The attractor recurrent neural network based on fuzzy functions: An effective model for the classification of lung abnormalities

Mohammad Bagher Khodabakhshi, Mohammad Hassan Moradi

Biomedical Engineering Department, Amirkabir University of Technology, Tehran 15875-4413, Iran

ARTICLE INFO

Keywords:
- Lung sounds
- Attractor recurrent neural networks
- Fuzzy functions
- Recurrence quantification analysis

ABSTRACT

The respiratory system dynamic is of high significance when it comes to the detection of lung abnormalities, which highlights the importance of presenting a reliable model for it. In this paper, we introduce a novel dynamic modelling method for the characterization of the lung sounds (LS), based on the attractor recurrent neural network (ARNN). The ARNN structure allows the development of an effective LS model. Additionally, it has the capability to reproduce the distinctive features of the lung sounds using its formed attractors. Furthermore, a novel ARNN topology based on fuzzy functions (FFs-ARNN) is developed. Given the utility of the recurrent quantification analysis (RQA) as a tool to assess the nature of complex systems, it was used to evaluate the performance of both the ARNN and the FFs-ARNN models. The experimental results demonstrate the effectiveness of the proposed approaches for multichannel LS analysis. In particular, a classification accuracy of 91% was achieved using FFs-ARNN with sequences of RQA features.

1. Introduction

Respiratory signal analysis is an efficient tool for the detection of pulmonary abnormalities and the assessment of lung disorders. As a common physical examination method, auscultation from different positions of the chest may be performed to listen to the lung sounds noninvasively. This method, which is performed by physicians with a stethoscope, has some limitations, such as the limited frequency response of the stethoscope, and the limited sensitivity of the human ear to lung sounds of lower frequencies [1]. However, electronic stethoscopes and digital microphones help overcome these limitations, and the lung sounds can be recorded for further signal-processing steps.

Based on the acoustic properties of the lung sounds, they can be categorized in one of the following groups: normal and abnormal (or adventitious). According to the definition by the American Thoracic Society (ATS), sounds are considered continuous if their duration is longer than 250 ms; otherwise, they are considered discontinuous [2]. Continuous adventitious sounds are further subdivided into wheezes and rhonchi. Additionally, discontinuous adventitious breath sounds are categorized into fine and coarse crackles. Each of the aforementioned lung sounds is associated with different types of lung disorders. The dominant frequency of the wheezes is above 400 Hz and they are heard in the patients suffering from asthma. The dominant frequency is 200 Hz or less for rhonchi, which are typically associated with chronic obstructive pulmonary disease (COPD) [2,3].

Computer analysis and data processing of recorded lung sounds are important tools for assessing lung sounds parametrization. Numerous investigations have been undertaken in the field of digital respiratory signal analysis. The methods based on the auto-regressive (AR) and multivariate auto-regressive (MAR) models [4–7], the Mel-frequency cepstral coefficient (MFCC) [8–10] approach, the methods based on the extraction of spectral features [1,4,11–13], and those relying on wavelet transform [10,14–17] are commonly applied for feature extraction. Another set of features used for the recognition of the lung abnormalities involving time-frequency representation (TFR) has been employed [13,18–20]. From the classifier point of view, artificial neural networks (ANNs) and K-nearest neighbour (KNN) are widely used for LS classification [3,6,15,21,22]. In the same way, the Gaussian mixture model (GMM) [10,16] and support vector machine (SVM) [1,23–25] as alternative classification methods were used in some other studies. A systematic review of the use of machine learning and artificial intelligence (AI) methods in LS analysis, including various types of features and classifiers, can be found in [3].

However, the effect of non-linear and recurrence quantification analysis (RQA) features for the dynamic modelling of respiratory sounds has rarely been considered [26]. These features, initially introduced by Marwan [27], characterize a given time series in its reconstructed state space (RSS). Although AI tools have often been utilized in LS classification, there are few studies dealing with the
application of AI to construct LS representative models [14,26]. In this work, the advantages of RQA features and AI methods are combined to achieve an effective model allowing the extraction of distinctive features of LS.

In recent studies, the application of wavelet neural networks (WNNs) [14] and fuzzy systems [26,28,29] for temporal behaviour modelling have been presented, and attractor recurrent neural networks (ARNNs) were used as auto-associative memory models [30,31]. A couple of studies served to prove that the network parameters could provide compact representation and distinctive information from the respiratory signals [14,26]. Motivated by the success of ARNNs in speech signal processing [30], in our work, we model the temporal behaviour of the recorded lung sounds using an ARNN and explore the distinctive features from the network parameters. Furthermore, given the capability of fuzzy approaches in lung sounds modelling [26,28], we extended ARNN in the fuzzy functions structure [28,32,33] (FFs-ARNN).

