Medical Engineering & Physics 36 (2014) 477–483

Contents lists available at ScienceDirect

Medical Engineering & Physics

journal homepage: www.elsevier.com/locate/medengphy

Detection of systolic ejection click using time growing neural network

Arash Gharehbaghi [a,*], Thierry Dutoit [b], Per Ask [a], Leif Sörnmo [c]

[a] Physiological Measurements (IMT), Department of Biomedical Engineering, Linköping University, Linköping, Sweden
[b] Faculty of Polytechnic, Mons University, Mons, Belgium
[c] Department of Electrical and Information Technology, Center for Integrative Electrocardiology, Lund University, Lund, Sweden

ARTICLE INFO

Article history:
Received 27 March 2013
Received in revised form 6 January 2014
Accepted 8 February 2014

Keywords:
Systolic ejection click
Time growing neural network
Time delay neural network
Heart sound

ABSTRACT

In this paper, we present a novel neural network for classification of short-duration heart sounds: the time growing neural network (TGNN). The input to the network is the spectral power in adjacent frequency bands as computed in time windows of growing length. Children with heart systolic ejection click (SEC) and normal children are the two groups subjected to analysis. The performance of the TGNN is compared to that of a time delay neural network (TDNN) and a multi-layer perceptron (MLP), using training and test datasets of similar sizes with a total of 614 normal and abnormal cardiac cycles. From the test dataset, the classification rate/sensitivity is found to be 97.0%/98.1% for the TGNN, 85.1%/76.4% for the TDNN, and 92.7%/85.7% for the MLP. The results show that the TGNN performs better than do TDNN and MLP when frequency band power is used as classifier input. The performance of TGNN is also found to exhibit better immunity to noise.

1. Introduction

Heart sound auscultation is a simple but very useful diagnostic technique for screening of heart problems. Since the technique is inexpensive, it plays an important role in home care and primary health centers. However, powerful algorithms are needed for detection and classification of heart sounds which explore temporal as well as spectral characterization. For the normal heart, the phonocardiogram is composed of two basic sounds in each cardiac cycle, namely, the first and the second heart sound, S1 and S2, respectively, which reflect different mechanical activities of the heart [1]. Abnormal heart sounds have either physiological origin, e.g., innocent murmurs, or pathological origin [2]. The systolic ejection click (SEC) is a representative of the latter type of origin and occurs after S1. It is observed in patients with dilated aorta or pulmonary artery, aortic stenosis (AS), pulmonary stenosis, bicuspid or flexible stenotic aortic, or pulmonary valve [3]. Fig. 1 shows a complete cardiac cycle from a normal child and a child with bicuspid aortic valve (BAV) disease with SEC.

Even for expert physicians, the detection of SEC by auscultation is a complicated task in children, especially since the heart rate is typically faster than in adults and that SEC can be confused with other physiological heart sounds [1]. At the same time, it is of vital importance to screen pediatric heart lesions before the age of 12 so that disease progress can be controlled by appropriate medical advice. Considerable research has been directed toward the screening of pathological heart sounds both in human and animal cases, see, e.g., [4–7], but not so to SEC detection in children. It is therefore essential to develop detection and classification methods which are particularly suited for this important group of patients.

Artificial neural networks (ANNs) are widely employed for pattern recognition and machine learning, being inspired by the human cognitive system [8]. Such networks play an important role in diverse fields such as biomedical engineering, control engineering, and electronic systems [9,10]. The nonlinear classification of a multi-layer perceptron (MLP) neural network makes it suitable for applications where learning is accomplished with less training data than for statistical classifiers. However, temporal properties of the input signal are not preserved by an MLP in contrast to hidden Markov model (HMM) and dynamic time warping. The time delayed neural network (TDNN) was first described as a transition invariant back-propagation network, and found to perform better than HMM on noisy speech recognition tasks [11]. With the TDNN, classification is based on successive, fixed length time windows from which the set of feature vectors are extracted. In certain applications, however, classification of short signals based on such an approach may not offer the needed time and frequency resolution and, as a result, the performance of the classifier is reduced.

This paper proposes the time growing neural network (TGNN) for classification of heart sounds, employing growing time windows in conjunction with an ANN classifier. The TGNN puts more

* Corresponding author. Tel.: +46 13 286754; fax: +46 13 101902.
E-mail addresses: arash.ghareh.baghi@liu.se, arash_gharahbaghi@yahoo.com (A. Gharehbaghi).

http://dx.doi.org/10.1016/j.medengphy.2014.02.011
1350-4533/

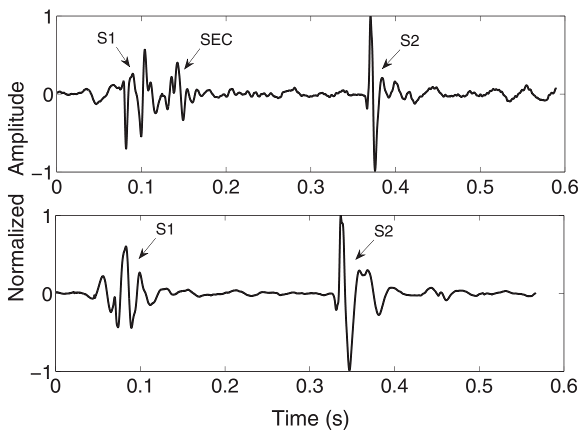

Fig. 1. A complete cardiac cycle of a normal child (bottom) and a patient with SEC (top).

emphasis on the first window as its contents are present in all other windows. This feature makes it suitable for SEC detection in which an abrupt change in frequency contents is heard as a click sound. The input signal is considered to be short-length as it comprises only the first heart sound and, in abnormal cases, also the SEC. The results obtained with the TGNN are compared to those of TDNN and MLP.

