

Recurrence Quantification Analysis as a Tool for Discrimination Among Different Dynamics Classes: The Heart Rate Variability Associated to Different Age Groups

Laurita dos Santos, Joaquim J. Barroso, Moacir F. de Godoy,
Elbert E.N. Macau and Ubiratan S. Freitas

Abstract We propose a classification method based on *recurrence quantification analysis* (RQA) combined with *support vector machines* (SVM). This method combines in an effective way various quantitative descriptors to allow a refined discrimination among dynamical non linear systems that presents dynamics which are very similar to each other. To show how effective this methodology is, firstly, based on synthetic data, it is applied on time series generated from the logistic map with nearby parameter values and in the chaotic regime. Next, it is applied to human biosignals, namely, heart rate variability (HRV) time series obtained from four groups of individuals (premature newborns, full-term newborns, healthy young adults, and adults with severe coronary disease). Roughly the proposed methodology works as follows: The signals are transformed into recurrence plots (RP) and a set of RQA statistical features (recurrence rate, determinism, averaged and maximal diagonal line lengths, entropy, laminarity, trapping time, and length of longest vertical line) are extracted to form the input vector for a SVM classifier. Results show that the method discriminates groups of different ages with classification accuracy better than 75 %. Given that heart rate continuously fluctuates over time and reflects different mechanisms to maintain cardiovascular homeostasis of an individual, the results obtained may allow to draw important information on the autonomic control of circulation in normal and diseased conditions.

Keywords Recurrence quantification analysis · Heart rate variability · Different age groups

L. dos Santos

Universidade do Vale do Paraíba - UNIVAP, São José dos Campos, São Paulo, Brazil
e-mail: lauritas9@gmail.com

J.J. Barroso · E.E.N. Macau (✉)

Instituto Nacional de Pesquisas Espaciais, São José dos Campos, São Paulo, Brazil
e-mail: elbert.macau@inpe.br

M.F. de Godoy

Faculdade de Medicina de so José do Rio Preto - Famerp, São José Dos, São Paulo, Brazil

U.S. Freitas

Complexe de Recherche Interprofessionnel en Aérothermochimie - CORIA,
Université de Rouen, Rouen, France

1 Introduction

Heart rate variability (HRV) is a non-invasive measure related to the balance of the activities of sympathetic and parasympathetic divisions of the autonomic nervous system [1]. This variability is normal and indicates the heart ability to response to the environmental and physiological stimuli [2]. The balance of nervous system activities results in a nonlinear behaviour of the HRV time series. In general autonomic and parasympathetic activities attenuate with age [3], which is related with reduction of the HRV [4] (comparing the normal healthy adult and older-age adult).

There are several methods for HRV analysis [5], for example, standard linear techniques (time and frequency domain analysis) and nonlinear methods (correlation dimension analysis, largest Lyapunov exponent, central tendency measure, Poincare plot). However, none of them, from our knowledge, is regarded as universally applicable or effective for all the cases related to HRV analysis. In this study we propose a methodology that is based on the recurrence plot and recurrence quantification analysis. In recent years, recurrence plot (RP) and recurrence quantification analysis (RQA) have been applied to study different dynamics systems [6, 7], in natural science, physics [8], biological systems, and physiological processes involving heart rate variability [9–12]. Given its intrinsic discrete character, RQA is particularly suited for the analysis of HRV time series and allows for a direct quantification of the complex dynamics of heart rhythm modulation [13, 14].

RQA is a useful tool and helps to understand the variation of the autonomic nervous system over time. The major advantage of RQA and recurrence plots (RPs) over standard HRV analysis are their applicability to non stationary data and also their sensitivity to subtle changes in the cardiovascular system dynamics. These aspects enable RPs to be used in the characterization of changes in the basic cardiovascular parameters during both physiological and pathological conditions. But the analysis of HRV time series using only RQA statistics is known as not being able to provide consistent enough information to achieve a suitable classification. And our goal here is to have an effective method that allows one enough sensitivity to properly differentiate systems with very similar dynamics. This desired amplification in the discrimination sensitivity using the SVM in combination with the RQA, will be shown here in the subsequent sections. In this work, combined with SVM, we evaluated RQA measures to discriminate and identify groups of different ages, including information about the system.