In our study, we consider three types of signals obtained from healthy subjects and subjects suffering from asthma and COPD. The respiratory signals obtained from each subject are modelled using an individual ARNN and an FFs-ARNN, so that the attractors of each network are formed. Then, the hidden-layer parameters of the ARNN, which correspond to the formed attractors, are utilized in the classification task. Data points representing a single inhale-exhale respiration cycle for each subject are fed in the ARNN as well as the FFs-ARNN for temporal modelling. Furthermore, the temporal sequences of their RQA features are considered as an alternative set of input data.

The rest of this paper is organized as follows: In Section 2, the detailed structure of the ARNN and its learning procedure are described. A brief review of basic FFs system and its structure are presented in Section 3. Our methodology, including the data acquisition protocol and the procedure followed to construct the feature vectors for the purpose of classification, is described in Section 4. Sections 5 and 6 are devoted to the results and discussion parts, respectively. Finally, the relevant conclusions are presented in Section 7.

2. Attractor recurrent neural network

The formation of continuous attractor dynamics in neural networks was introduced earlier, in the context of speech signal for nonlinear de-noising and classification [30,31,34]. The proposed model in this paper, which is used to model the respiratory signals, is inspired by the ARNN idea and the application of attractors for distinguishing three types of subjects. In the training procedure of the network, for each input sample, an individual kernel function is created to discriminate the sample with soft boundaries, so the network has numbers of attractors and basins of attraction corresponding to each sample. Therefore, the network learns input patterns as attractors and it has the capability of converging inputs located in each basin of attraction to the corresponding attractors by cycling in the net [30]. The dynamics of the attractors in ARNNs appear in the network parameters and we aim to illustrate their capability to classify lung abnormalities.

2.1. The structure of ARNN

ARNN is equipped with two types of full connections in its structure, one of which is the conventional link employed in feed-forward neural networks, namely the make-forward path, and the other one is the recurrent connection connecting the hidden-layer nodes to each other in order to form the attractors. Suppose \((x_i, z_i)\) represents an input-output pair, then:

\[
x_i = (x_{i1}, x_{i2}, \ldots, x_{in}), \quad z_i = (z_{i1}, z_{i2}, \ldots, z_{io}), \quad i = 1, 2, \ldots, N
\]

where \(N\), and \(N\) are the numbers of input pattern and output pattern dimensions, and the number of samples respectively. The structure of the network consists of input, intermediate (hidden), and output layers.

In order to form the attractors in the network, the reconstructed input in the hidden-layer neurons are calculated as follows:

\[
\tilde{x}_i = (1 - \gamma) x_i + \gamma y_i, \quad \gamma = \left( \begin{array}{c} \gamma_1, \gamma_2, \ldots, \gamma_m\end{array} \right), \quad y_i = (y_{i1}, y_{i2}, \ldots, y_{im}), \quad i = 1, 2, \ldots, N
\]  \(\text{(2)}\)

where, \(\tilde{x}_i\), \(y_i\) are the reconstructed input and the hidden-layer output corresponding to the \(i\)th input pattern, \(m\) is the number of nodes in the hidden layer, \(\gamma\) is the reconstruction coefficient, which is set to 0.7 according to [30,34]. As shown in Eq. (2), \(\tilde{x}_i\) is constructed by 70 per cent of the outputs of the recurrent connections and 30 per cent of input patterns \(x_i\), implying that in each epoch of training \(\tilde{x}_i\) is reconstructed based on its corresponding input pattern in the previous epoch. By feeding the reconstructed input to the network, the reconstructed output of the hidden layer for the \(i\)th input is computed as:

\[
\tilde{y}_i = f(\tilde{x}_i, y_i), \quad \tilde{y}_i = (\tilde{y}_i, \tilde{y}_2, \ldots, \tilde{y}_o)
\]  \(\text{(3)}\)

where \(f = \frac{1-e^{-x}}{1+e^{-x}}\) is a non-linear bipolar sigmoid function. Substituting Eq. (2) in Eq. (3), the reconstructed output of the hidden layer in each iteration is given by:

\[
\tilde{y}_i = f(1 - \gamma) x_i + \gamma y_i + \gamma y_i + y_i
\]  \(\text{(4)}\)

Here, \(y_{i1}\), is the weight matrix connecting the input layer to the hidden layer. After transferring the product of \(y_i\), \(y_i\) to \(y_i\), (Eq. (4)), the self-recurrent structure of the network is formed and the reconstructed output of the hidden layer in each epoch will be trained using its values from the previous epochs.

