecture. In this figure, $\mathbf{x}(t)$ is the input pattern, and $\mathbf{X}(f)$ and $\mathbf{H}(f)$ are the Fourier transform of $\mathbf{x}(t)$ and the correlation filter, respectively. Also, 'FT' and 'P/S' stand for the computing procedures for the Fourier transform and power spectrum, respectively. $\otimes$ represents the point-wise multiplication.

3. Data collection

3.1. Collection of sleep breathing sounds

Fifteen randomly chosen participants with snoring symptoms were asked to wear a neck band equipped with a small recording device with a condenser microphone during overnight sleep. The microphone was a unidirectional device faced the neck, which minimized noise disturbances from the environment. The recorded signals were then digitized with the sampling frequency of 8 kHz and a 16-bit resolution as mono channel raw data.

3.2. Pattern data set

From the long breathing sound signal recorded during overnight sleep, the patterns of snoring sounds and normal respiration were collected. By manual analysis on the recorded sound, it was found that snoring sounds occurred at the inhalation phase of the breathing cycle, while normal breathing sounds occurred at the exhalation phase or intermediate of those two phases (usually no sound).
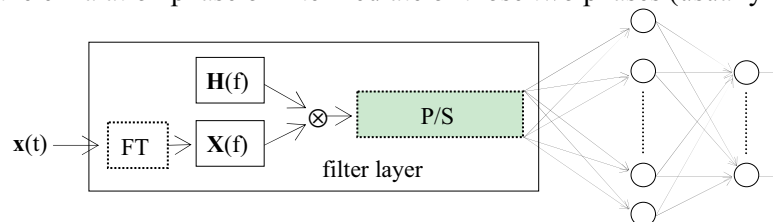

Fig. 1. f-MLP basic structure.

Firstly, the snoring zones were determined by auditory examination, and then the pattern vectors for snoring sounds were segmented out using a sliding window of 1024 points. For the next pattern, the sliding window moved 512 points. This sliding segmentation process was repeated over the central region of a snoring zone to have a gap between its neighbor non-snoring zones. This procedure was repeated for the other snoring zones of the breathing sound signal. Using the same sliding window segmentation procedure, the pattern vectors for normal breathing sounds were collected from the non-snoring zones, which include the exhalation phases and normal inhalation phases. The pattern vector segmentation was applied to the sleep breathing sound files obtained from the participants. Then, several thousand pattern vector sets were obtained. This pattern data set was then used for running the experiments by randomly choosing 1,000 patterns for each category. 70% of the data set was used for training the MLP classifiers, while the rest 30% unseen data was used for testing the trained classifier.

4. Experimental results and analysis

4.1. Experimental set-up

In the experimental studies, both the correlational filter multilayer neural network (f-MLP) and the ordinary multilayer neural network (o-MLP) were trained with the training pattern set and then tested with the test set. The f-MLP consisted of 2 layers of the filter layer and the output layer, which has only a single hidden layer of the correlational filter and the output layer. The filter layer had the structure of 1024 units with complex number architecture. The output layer had only 2 units, which represent the 2 classes of snoring and non-snoring. Other network configurations and training criterion were the same as in [9].

The o-MLP was also trained and tested with the same data set for comparison purposes. However, the size of the neural network (i.e., the number of parameters such as weights and biases) and the training conditions cannot be fairly compared with those of f-MLP. Thus, nearly optimal conditions for the network size and training process were exhaustively found, and the o-MLP network of a single hidden layer with 41 hidden nodes was chosen.

4.2. Classification results

To avoid overtraining, cross-validation with the test dataset was used as an additional stopping criterion with the predefined maximum number of training epochs and minimum sum-squared-error
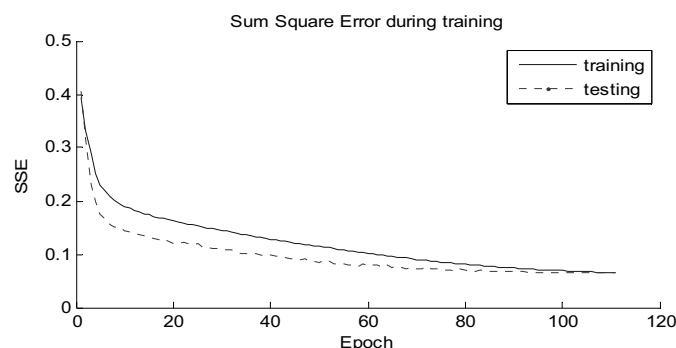

Fig. 2. Training and testing curve of f-MLP.