2 Materials and Methods

2.1 *Experimental Database*

The study comprised a total of 148 tachograms divided into four groups: 26 full-term newborns (FNB) (8 days on average), 48 premature newborn (PNB) (± 27.4 days), 61 healthy young adults (HYA) (20.7 ± 1.6 years), and 61 adults in preoperative

evaluation for coronary artery bypass grafting for severe coronary disease (SCD) (58.4 ± 10.2 years). All tachograms are from databases from previous studies of Transdisciplinary Nucleus for Chaos and Complexity Study (NUTECC/Brazil) [15, 16]. There are time series with 15 min up to 1 h recording period from patients in a supine rest position without visual and sound stimulations.

The equipment used to collect signal was Polar Monitor (S810i or RS800), which has been proven [17–19] to be feasible and reliable for measuring HRV according to recognized standards [20]. At a sampling rate of 1000 Hz, this device captures successive intervals between heartbeats, namely NN, in the normal sinus rhythm (i.e., initiated by the sinoatrial node). All these studies were approved by the respective ethic committee. All NN time series were filtered to remove artifacts using an adaptive filter which takes into account the peculiarities of the signal to be analyzed [5].

2.2 Recurrence Quantification Analysis

Defined as a repeated occurrence in time of a given state of a system, recurrence is a basic attribute of many dynamical systems. It means that along the time a trajectory comes repetitively close in the state space of points previously visited. Embedding the time series in a appropriate dimensional space and then plotting in a matrix the recurrences according to a tolerance rule results a recurrence plot (RP), which is a graphical representation of the recurrences in the dynamical system. The visual features of such plots are appealing and reveal patterns not previously viewed in the original series [13].

RP represents the autocorrelation in the signal at all possible time scales. Since the diagonal marks the identity in time, long-range correlations are associated to points far from the main diagonal, whereas the elements near the principal diagonal correspond to short-range correlations. Diagonals reflect the repetitive occurrence of similar states in the system dynamics and express the similarity of system behavior in two distinct time sequences. To quantify such features, recurrence quantification analysis (RQA) has been introduced for measuring quantitative information contained in recurrence plots [21].

For instance, the density of recurrence points in a recurrence plot is defined as recurrence rate (RR), giving the probability that a specific rate will recur. Parameters based on the diagonal lines are determinism (DET, the percentage of recurrence points forming diagonals from all recurrence points), averaged diagonal line length (L), maximal diagonal line length (Lmax), and entropy (which denotes the Shannon entropy of the histogram of the lengths of diagonal segments and thus indicates the complexity of the deterministic structure of the system).

Verticals are also important structures in a RP in that they reflect the persistence of one state during some time interval. The parameters derived from vertical lines are laminarity (LAM, the proportion of recurrence points forming verticals), trapping time (TT, the mean length of vertical lines), and the maximal length of a vertical, Vmax. Low TT, LAM, and Vmax values imply high complexity in the systems

dynamics, since the state of the system stays only for a short time in a state similar to the previously occurring state. Theoretically, diagonal and vertical structures would not occur in random (stochastic) as opposed to determinist process [7].

2.3 SVM Classifier

Support Vector Machines (SVMs), developed by [22], are supervised learning techniques used for classification, regression analysis and learning tasks. Such techniques can be applied to the solution of problems related to text categorization, image analysis, and bioinformatics [23]. The main idea behind this classifier is to construct a hyperplane that maximizes the distance (so-called margin) to the nearest data points pertaining to two classes as pictured in Fig. 1.

The classifier was trained from the previously discussed dataset which, in an empirical way, was divided into the training and test sets enumerated in Table 1.