The detailed structure of the proposed model is shown in Fig. 1. The structure of the hidden layer is split into two sub-layers, namely the primary hidden layer (PHL) and the reconstructed hidden layer (RHL). In addition, there is another distinct output layer, which is responsible for forming the attractors. At the beginning of the network, the training values of the hidden layer nodes are set to zero and their values are constructed based on the input values (Fig. 1-a). In the subsequent epochs, the connections from the inputs and outputs are added to generate the reconstructed output of the hidden layer to form attractors (Fig. 1-b). By adopting the input patterns in the network, the movement of the pattern towards the accurate attractors will be guaranteed [30,31].

According to the aforementioned structure, the input-output relations can be written as:

\[
x_{io} = f(x_i, x_i), y_{io} = (y_{i1}, y_{i2}, \ldots, y_{io}), \quad i = 1, 2, \ldots, N
\]  \(\text{(5)}\)

\[
y_i = f(1 - y_i, x_i, x_i, y_i + y_i, y_i, y_i)
\]  \(\text{(6)}\)

\[
z_i = f(y_i, y_i)
\]  \(\text{(7)}\)

where \(y_{i1}, y_i,\) and \(z_i\) are the PHL output, the RHL output, and the output of the network, respectively. In each iteration, the second output of the network and the updated value of PHL are calculated as follows:

\[
x_i = (1 - \gamma) x_i + \gamma y_i + y_i + y_i, \quad x_i = (y_{i1}, y_{i2}, \ldots, y_{io}), \quad i = 1, 2, \ldots, N
\]  \(\text{(8)}\)

\[
x_{io} = y_i
\]  \(\text{(9)}\)

where \(x_i\) is the second output of the network that is utilized to form the attractors. In addition to \(y_i\), there is another weight matrix \(\gamma_{io}\), that connects the hidden-layer nodes to the output nodes, thereby providing forward weights of the network. The purpose of our study is to utilize the forward path parameters \(\gamma_{io}\) of the ARNN for non-linear mapping of the network inputs to its output in order to construct the appropriate models for them. In addition, the recurrent parameter of ARNN, \(\gamma_i\) will be used to form the attractors and classify the recorded pulmonary abnormalities.
2.2. ARNN learning

The selection criteria for the two objective functions employed to train the proposed network are as follows: The system should map the input patterns to the output with minimum matching error and the network should create the appropriate attractors to identify the discriminating features. These objective functions, which are the sum of the squared error related to the forward path of the network and the sum of the squared error associated with reconstruction of the input pattern in the hidden layer, are given by:

\[
E_1 = \frac{1}{2} \sum_{i=1}^{m} (d_i - z_i)^2
\]

\[
E_2 = \frac{1}{2} \sum_{j=1}^{m} (\eta_j - \gamma_j)^2
\]

where \(d\) is the desired output. As Eq. (11) depicts, in order to form each training sample as an individual attractor, it is necessary to minimize the error between the reconstructed input and the output of PHL in the first epoch. In other words, based on this error function, each data point gradually converges to the first value of the PHL, providing the desired value of the network for creating attractors.

In the present study, the back-propagation (BP) algorithm has been applied to learn parameters, as illustrated in Fig. 2. This figure depicts \(w_F\) and \(v_F\), which are updated based on \(E_1\) and \(E_2\), respectively, while \(\eta_1\) uses both of the error functions.

Fig. 1. The structure of the ARNN. The hidden layer of the network is split into PHL and RHL. a) In the first epoch of training, the values of the hidden-layer nodes are set to zero. b) The reconstructed output of the network is generated to form the attractors.

Fig. 2. ARNN back-propagation algorithm. \(w_F\) and \(v_F\) are adjusted based on \(E_1\) and \(E_2\), respectively, and \(\eta_1\) uses both of the error functions.
where $\eta$ and $\alpha$ are the learning and momentum rates, respectively. The partial derivatives in the above equations are determined as:

$$\frac{\partial E_1}{\partial \eta} = y. (1 - z^2). (d - z)$$  \hspace{1cm} (15)

$$\frac{\partial E_2}{\partial \eta} = y. (x. \eta - y)$$  \hspace{1cm} (16)

3. Fuzzy functions

The fuzzy-functions (FFs) system, which was introduced by Turksen [32], represents an alternative version of the fuzzy rule-based (FRB) systems. Due to the multidimensional structure of FFs, combination operators such as T-norm and S-norm have been eliminated and the antecedent parts of FFs rules have simpler structures compared with the conventional FRB systems. Thus, the membership values of input vectors are obtainable based on any fuzzy clustering algorithms such as FCM or its variations. Moreover, the achieved membership values alongside the input patterns allow the generation of augmented input patterns, which can be employed to build FFs in the subsequent of the rules. The FFs can be estimated by different regression methods such as the least squared estimation (LSE) [32], support vector machine (SVM) [35], and multi-adaptive regression spline (MARS) [28]. One of the main purposes of this study is to apply ARNNs as the regression function in the structure of FFs, thereby allowing the introduction of the FFs-ARNN model for lung sounds.