Table 1
Comparison of classification performance between f-MLP and o-MLP

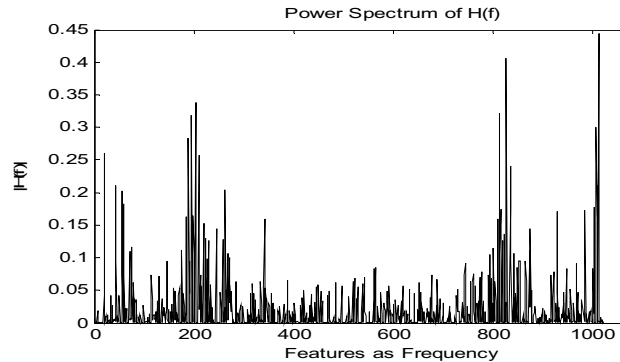| trials	1		2		3		4		5	
	o-MLP	f-MLP	o-MLP	f-MLP	o-MLP	f-MLP	o-MLP	f-MLP	o-MLP	f-MLP
training	93.9	96.8	97.0	96.9	93.0	97.6	98.4	97.0	96.6	97.0
testing	84.0	96.6	77.8	94.6	81.0	96.3	82.0	96.5	86.2	94.8
difference	9.9	0.2	19.2	2.3	12.0	1.3	16.4	0.5	10.4	2.2

Fig. 3. Double-side power spectrum of the filter layer in f-MLP.

(SSE). Thus, the training process terminated when the testing error did not decrease during certain training epochs. As shown in Figure 2, most of the training processes for f-MLP stopped between 90 and 110 epochs.

Table 1 shows the comparison results of the 5 repeated experimental studies for both f-MLP and o-MLP. Note again that for each trial, 1,000 training data sets for each category (snoring and non-snoring) were randomly retrieved from the pattern data pool of several thousands, as described in section 3.2. Then, 70% were used for training and 30% were used for testing.

The training accuracy for both MLP classifiers was over 93%, although f-MLP usually achieved higher accuracy. However, for the testing set, correction rates of f-MLP were always higher than that of o-MLP with the maximum difference of 16.8% and the minimum difference of 8.6%. Also, in terms of the generalization capability, which can be evaluated by the difference between training accuracy and testing accuracy, f-MLP is superior to o-MLP by showing the maximum difference of 2.3%. However, o-MLP mostly shows the difference higher than 10%, which indicates poor generalization capability.

4.3. Analysis

As shown by a high testing accuracy over 96%, the f-MLP successfully extracted and classified the frequency information from the breathing sound during sleep. Figure 3 shows the double-side power spectrum of coefficients in the filter layer; the horizontal axis shows the number of features extracted in the filter layer (not the real frequency index of the signal). As shown in Figure 3, at least 2 feature (frequency) ranges around 200 and 810 of the x-axis index have been activated in this layer to give more discriminant power for the next layer. Figure 4 statistically shows the distribution of the power spectrum of filter output for every pattern. The dots (.) and the stars (∗) in this figure represent the mean values for the classes, while the error bars indicate the variances. The ratio of the distance between two means to the sum of variances for each feature index can be considered as the

discrimination power of that feature: the higher the ratio, the more discrimination power. Thus, separation power of each feature can be defined by the following equation:

$$separation = \frac{|\mu_1 - \mu_2|}{|var_1| + |var_2|}$$ (1)

where μ_i indicates the mean value and var_i does the variances of class i ($i = 1, 2$), respectively. If *separation*>1, the two classes are highly separable with that feature. However, even though the class's distributions overlap, the neural network can perform its task by their nonlinear classification ability.

Figure 4 explains why some of the features are highly activated at the hidden layer output. Two feature index ranges of [20 80] and [120 220] show higher discrimination power. Also, by close examination, some other points such as 270 and 350 can be found as good features for classification. By converting these feature indices to real frequency values, the interesting frequency zones are [156 625] Hz, [937 1718] Hz, 2100 Hz, 2735 Hz. It can be also found that the snoring sounds mostly have higher power in the first range of [156~625] Hz and have lower power in other regions, especially in [937~1718] Hz.

5. Conclusion and future works

A new correlational filter multilayer Perceptron neural network (f-MLP) method for detecting snoring was introduced. Respiratory sounds were directly analyzed and classified based on their frequency characteristics. Through this research, f-MLP produced a superior performance to the conventional multilayer Perceptron neural network (o-MLP) by producing the testing accuracy of 97% or higher. It was also experimentally proven that the generalization capability of f-MLP was better with smaller differences between training accuracy and testing accuracy, which could not be reached by o-MLP.