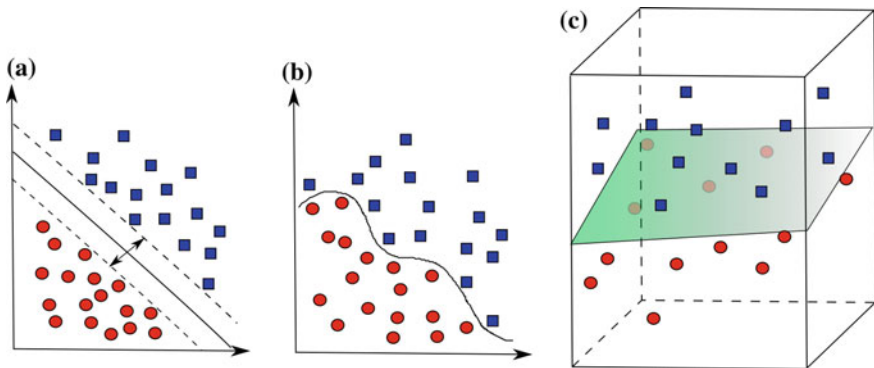


Fig. 1 Examples of separation of two classes using an SVM classifier: **a** two classes linearly separable, **b** two classes with nonlinear separation, and **c** separation achieved by a hyperplane in a high-dimensional space

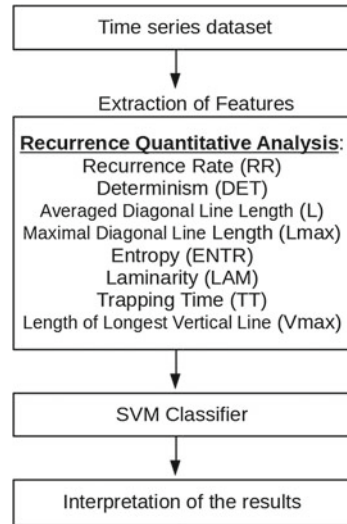
Table 1 Number of training and test sets employed in the SVM classifier

Groups ^a	Training set	Test set	Total of cases for comparison
FNB (26) and PNB (48)	17 and 17	9 and 9	26 and 26
FNB (26) and HYA (61)	17 and 17	9 and 9	26 and 26
PNB (48) and HYA (61)	30 and 30	18 and 18	48 and 48
FNB (26) and SCD (61)	18 and 18	8 and 8	26 and 26
PNB (48) and SCD (61)	30 and 30	18 and 18	48 and 48
HYA (61) and SCD (61)	45 and 45	16 and 16	61 and 61

^aFNB full-term newborn, PNB premature newborn

HYA—healthy young adult, SCD—adult with severe coronary disease

Fig. 2 Structure of the methodology for discrimination of HRV clinical groups



The class label (PNB, FNB, HYA, SCD) for each NN interval time series was assigned by a cardiologist.

For each time series, eight RQA features were extracted to form the input for the classification step (Fig. 2). The SVM classifier was assessed from the LIBSVM open library [23] and executed 100 times for each RQA feature for comparison between two clinical groups. Detailed information about learning and classification algorithm can be found in [22, 23]. For each execution of the code, the training and test cases were randomly selected from which was obtained the average accuracy, defined as the percentage ratio of the number of cases correctly classified to the total number of cases used for classification.

3 RQA Plus SVM: Discriminating Almost Similar Dynamics

To show the effectivity of the proposed the methodology RQA+SVM, we used the logistic map time series ($x_{n+1} = r * x_n * (1 - x_n)$) for values $r = 3.68$, $r = 3.7$ and $r = 3.9$ (see Fig. 3). For each value of the dynamic parameter r , 30 time series were generated, each one with 2,000 points (the first 200 points were discarded to allow transients to die out), with $x(0) \in [0.1, 0.8]$, and an incremental step $\Delta x(0) = 0.0241$.

For the study of the RQA measures, the RP parameters to the logistic map were selected embedding dimension ($m = 1$), delay ($\tau = 1$) and threshold radius ($\varepsilon = 0.1$). Details about these values are given in [7].