**Fig. 3.** a) The structure of the proposed FFs-ARNN. b) The detailed structure of fuzzy regression functions.

**Fig. 4.** Sensor array attached on the posterior thoracic surface of the subjects.
transformations of the membership values to their corresponding primary vectors, augmented input patterns are created as:

$$
\phi_u = \left[ x_i, u_i, T(u_i) \right]
$$

(19)

where $u_i$ is the membership degree of the $i$th input data in the $h$th cluster and $T$ is the transformation function applied to the membership values. Detailed information regarding the selection of appropriate transformation functions is available elsewhere [32]. To estimate FFs for each cluster, an ARNN model is applied, which uses $\phi_i = 1, 2, \ldots, N$. Finally, the overall output of the system is calculated by the weighted average of each cluster output as follows:

$$
F(x_i) = \frac{\sum_{h=1}^{H} u_h(x_i) \cdot FF_h(\phi_h)}{\sum_{h=1}^{H} u_h(x_i)}
$$

(20)

where $F$ is the output of the network and $FF_h$ is the fuzzy function of the $h$th cluster.

4. Methodology

4.1. Subjects and data acquisition

Respiratory sounds were recorded from a total of 83 subjects at the Department of Pneumology, Shariati Hospital, Tehran. The researchers, however, are from the bioinstrumentation and biological signal processing lab of the Amirkabir University of Technology. The resulting database [26,36] contains LS from 27 patients suffering from COPD, 31 asthma patients which members of both groups aged within 25–55 years, and 25 healthy individuals of 20–40 years of age who were non-smokers and had no history of serious pulmonary disorders. Before
starting the data acquisition process, all subjects had provided the recording personnel with informed consent. The data was collected from the subjects who were made to wear a 3×2 sensor array, in the seated position, in a regular room. The microphones were attached to the posterior side of the thorax and were aligned symmetrically with respect to the spine, as shown in Fig. 4.

The subjects were asked to stay relaxed and were instructed to breathe normally. Each recording contained three breathing cycles with a one-second pause after each cycle. The sounds were recorded using an AD Instrument data acquisition block with a sampling rate of 10 kHz and an analogue RC filter circuit with a cut-off frequency equal to 50 Hz was used for the driving microphones. Since the main frequency components of LS are between 100 Hz and 1 kHz, the data was digitally filtered by a band-pass FIR filter. Fig. 5 depicts the effect of filtering the raw signal (first row) wherein heart sound (HS) appears in the lower frequencies of the spectrogram, while it is dramatically decreased in the filtered signal (second row).

Fig. 6(a) provides an example of two successive respiratory cycles
from a healthy subject acquired at ch5. As can be seen, the amplitude of LS in the inspiratory phase is higher in comparison to the expiratory phase. On the other hand, Fig. 6(b) shows the LS from a patient on the same channel where spikes occur in the inspiratory phase.

### Table 1
The classification results of the different approaches for each channel. The samples are categorized in three distinct groups (healthy, asthma, and COPD).

