ABSTRACT

Pattern analysis of blood oxygen saturation is important for gaining insights into the cardiorespiratory control system, real-time monitoring during operations, identifying potential predictors for the diagnosis of disease severity, and improving the hospitalization of patients with critical chronic diseases. This paper investigates the use of nonlinear dynamics features for machine learning and classification of blood oxygen saturation signals in healthy young and healthy old subjects. The validation of the feature reliability for the signal variability analysis has a clinical implication for differentiating blood oxygen saturation in patients with respect to the particular influence of aging, when patient’s data become available.

Index Terms — Blood oxygen saturation, multiscale entropy, time-shift multiscale entropy, fuzzy recurrence plots, support vector machines.

I. INTRODUCTION

Hemoglobin is the protein molecule in red blood cells, which carries oxygen from the lungs throughout the body tissues and returns carbon dioxide from the tissues back to the lungs. A pulse oximeter non-invasively uses two frequencies of light, which are red and infrared, to determine the percentage of hemoglobin in the blood that is saturated with oxygen. This percentage is called blood oxygen saturation, or $SpO_2$. Pulse oximetry is particularly convenient for noninvasive continuous measurement of blood oxygen saturation. Pulse oximetry is a useful method for diagnosing if a patient’s oxygenation is unstable, and evaluating the effectiveness of or the need for supplemental oxygen in many clinical settings, including critical care [1], during and after surgery [2], and recovery [3]. Because pulse oximeters are simple to use and have the ability to provide continuous and immediate oxygen saturation values, these monitors become important devices adopted in emergency medicine [4], and also very useful for patients with acute cardiac problems [5], especially chronic obstructive pulmonary disease [6], or for diagnosis of some sleep disorders such as apnea and hypopnea [1], [7].

Given the importance of the measure of blood oxygen saturation, study on pattern analysis of oxygen saturation variability has been rarely found in literature. A recent study [8] has reported on the use of fractals and nonlinear dynamics methods for characterizing the pattern of oxygen saturation in healthy young and healthy old populations with an expectation to identify its potential application for differentiating patterns in diseased patient subsets.

This paper investigates the use of three nonlinear dynamics methods, which are the multiscale entropy, time-shift multiscale entropy, and fuzzy recurrence plots, for characterizing patterns of blood oxygen saturation signals in healthy subjects with respect to the effect of aging. The performance of each method is cross-validated with a pattern classifier known as the least-squares support vector machines. The remaining of the paper is organized as follows. Section II describes the three nonlinear dynamics methods. Section III includes the presentation and discussion of the experimental results. Finally, Section IV is the conclusion of the research finding.

II. METHODS

II-A. Multiscale entropy

The multiscale entropy (MSE) [9] was developed to measure the irregularity of time series at multiple scales by computing the sample entropy (SampEn) [10] for a set of averaged non-overlapping time points of the original time series. In other words, the MSE works by applying a “coarse-graining” process to the original time series to generate several time series of different scales, and then computing SampEn values for all coarse-grained time series, which are plotted as a function of the scale factor. The mathematical expression of MSE can be established by considering a time series $X$ of length $N$ taken at regular intervals: $X = (x_1, x_2, \ldots, x_N)$, and for a scale factor $\tau$, a new time series $X^\tau$ is created by the MSE as follows [9]:

$$X^\tau_j = \frac{1}{\tau} \sum_{i=(j-1)\tau+1}^{j\tau} x_i, \quad 1 \leq j \leq N/\tau,$$

(1)

where each of which is to be computed by SampEn.

SampEn is a measure of irregularity in time series. The formulation of SampEn is briefly described as follows [10]. Consider a time series $X$ of length $N$ taken
at regular intervals: \( X = (x_1, x_2, \ldots, x_N) \), and a given embedding dimension \( m \), a set of newly reconstructed time series from \( X \), denoted as \( Y^m \), can be established as \( Y^m = (y^m_1, y^m_2, \ldots, y^m_{N-m+1}) \), where \( y^m_i = (x_i, x_{i+1}, \ldots, x_{i+m-1}) \), \( i = 1, 2, \ldots, N - m + 1 \). The probability of vector \( y^m_i \) being similar to vectors \( y^m_j \) is computed as \( (N - m - 1)^{-1} \) times the number of vectors \( y^m_j \) within a similarity tolerance of \( y^m_i \), where self-matches are excluded, and mathematically expressed as follows

\[
B^m_i(r) = \frac{1}{N - m - 1} \sum_{j=1}^{N-m} H[d(y^m_i, y^m_j)], i \neq j, \quad (2)
\]

where \( r \) is a real positive value for the similarity tolerance, and \( H(d(y^m_i, y^m_j)) \) is the Heaviside function, defined as

\[
H[d(y^m_i, y^m_j)] = \begin{cases} 
1 & : d(y^m_i, y^m_j) \leq r \\
0 & : d(y^m_i, y^m_j) > r 
\end{cases} \quad (3)
\]