To the SVM classifier three groups are assigned (according r values: $r = 3.68$, $r = 3.7$ and $r = 3.9$). We used 21 time series for each group of the training set

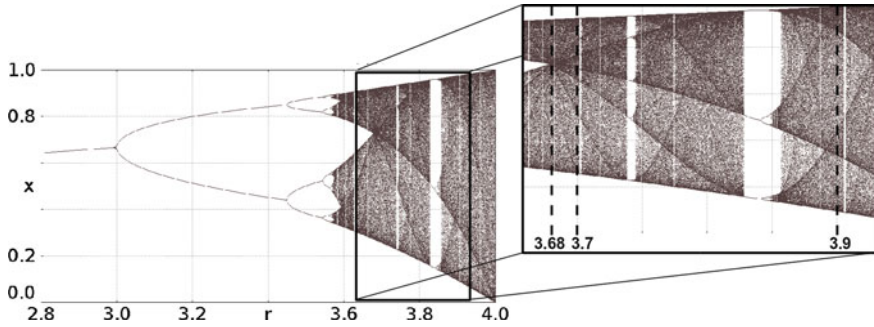


Fig. 3 Bifurcation diagram of the logistic map with a zoomed in view in the region of the r parameters chosen: $r = 3.68, r = 3.7$ and $r = 3.9$

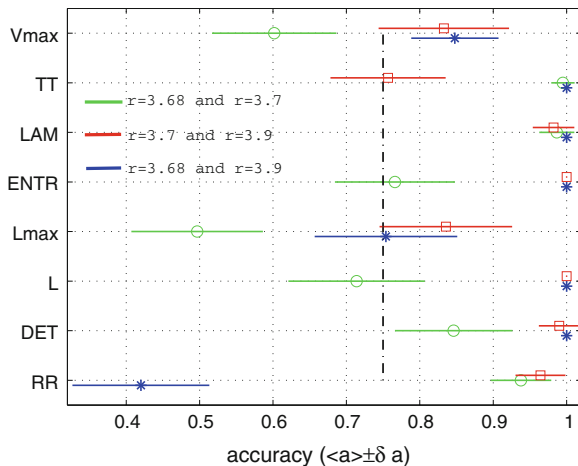


Fig. 4 Average accuracy and standard deviation ($\langle a \rangle \pm \delta a$) obtained by the SVM of the comparison between two time series groups to Logistic Map. (A) $r = 3.68$ and $r = 3.7$, (B) $r = 3.7$ and $r = 3.9$ (red line); $r = 3.68$ and $r = 3.9$ (blue line). The proper level of accuracy (75 % or higher) is indicated by points to the right of the dashed vertical line

and 9 to the test set. The average accuracy and standard deviation obtained by SVM are displayed in Fig. 4. In this figure an accuracy equal to 1 means that all of the cases tested (100 %) were correctly classified, while a zero value means that all cases were not properly classified. For accuracy values above the threshold of 75 %, the dynamics of the analyzed groups are considered to be similar. We can observe that for the pair of groups ($r = 3.68, r = 3.7$) the RQA features are more similar than the pairs of groups ($r = 3.68, r = 3.9$) and ($r = 3.7, r = 3.9$). These results demonstrate the ability of the methodology RQA + SVM to differentiate groups with almost similar dynamics.

4 Application: Using HRV to Discriminate Physiological Age

The main objective of this study was to analyze RQA measures as a tool to discriminate HRV time series recorded from different clinical groups. Typical HRV time series and RP patterns are shown in Figs. 5 and 6, from which the peculiarities of each recurrence plot and the corresponding HRV series are noticeable. For these plots and throughout the present study the RP parameters were selected as: $m = 3$, $\tau = 3$, and $\varepsilon = 8$. The choice of embedding dimension ($m = 3$) was based on results from the false nearest neighbor method [24]. We chose the minimum value for m that presented minimum percentage of false neighbors. This value was adopted for the time series analyzed, standardizing all the data set. Time delay for embedding was set at the first minimum of the mutual information function [25], since the embedded signals have the minimum overlapping information. The tolerance level, following the recommendation of [26], was selected to ensure the percentage of recurrence points lying between 0.1 and 0.2 % to obtain reliable values for the RP parameters. Detailed discussions about the RP parameters are found in [13, 14, 26].