<table>
<thead>
<tr>
<th>FVs extraction technique</th>
<th>Phase</th>
<th>Ch1 (%)</th>
<th>Ch2 (%)</th>
<th>Ch3 (%)</th>
<th>Ch4 (%)</th>
<th>Ch5 (%)</th>
<th>Ch6 (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AR coefficients [4,6]</td>
<td>Training</td>
<td>89.67</td>
<td>86.36</td>
<td>82.58</td>
<td>71.67</td>
<td>86.67</td>
<td>79.51</td>
</tr>
<tr>
<td></td>
<td>Testing</td>
<td>85.50</td>
<td>82.76</td>
<td>73.62</td>
<td>54.17</td>
<td>84.41</td>
<td>71.27</td>
</tr>
<tr>
<td>Wavelet transform [10,15]</td>
<td>Training</td>
<td>88.50</td>
<td>85.00</td>
<td>86.33</td>
<td>74.02</td>
<td>86.26</td>
<td>78.62</td>
</tr>
<tr>
<td></td>
<td>Testing</td>
<td>83.33</td>
<td>75.00</td>
<td>70.83</td>
<td>58.33</td>
<td>79.17</td>
<td>62.50</td>
</tr>
<tr>
<td>MFCC [8–10]</td>
<td>Training</td>
<td>90.42</td>
<td>91.04</td>
<td>84.33</td>
<td>70.53</td>
<td>83.26</td>
<td>73.20</td>
</tr>
<tr>
<td></td>
<td>Testing</td>
<td>84.26</td>
<td>80.35</td>
<td>75.00</td>
<td>55.25</td>
<td>78.42</td>
<td>60.91</td>
</tr>
<tr>
<td>Temp-ARNN</td>
<td>Training</td>
<td>51.88</td>
<td>53.89</td>
<td>63.42</td>
<td>53.55</td>
<td>45.11</td>
<td>55.03</td>
</tr>
<tr>
<td></td>
<td>Testing</td>
<td>40.93</td>
<td>46.00</td>
<td>37.96</td>
<td>38.30</td>
<td>37.30</td>
<td>41.82</td>
</tr>
<tr>
<td>Temp-FFsARNN</td>
<td>Training</td>
<td>50.11</td>
<td>61.03</td>
<td>70.24</td>
<td>61.22</td>
<td>72.34</td>
<td>68.21</td>
</tr>
<tr>
<td></td>
<td>Testing</td>
<td>47.95</td>
<td>56.47</td>
<td>65.13</td>
<td>58.08</td>
<td>67.12</td>
<td>66.29</td>
</tr>
<tr>
<td>RQA [27]</td>
<td>Training</td>
<td>62.70</td>
<td>67.72</td>
<td>53.68</td>
<td>50.08</td>
<td>70.71</td>
<td>51.12</td>
</tr>
<tr>
<td></td>
<td>Testing</td>
<td>61.37</td>
<td>65.31</td>
<td>49.81</td>
<td>43.73</td>
<td>64.98</td>
<td>47.78</td>
</tr>
<tr>
<td>RQA-ARNN</td>
<td>Training</td>
<td>80.39</td>
<td>81.48</td>
<td>70.35</td>
<td>75.72</td>
<td>81.09</td>
<td>75.74</td>
</tr>
<tr>
<td></td>
<td>Testing</td>
<td>77.83</td>
<td>79.19</td>
<td>69.22</td>
<td>75.24</td>
<td>77.92</td>
<td>75.40</td>
</tr>
<tr>
<td>RQA-FFsARNN</td>
<td>Training</td>
<td>90.52</td>
<td>92.14</td>
<td>72.02</td>
<td>69.49</td>
<td>93.49</td>
<td>72.42</td>
</tr>
<tr>
<td></td>
<td>Testing</td>
<td>88.87</td>
<td>91.75</td>
<td>71.73</td>
<td>69.17</td>
<td>91.68</td>
<td>71.78</td>
</tr>
</tbody>
</table>

### Table 2
Sensitivity and specificity values in the best three channels.

<table>
<thead>
<tr>
<th>FVs extraction technique</th>
<th>Phase</th>
<th>Ch1</th>
<th>Ch2</th>
<th>Ch5</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sensitivity (%)</td>
<td>Specificity (%)</td>
<td>Sensitivity (%)</td>
<td>Specificity (%)</td>
</tr>
<tr>
<td>AR coefficients [4,6]</td>
<td>Training</td>
<td>100</td>
<td>88.43</td>
<td>99.09</td>
</tr>
<tr>
<td></td>
<td>Testing</td>
<td>95.36</td>
<td>85.11</td>
<td>93.25</td>
</tr>
<tr>
<td>Wavelet transform [10,15]</td>
<td>Training</td>
<td>92.02</td>
<td>81.43</td>
<td>89.37</td>
</tr>
<tr>
<td></td>
<td>Testing</td>
<td>90.41</td>
<td>81.86</td>
<td>83.60</td>
</tr>
<tr>
<td>MFCC [8–10]</td>
<td>Training</td>
<td>97.32</td>
<td>88.57</td>
<td>87.85</td>
</tr>
<tr>
<td></td>
<td>Testing</td>
<td>93.53</td>
<td>85.38</td>
<td>80.19</td>
</tr>
<tr>
<td>Temp-ARNN</td>
<td>Training</td>
<td>96.99</td>
<td>27.92</td>
<td>97.01</td>
</tr>
<tr>
<td></td>
<td>Testing</td>
<td>94.62</td>
<td>16.07</td>
<td>85.56</td>
</tr>
<tr>
<td>Temp-FFsARNN</td>
<td>Training</td>
<td>93.21</td>
<td>37.53</td>
<td>95.81</td>
</tr>
<tr>
<td></td>
<td>Testing</td>
<td>91.12</td>
<td>32.70</td>
<td>89.57</td>
</tr>
<tr>
<td>RQA [27]</td>
<td>Training</td>
<td>100</td>
<td>85.96</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>Testing</td>
<td>97.78</td>
<td>82.01</td>
<td>95.52</td>
</tr>
<tr>
<td>RQA-ARNN</td>
<td>Training</td>
<td>100</td>
<td>93.42</td>
<td>95.00</td>
</tr>
<tr>
<td></td>
<td>Testing</td>
<td>100</td>
<td>83.33</td>
<td>95.00</td>
</tr>
<tr>
<td>RQA-FFsARNN</td>
<td>Training</td>
<td>100</td>
<td>90.28</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>Testing</td>
<td>100</td>
<td>82.14</td>
<td>100</td>
</tr>
</tbody>
</table>