The distance between the two vectors is obtained by using the Chebyshev distance or the \( L_{\infty} \) metric, where the distance between two vectors is the largest of their differences along any coordinate dimension and mathematically expressed as

\[
d(y^m_i, y^m_j) = \max_k(|x_{i+k-1} - x_{j+k-1}|), k = 1, 2, \ldots, m. \quad (4)
\]

The probability of pairs of vectors or data points of length \( m \) having the Chebyshev distance \( \leq r \), denoted as \( B^m(r) \), is expressed as

\[
B^m(r) = \frac{1}{N - m} \sum_{i=1}^{N-m} B^m_i(r). \quad (5)
\]

Similarly, \( A^m(r) \) is defined as \( (N - m - 1)^{-1} \) times the number of vectors \( y^m_{i+1} \) within a similarity tolerance of \( y^m_{i+1} \), where \( j = 1, \ldots, N - m, j \neq i \), and setting

\[
A^m(r) = \frac{1}{N - m} \sum_{i=1}^{N-m} A^m_i(r). \quad (6)
\]

SampEn is defined as

\[
\text{SampEn}(m, r, N) = -\log \left[ \frac{A^m(r)}{B^m(r)} \right], \quad (7)
\]

where \( A^m(r) \leq B^m(r) \), which is imposed by the Chebyshev distance.

II-B. Time-shift multiscale entropy

The time-shift multiscale entropy [11], denoted as TSME, is based on the Higuchi’s fractal dimension (HFD) [12]. The HFD computes the “mean length” of the curve of a time series by constructing a set of new time series that has the property of a fractal curve over all time scales as each time series can be considered a reduced scale form of the whole. A set of new time series constructed from the original time series by the HFD, which are utilized in this study, are based on the consideration of the phase distribution. This phase distribution can reveal strong effects of the irregularity of time series [13].

The new time series generated by the HFD are constructed as follows by once again considering the time series \( X \) of length \( N \): \( X = (x_1, x_2, \ldots, x_N) \). Let \( \beta \) and \( k \) be positive integers, where \( \beta = 1, 2, \ldots, k \), then \( k \) new time series can be generated using the following equation [12]:

\[
X^\beta_k = (x_{\beta}, x_{\beta+1}, \ldots, x_{\beta+k-1}), \quad (8)
\]

where \( \lfloor \frac{N-\beta}{k} \rfloor \) is the “floor” function that rounds \( \frac{N-\beta}{k} \) to the largest integer not exceeding \( \frac{N-\beta}{k} \).

From Equation (8), \( \beta \) and \( k \) indicate the initial time point and time interval, respectively, that is for a given time interval \( k \), \( k \) new time series are constructed using \( k \) time shifts.

The TSME method works by constructing \( k \) time-shift series for a given time interval \( k \), then computing SampEn for all time-shift time series, denoted as \( TSME^\beta_k \). \( \beta = 1, \ldots, k \). The TSME for each \( k \), denoted as \( TSME_k \), \( k = 1, \ldots, k_{\text{max}} \), is defined as the average value of all \( TSME^\beta_k \), that is

\[
TSME_k = \frac{1}{k} \sum_{\beta=1}^{k} TSME^\beta_k. \quad (9)
\]

II-C. Fuzzy recurrence plots

The method of fuzzy recurrence plots (FRP) [14] was developed as an extension of the method of recurrence plots (RP) [15], where the output of the former results in a grayscale image and the latter is a binary image. Each pixel in the FRP or RP indicates the recurrence of a corresponding state of a dynamical system. In mathematical terms, let \( X = \{x_i\} \) be a set of phase-space states, in which \( x_i \) is the \( i \)-th state of a dynamical system in \( m \) embedding dimensions. An RP is an \( N \times N \) image in which a pixel \( (i, j) \in \{0, 1\}, i = 1, \ldots, N, j = 1, \ldots, N, \) is 0 if \( x_i \) and \( x_j \) are considered to be closed to each other [15]. A symmetrical RP is defined as [16]:

\[
RP(i, j) = H(\epsilon - \|x_i - x_j\|), \quad (10)
\]

where \( RP(i, j) \) is an element \( (i, j) \) of the recurrence image \( RP \), \( \epsilon \) is a similarity threshold, and \( H(\cdot) \) is the Heaviside function.
step function or the unit step function that yields either 0 or 1 if \((\epsilon - \|x_i - x_j\|) < 0\) or otherwise, respectively.