2. Datasets

2.1. Real data

Heart sound recordings of 10-s duration were acquired from 40 children, ranging in age from 3 to 9 years, being referred to the Children's Medical Center of Tehran, Iran. Medical trials were accomplished according to the guidelines of Tehran University of Medical Science which is in compliance with the declaration of Helsinki. Informed consent as well as assent (for children under 15) for the collection and use of data was given by the individuals or their guardians prior to the data recording. Twenty of these children defined the training dataset, whereas the remaining 20 children defined the test dataset, see Table 1. All recordings were segmented under the supervision of pediatric cardiologists who used echocardiography and complementary tests as gold standard. The S1 segment served as the input signal for classification, including SEC for abnormal recordings. The training and test datasets were approximately of equal sizes as they consisted of 285 and 329 cardiac cycles, respectively, i.e., a total of 614 cycles.

For data acquisition, a Welch Allyn Meditron stethoscope was used together with a laptop equipped with a 16-bit sound card using a sampling rate of 44.1 kHz. A high sampling rate was used to avoid signal distortion during replay. However, the signals were later down-sampled to 2 kHz for classification.

Table 1
Training and test datasets.

Description	Training	Test
Number of recordings from normal	10	10
Number of recordings from patients with SEC	10	10
Total number of cardiac cycles with SEC	133	168
Total number of cardiac cycles without SEC	152	161
Number of abnormal recordings with AS	6	6
Number of abnormal recordings with BAV	6	8
Minimum age (years)	3	3
Maximum age (years)	8	9

2.2. Noise study

Noise immunity of the three classifiers was studied by adding either colored Gaussian noise, physiological or external noise to the heart sound recordings. Colored noise was produced by filtering white Gaussian noise with an FIR low-pass filter with cutoff frequency at either 1000 Hz (type #1) or 250 Hz (type #2). Type #1 noise had the same bandwidth as the heart sound signals, whereas type #2 noise covered the spectral band over which the spectral features of the present method were computed.

Physiological noise was obtained by extracting the diastolic part from the heart sounds recordings of normal subjects having a low level of background noise (assessed visually and auditively). The extraction resulted in a total of 52 cycles from 7 subjects. A randomly selected noise cycle was then added to a heart sound signal, ensuring that the respective lengths were matched and that the noise did not originate from the target heart sound recording.

External noise was recorded by the above-mentioned equipment during "background conversation", but without the inclusion of heart sounds. Segments at random locations with suitable length were then extracted from the external noise recording and added to the heart sound signal.

3. Methods

3.1. TGNN classifier

The TGNN classifier is based on windowing of the observed signal $x(n)$, using time windows with a fixed starting point but with growing lengths. The shortest window contains N_0 samples, and then the length grows in multiples of G samples until the entire signal, of length N, is included. N_0 is ideally selected such that the S1 segment is incorporated (typically 100 ms). However, the contents of an incomplete S1 are included in the subsequent growing windows, acting as a compensatory factor for small values of N_0. The total number of analyzed windows is equal to K, chosen such that N is not exceeded. Each window is characterized by its spectral content using the periodogram,

$$X(\omega, k) = \frac{1}{N_0 + kG} \left| \sum_{n=0}^{N_0+kG-1} x(n)w(n)e^{-j\omega n} \right|^2, \tag{1}$$

for $k = 0, \ldots, K-1$. The weighting function $w(n)$ is introduced for the purpose of reducing side lobe leakage [12]. For each window $[0, N_0 + kG - 1]$, the spectral power in adjacent frequency bands, all having bandwidth $\Delta\omega$, are computed and employed as features for classification,i.e.,

$$P_{kj} = \int_{\omega_0+(j-1)\Delta\omega}^{\omega_0+j\Delta\omega} X(\omega, t)\, d\omega, \tag{2}$$

where $j = 1, \ldots, J$ and ω_0 denotes the lowest frequency for analysis.

The resulting set of features $\{P_{kj}(N_0,G)\}$ constitutes the input layer of the TGNN and depends on the design parameters N_0 and G. In practice, the critical design parameter is K since the growth factor G is linked to K by $G = L/K$ and N_0 is assumed to be an integer multiple of G:

$$N_0 = N_R \cdot \frac{L}{K}, \tag{3}$$

where L denotes the signal length and N_R (<K) the relative length of the shortest window. Thus, the design parameters become K and N_R, where K is the total number of temporal divisions of the input signal, and N_R is the number of temporal divisions assumed as the initial window length.

The neural network itself is a nonlinear function $f(P_{kj}(N_R,K);S)$ that is defined by the activation functions of the hidden and output