### Table 3
Comparison of statistical measures between different methods.

<table>
<thead>
<tr>
<th>FVs extraction technique</th>
<th>AUC</th>
<th>F1 Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>AR coefficients [4,6]</td>
<td>0.95</td>
<td>0.9397</td>
</tr>
<tr>
<td>Wavelet transform [10,15]</td>
<td>0.90</td>
<td>0.8387</td>
</tr>
<tr>
<td>MFCC [8–10]</td>
<td>0.94</td>
<td>0.801</td>
</tr>
<tr>
<td>RQA-ARNN</td>
<td>0.964</td>
<td>0.9262</td>
</tr>
<tr>
<td>RQA-FFsARNN</td>
<td>0.9778</td>
<td>0.9421</td>
</tr>
</tbody>
</table>

### Table 4
Execution time of some different approaches.

<table>
<thead>
<tr>
<th>FVs extraction technique</th>
<th>Time (sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AR coefficients [4,6]</td>
<td>38.224</td>
</tr>
<tr>
<td>Wavelet transform [10,15]</td>
<td>29.40</td>
</tr>
<tr>
<td>MFCC [8–10]</td>
<td>33.58</td>
</tr>
<tr>
<td>RQA [27]</td>
<td>26.35</td>
</tr>
<tr>
<td>RQA-FFsARNN</td>
<td>46.74</td>
</tr>
</tbody>
</table>
4.2. Time-series sequences

In the pre-processing step, initially, a single respiratory cycle including an inhalation and an exhalation for each subject was separated by listening to the recorded lung sounds, so that the temporal modelling could be carried out. To this end, based on Taken’s theorem, the RSS of the respiratory time series should be generated by its estimated embedding dimension, \( m_d \) and delay \( \tau \). Therefore, points of the time series are embedded in another space with dimension \( m_d \) as:

\[
s(t) = [x(t), x(t + \tau), \ldots, x(t + (m_d - 1)\tau)]
\]  

(21)

where \( m_d \), \( d \), and \( \tau \) are the \( i \)-th state vector, the embedding dimension, and the time delay of the RSS, respectively. In order to model the time series through our proposed method, the last value of each state vector was considered as the desired value with the remaining regarded as input vectors.

4.3. Sequences of features (RQA)

RQA features are useful for revealing the non-linear behaviour of the lung sounds. In order to reproduce the sequences of the RQA feature vectors (FVs), it is necessary to split the signals into a finite number of windows and calculate the features in each window. The length of the window was selected to be 500 ms with a 20 ms shift between the neighbouring windows [26]. Based on the time duration of the adventitious sounds, we select a proper value for window size and shift value to be assured so that each window can contain a complete event.

The RQA features are quantified based on the recurrence plots (RPs) of the signals. Several measures based on the recurrence point densities at different lines of RPs were introduced to obtain the RQA. Motivated by the results reported in [26], we have used the following five RQA features to produce sequences of FVs:

- a. Length of longest horizontal (vertical) line: shows the maximum laminar phase in the system
- b. (Intermittency).
- c. Recurrence period density entropy: calculates the normalized entropy of the recurrence time distribution of time series.
- d. Transitivity: determines the probability of going from anywhere to any other point in the state space for the trajectory of the system.
- e. Determinism: introduces a measure for determinism or predictability of the system.

Detailed information about the definition and calculation of the above features is available elsewhere [27]. All of the six aforementioned features were calculated for each window of the LS signal and their temporal variations were modelled through the proposed method. The first four features were considered as inputs in order to model the fifth one.
4.4. The classification scheme

The FVs extracted from the models which served as the parameters corresponding to the network attractors are fed to a classifier. As a common classification technique, a three-layer MLP network with a back-propagation algorithm was implemented. The number of hidden-layer nodes was considered to be five, based on the best performance of the network according to the mean squared error (MSE) criterion. Obviously, the number of input and output layer nodes was defined by the dimension of the FVs and the number of classes, respectively. In ARNN modelling, the number of PHL nodes, \( m \), defines the length of the vector for each subject. Moreover, in the FFs models, this value is dependent on the number of clusters, \( c \), as well. In this work, for the FFs modelling, the values of \( j \) and \( c \) were set to 3 and 2 respectively. The FV set was divided into the training and the testing phases. In order to achieve an insight into how the classifiers will generalize to an independent dataset, a multiple K-fold cross-validation approach with different K values was applied. The extracted feature set contains three distinct categories, namely the healthy individuals, the subjects suffering from COPD, and the asthma patients. The performance of the aforementioned classifiers was evaluated based on the percentage of the correctly classified people in each of the three classes. Moreover, two alternative statistical measures were calculated in order to evaluate the effectiveness of the proposed method: the true positive ratio (TPR), defined as the percentage of patients who are correctly identified, also known as sensitivity, and the true negative ratio (TNR), or specificity, which is the percentage of subjects who are classified as not suffering from any illness. For calculating the last two measures, it is necessary to consider the samples as belonging to normal and abnormal groups. To this end, all samples associated with the healthy subjects were labelled as members of the normal class, while the other samples were considered to belong to the abnormal class.