As an RP is a binary image, its textual information is therefore limited. The concept of a fuzzy recurrence plot (FRP) [14] allows to construct the recurrence image that takes values in [0, 1] to enhance the texture and to relax the selection of the critical threshold \(\epsilon\) required by the RP analysis. Let \(V = \{v\}\) be the set of fuzzy clusters of the states. A fuzzy (binary) relation \(R\) from \(X\) to \(V\) is a fuzzy subset of \(X \times V\) characterized by a fuzzy membership function \(\mu \in [0, 1]\). This fuzzy membership grade expresses the strength of relation of each pair \((x, v)\) in \(R\) that has the following properties [17]: (1) reflexivity: \(\mu(x, x) = 1, \forall x \in X\), (2) symmetry: \(\mu(x, v) = \mu(v, x), \forall x \in X, \forall v \in V\), and (3) transitivity: \(\mu(x, z) = \vee_v[\mu(x, v) \wedge \mu(v, z)], \forall x \in X, \forall z \in Z\), which is called the max-min composition, where the symbols \(\vee\) and \(\wedge\) stand for max and min, respectively.

The fuzzy \(c\)-means (FCM) algorithm [18] can be applied to obtain the fuzzy clusters of the phase-space states to determine the closeness between the states and their fuzzy cluster centers, based on which the inference of the similarity between the pairs of the states can be made using the max-min composition of a fuzzy relation. Let \(\{x_1, x_2, \ldots, x_N\}\) be a set of phase-space states of the system, the FCM algorithm tries to minimize the following fuzzy objective function using a numerical solution [18]:

\[
J(U, Z) = \sum_{i=1}^{N} \sum_{j=1}^{c} (\mu_{ij})^w [d(x_i, z_j)]^2, \tag{11}
\]

where \(c\) is the number of clusters, \(1 < c < N\), \(w \in [1, \infty)\) is the fuzzy weighting exponent, \(U = [\mu_{ij}], i = 1, \ldots, N, j = 1, \ldots, c\), is the matrix of the fuzzy \(c\)-partition, \(Z = (z_1, z_2, \ldots, z_c)\) is the vector of cluster centers, \(z_j\) is the center of cluster \(j\), and \(d(x_i, z_j)\) is any inner-product induced norm metric.

The fuzzy objective function expressed in Equation (11) is subject to

\[
\sum_{j=1}^{c} \mu_{ij} = 1, i = 1, \ldots, N \tag{12}
\]

where \(\mu_{ij} \in [0, 1], i = 1, \ldots, N, j = 1, \ldots, c\).

In order to optimally determine \(U\) and \(Z\), a numerical solution to the minimization of the objective function \(J(U, Z)\) is by an iterative process of updating \(U\) and \(Z\) until some convergence is reached. Further details about the construction of a FRP can be obtained in [14].

### III. RESULTS

The MSE, TSME, and FRP were applied to extract features of \(SpO_2\) signals using the Pattern Analysis of Oxygen Saturation Variability database of the PhysioBank databases [8], that can be publically accessed at https://www.physionet.org/physiobank/database/ovs/. This database contains one hour oxygen saturation measurements of 36 healthy subjects. The young population (< 35 years) includes 20 subjects (mean age = 21.0 years, standard deviation = 1.36 years), and the old population (≥ 35 years) includes 16 subjects (mean age = 50.0 years, standard deviation = 10.4 years).

To extract the MSE feature, we set \(m=1, r = 0.5 \times \sigma\), where \(\sigma\) is the standard deviation of the signal, and \(r = 5\). For the extraction of the TSME feature, we set \(m=1, r = 0.5 \times \sigma\), and \(k_{max} = 5\). For the construction of FRP, the number of fuzzy clusters \(c = 10\), and 19 features of the gray-level co-occurrence matrix (GLCM) [19] were extracted. The 20 FRP-based GLCM features are: autocorrelation (F1), cluster prominence (F2), cluster shade (F3), contrast (F4), correlation (F5), difference entropy (F6), difference variance (F7), dissimilarity (F8), energy (F9), entropy (F10), homogeneity (F11), information measure of correlation 1 (F12), information measure of correlation 2 (F13), inverse difference (F14), maximum probability (F15), sum average (F16), sum entropy (F17), sum of squares variance (F18), and sum variance (F19), which were described in [20]. One signal from the young population and two signals from the old population were excluded in the FRP analysis due to the inability to compute several of the GLCM features. Figure 1 shows the FRPs of the first 2000 samples of the \(SpO_2\) signals of a healthy young (female, 23 years old) and a healthy old (female, 43 years old) subjects, based on which the GLCM features were extracted.