In the future works, further research should be conducted on this snoring detection method. For this research, more sleep breathing sounds should be collected from various populations. Also, comparison
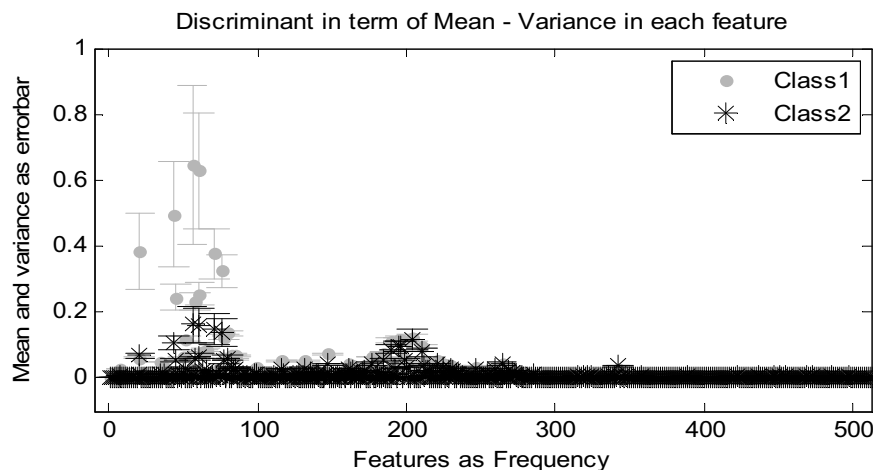

Fig. 4. Statistical distribution of a single-side power spectrum for each feature index.

with a standardized machine approved by an official organization such as the Food and Drug Administration (FDA) should be performed to promote real-world applications. For reducing the computational burden, the f-MLP classifier can be applied only for the frequency ranges with higher discrimination power for the classification.

Acknowledgment

References

[1] F. Dalmasso and R. Prota, Snoring: Analysis, measurement, clinical implications and applications, European Respiratory Journal **9** (1996), 146-159.

[2] M. Cavusoglu, M. Kamasak, O. Erogul, T. Ciloglu, Y. Serinagaoglu and T. Akcam, An efficient method for snore/nonsnore classification of sleep sounds, Physiological Measurement **28** (2007), 841-853.

[3] E. Dafna, A. Tarasiuk and Y. Zigel, Automatic detection of whole night snoring events using non-contact microphone, PLoS One **8** (2013), e84139.

[4] H.K. Lee, J. Lee, H. Kim, J.Y. Ha, and K.J. Lee, Snoring detection using a piezo snoring sensor based on hidden Markov models, Physiological Measurement **34** (2013), N41-49.

[5] R.J. Beeton, I. Wells, P. Ebden, H.B. Whittet and J. Clarke, Snore site discrimination using statistical moments of free field snoring sounds recorded during sleep nasendoscopy, Physiological Measurement **28** (2007), 1225-1236.

[6] A. Hossen, D. Jaju, B. Al-Ghunaimi, B. Al-Faqeer, T. Al-Yahyai, M.O. Hassan and M. Al-Abri, Classification of sleep apnea using wavelet-based spectral analysis of heart rate variability, Technology and Health Care **21** (2013), 291-303.

[7] Y. Maali and A. Al-Jumaily, Automated detecting and classifying of sleep apnea syndrome based on genetic-SVM, International Journal of Hybrid Intelligent Systems **9** (2012), 203-210.

[8] T.L. Nguyen et al., A multilayer neural network for classification of frequency information dominant patterns, Proceeding of the 15th International Conference on Advances in Neural Networks, Gdansk, Poland, 2014, pp. 62-68.

[9] T.L. Nguyen et al., Feature extracting correlation filter trained by perceptron, Proceedings of the 2014 International Conference on Systems, Control, Signal Processing and Informatics II (SCSI '14), Praha, Czech Republic, 2014, pp. 82-85.

[10] R. Jane, J. Sola-Soler, J.A. Fiz and J. Morera, Automatic detection of snoring signals: Validation with simple snorers and OSAS patients, Proceedings of the 22nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society **4** (2000), 3129-3131.

[11] R. Jane, J.A. Fiz, J. Sola-Soler, S. Blanch, P. Artis and J. Morera, Automatic snoring signal analysis in sleep studies, Proceedings of the 25th Annual International Conference of the IEEE Engineering in Medicine and Biology Society **1** (2003), 366-369.

[12] B.R. Snider and A. Kain, Automatic classification of breathing sounds during sleep, IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2013, pp. 699-703.