Figs. 7 and 8 show a general overview of our methodology for modelling the signals using ARNNs and FFs-ARNN respectively, in order to classify the subjects into distinct groups. For each subject, an individual model was constructed. Following the termination of the training procedure, the network coefficients associated with the formed attractors were selected to perform the classification task.

5. Results

As indicated in Figs. 7 and 8, the proposed modelling scheme allows
the extraction of FVs based on the model’s attractors. Setting $m=3$ and $c=2$, the length of FVs turns out to be 6 for both ARNN and FFs-ARNN modelling approaches. These parameters were chosen empirically due to the fact that their high values cause some problems regarding an increase in the training time of the model and, also, the large number of features leads to low performance of the classifier in the test part because of overtraining.

Table 1 summarizes the classification results of the lung sounds based on the K-fold strategy, with $K=4$. This table compares the percentage of correct classifications based on our proposed approach under the following three conditions: the FVs corresponding to RQA features, the FVs extracted from temp-ARNN and temp-FFsARNN based on temporal modelling of the time series, as well as the FVs obtained from the RQA-ARNN and RQA-FFsARNN using temporal modelling of the RQA sequences. Table 1 also shows the classification results for the separate channels. In addition, four methods namely RQA [27], AR coefficients [4,6], wavelet transform [10,15], and MFCC [8–10] are compared with others through the table. Similar to RQA, these features were calculated in the subsequent temporal windows. Also, in order to compare our methodology with the methods stated in the literature, the length of FVs was set to 6 for all of them. To achieve this goal, we calculated these features in specific temporal windows and averaged across all windows. According to Table 1, the highest performance was achieved by the RQA-FFsARNN method, demonstrating correct classification results in 91.75% of the cases in the testing phase (channel 2). The results of the testing phase indicate that in the best FVs extraction technique, the performance of the methods is superior in channels 1, 2, and 5, compared to the other channels.

Table 2 provides the values of the sensitivity and specificity in channels 1, 2, and 5. Moreover, area under curve (AUC) and F1-scores, as two important statistical measures, are listed in Table 3. Based on the RQA features, the performance of the joint RQA-ARNN techniques are superior in discriminating between the healthy and the non-healthy subjects. In addition, the execution time of our best approaches against others is illustrated in Table 4, containing the mean value of the run time for all test data.

In Fig. 9, the performance of the different techniques relying on the extraction of FVs by combining the results obtained from different channels is compared. To this end, the classification was made based on the results indicated by voting among six channels. Table 5 indicates the confusion matrix of the best method in the training and the testing parts. It can be seen that the method correctly distinguishes between the healthy and the non-healthy subjects.

In order to evaluate the performance of the proposed RQA-
FFsARNN, our methodology was tested on the recorded LS from new subjects. The new dataset contains LS from six patients suffering from COPD, nine asthma patients, and six healthy subjects; these data were never seen by the classifier. Tables 6 and 7 indicate the confusion matrix of RQA-FFsARNN and the performance of all methods for the new subjects, respectively. The efficiency of our methodology in comparison with the other approaches is verified through these tables.

The analysis of variance (ANOVA) was performed to probe significant differences among the extracted FVs. Figs. 10–12 illustrate the results of the one-way ANOVA as well as box plots for the RQA-ARNN, RQA-FFsARNN, and Temp-ARNN methods, respectively. For simplicity, only the ANOVA test was performed for each dimension corresponding to ch5 and the samples were considered to be in two categories, namely healthy and non-healthy. Based on the p-values of the ANOVA test listed in Table 8, for the RQA-ARNN method and in dimensions 1, 2, and 3, the mean values of data in each category are significantly different from those of the other groups. In addition, significant differences between those groups are observed in the RQA-FFsARNN method in dimensions 1, 3, and 6. In contrast, the reported p-values for all channels in the Temp-ARNN method are higher than those of the other methods, implying that in this method the difference is not meaningful.