| Table I: Mean values of MSE features with 5 scales (S1, ..., S5) of \(SpO_2\) signals. |
|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| Age Group     | S1             | S2             | S3             | S4             | S5             |
| Young         | 0.1412         | 0.1167         | 0.2682         | 0.2245         | 0.3806         |
| Old           | 0.1605         | 0.1049         | 0.2227         | 0.2101         | 0.3275         |

| Table II: Mean values of TSME features with 5 scales (S1, ..., S5) of \(SpO_2\) signals. |
|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| Age Group     | S1             | S2             | S3             | S4             | S5             |
| Young         | 0.1411         | 0.1128         | 0.2340         | 0.2096         | 0.3222         |
| Old           | 0.1604         | 0.1001         | 0.2352         | 0.1902         | 0.3058         |

| Table IV: Cross-validation results of LS-SVM classification of \(SpO_2\) signals of healthy young and healthy old subjects obtained from MSE, TSME, and FRP methods in terms of sensitivity (SEN), specificity (SPE), area under ROC curve (AUC), and accuracy (ACC). |
|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| Method         | SEN (%)         | SPE (%)         | AUC             | ACC (%)         |
| Leave-one-out  |                 |                 |                 |
| MSE            | 100             | 100             | 1               | 100             |
| TSME           | 75.00           | 75.00           | 0.7562          | 75.00           |
| FRP            | 100             | 100             | 1               | 100             |
| 10-fold        |                 |                 |                 |
| MSE            | 81.25           | 80.00           | 0.8469          | 80.56           |
| TSME           | 87.50           | 100             | 0.9938          | 94.44           |
| FRP            | 100             | 100             | 1               | 100             |
Table III: Mean values of 10 FRP-based GLCM features (F1, ..., F10) of \( SpO_2 \) signals.

<table>
<thead>
<tr>
<th>Age Group</th>
<th>F1</th>
<th>F2</th>
<th>F3</th>
<th>F4</th>
<th>F5</th>
<th>F6</th>
<th>F7</th>
<th>F8</th>
<th>F9</th>
<th>F10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Young</td>
<td>35.8373</td>
<td>364.4650</td>
<td>-13.2509</td>
<td>22.9524</td>
<td>-0.2556</td>
<td>0.7362</td>
<td>10.5456</td>
<td>3.3060</td>
<td>0.3885</td>
<td>1.1335</td>
</tr>
<tr>
<td>Old</td>
<td>36.6714</td>
<td>350.1399</td>
<td>9.2761</td>
<td>20.7297</td>
<td>-0.2325</td>
<td>0.7623</td>
<td>10.1144</td>
<td>3.0088</td>
<td>0.4182</td>
<td>1.1118</td>
</tr>
</tbody>
</table>

Fig. 1: FRPs of a healthy young and a healthy old subjects.

Tables I, II, and III show the mean values of the MSE, TSME, and 10 FRP-based GLCM features of the first 2000 samples of \( SpO_2 \) signals of the young and old populations, respectively. For the SME, at scale 1, the SampEn value of the young population is lower than that of the old population, but from scales 2-5, the SampEn values of the young population become higher than those of the old population. For the TSME, the SampEn values at scale 1 of the young and old populations are very similar to those given by the MSE, and the SampEn values at larger scales of the young population tend to be higher than those of the old population. The 10 FRP-based GLCM features show more discriminating values than the MSE and TSME between the young and old populations, particularly the cluster shade (F3) showing negative value (-13.2509) for the young and positive value (9.2761) for the old population.

To test the performance of the \( SpO_2 \) features obtained from the three nonlinear dynamics methods, the least-squares support vector machines (LS-SVM) method [21] was used for pattern classification of the \( SpO_2 \) features between the two populations. Table IV shows the leave-one-out and 10-fold cross-validation results in terms of sensitivity, specificity, area under the receiver characteristic curve (ROC) of the three methods, where the FRP outperforms both the MSE and TSME in both cross-validations. The MSE performed better than the TSME in the leave-one-out cross-validation, and the TSME gave better results than the MSE in the 10-fold cross-validation.

IV. CONCLUSION

The three nonlinear dynamics methods have been presented and tested for the analysis and classification of \( SpO_2 \) signals obtained from healthy young and old populations. Among the three methods, the FRP appears to be the most effective and robust for characterizing patterns with respect to short signal variability in different age groups. The high performance of the FRP is due to the capture of textural information provided by the grayscale images of the two groups. The finding may have an important implication for gaining insight into patterns of \( SpO_2 \) in patients of the diseases to improve diagnosis and clinical treatment.

V. REFERENCES