Due to the six-dimensional structure of the extracted features, we have performed multivariate analysis of variance (MANOVA). The MANOVA test indicated p-values of 0.2318, 7.4079e−04, and 1.2894e−06 for the Temp-ARNN, RQA-ARNN, and RQA-FFsARNN methods, respectively. Therefore, the joint RQA-ARNN modelling approach allows meaningful discrimination of data samples outperforming the static RQA and also the joint Temp-ARNN methods.

![Fig. 12. The box plots of FVs extracted from Temp-ARNN method in different dimensions (Dim 1, 2, ...).](image-url)
6. Discussion

The results generally indicate that the best performance was attained by the joint RQA-ARNN-based methods, namely RQA-ARNN and RQA-FFsARNN. The FFs-based models allow the localization of the input samples within predefined sub-spaces so that the reconstructed models can be constructed by the integration of the modelling procedures in these subspaces. Since each subspace consists of input patterns corresponding to a specific linguistic variable, the local linguistic model serves as a discriminating model with performance superior to the non-fuzzy models.

It is worth noting that our methodology explores the dynamic information embedded in the temporal changes of the respiratory signals. Also, we tried to indicate the dynamic behaviour of the RQA features in term of novel FFs-ARNN parameters as a new set of features. Therefore, instead of using RQA in all the temporal windows as a set of features, we had just considered their temporal changes in the successive windows. In other words, in our methodology, the length of FV becomes less than the conventional feature extraction techniques.

The idea of using the RQA sequences as inputs to the model is that it would allow discrimination of lung sounds in the reconstructed phase space as is indicated by Figs. 10 and 11. The classification performance of temporal modelling strategies is dramatically inferior to that of the RQA-based approaches. However, static RQA features lead to a better performance than that achievable by temporal modelling.

In this study, all ARNN-based models are constructed by using data acquired from separate channels, in such a way that each model corresponds to one location of the thorax. The results of the combined local models are assessed by voting among all channels. Consequently, it is advisable to consider the correlation between channel information in sequence modelling in order to develop general models with more robust FVs.

Comparing the results presented in this work with those presented in the literature cited in the introduction part involving respiratory signal processing would be difficult due to the differences in the acquisition hardware, the variations in the modelling methodology employed, and the variety of the pulmonary abnormalities to be detected. However, we implemented some of the well-known methodologies on our database. As a conventional method, the AR coefficients were selected based on two distinctive studies. In one of them [6], the pathological and the healthy subjects were analysed only in normal and abnormal classes, and the acquired LS only recorded from two positions of the chest. In the work of Charleston et al. [4], the effectiveness of univariate autoregressive (UAR) models was verified for multichannel recording. Since the extracted features of all the channels in each time window were concatenated in [4], the dimension of the obtained FV would be much higher than the proposed RQA-FFsARNN. The static analysis of the lung sounds was implemented in [10,15] using wavelets transform and the statistical features of subbands were considered as a set of features. In addition, we have compared the MFCC coefficients of LS, which were examined on the respiratory cycles in [8,10]. Comparing the results of these approaches with our new insight into the dynamical modelling verified that RQA-FFsARNN has the higher performance in distinguishing between the three types of subjects while it uses fewer lengths of FV.

There have been three separate studies [26,28,36] using the same database as employed in this work, one of which [26] was related to pulmonary sound classification. In particular, Goudarzi et al. considered the RQA sequences as the input and extracted the FVs based on the recurrent links of a novel form of FFs [26]. Their method identifies three types of signals with a correct classification rate of approximately 75% using the recorded sounds from a single channel, while our multichannel approach leads to a better classification rate. In addition, we report the sensitivity and specificity as well. The evaluation of these statistical measures is important in order to confirm the ability of the proposed method to correctly distinguish healthy from non-healthy subjects.

7. Conclusion

An automated scheme for classification of the recorded multichannel respiratory signals was developed. All the signals collected from the subjects were modelled using the attractor recurrent neural networks (ARNN). Besides the introduction of the fuzzy structure of the ARNN (FFs-ARNN), the ability of the attractor parameters in both structures (fuzzy and non-fuzzy) were examined to distinguish three types of lung sounds. The network attractors are formed whenever two types of sequences are fed as inputs, namely the temporal sequence of time series and the RQA feature sequences. Results with a high level of accuracy were achieved by both fuzzy and non-fuzzy structures of the ARNN models with RQA sequences applied as the inputs. In contrast, the worst results were associated with the time-series sequence modelling features. In general, the results verified that dynamic information embedded in the temporal changes of the respiratory system could represent distinctive information of the abnormalities through a fewer number of features. Although the results demonstrate that the local models for each recording channel exhibit encouraging performance in the classification of healthy and non-healthy subjects, combining the local information of the microphones could serve as a promising approach to improve the three-class differentiation scheme.

References