For each group, the extracted RQA features are displayed in Figs. 7 and 8. We can notice that for the pairs of groups (SCD, HYA) and (FNB, PNB) the RQA features are similar. Then to further examine the ability of RQA features to differentiate groups of different ages we applied SVM classification.

A similar plot to those for the RQA measures (Figs. 7 and 8) is obtained when using the mean value and standard deviation of the HRV time series. We see in Fig. 9a that the groups FNB and PNB can be distinguished from the groups SCD and HYA in terms of the average values of the NN intervals. But since an NN interval gives

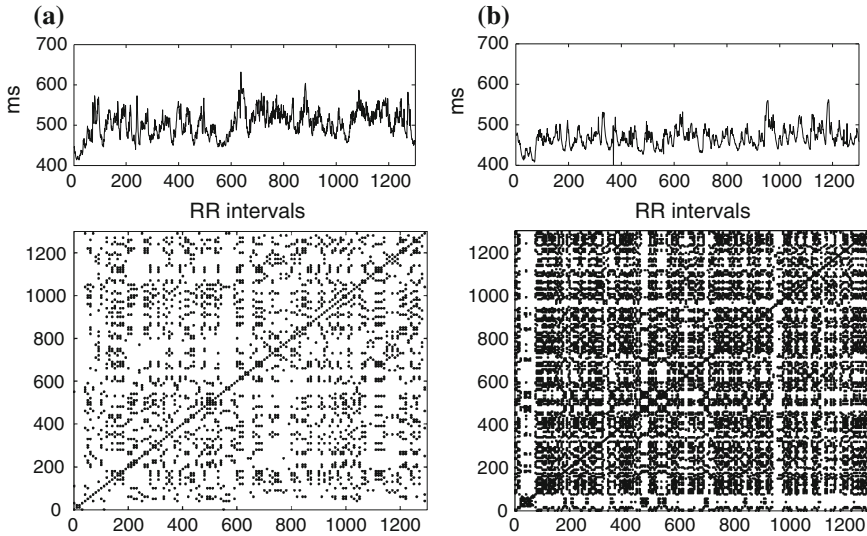


Fig. 5 Examples of NN time series and RPs for **a** FNB and **b** PNB groups (embedding dimension = 3, delay = 3 and threshold = 8)

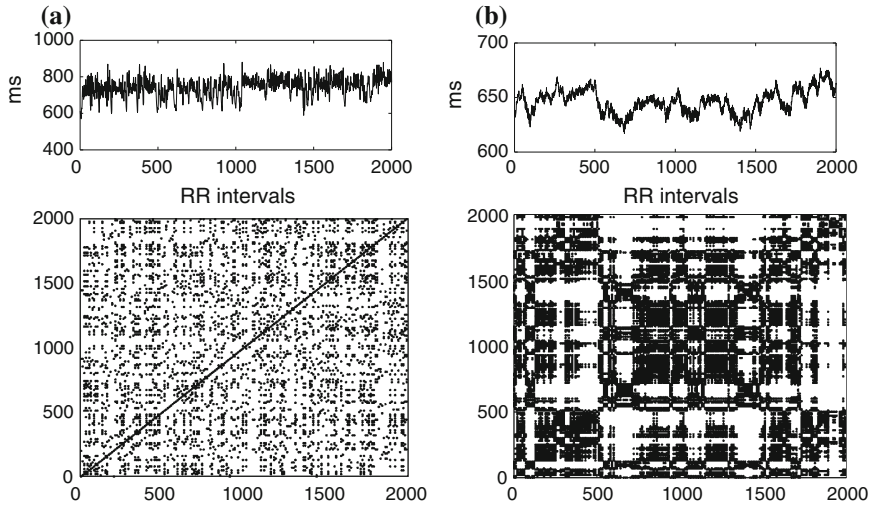


Fig. 6 Examples of NN time series and RPs for **a** HYA and **b** SCD groups (embedding dimension = 3, delay = 3 and threshold = 8). Example of RPs for each time series groups (embedding dimension = 3, delay = 3 and threshold = 8)

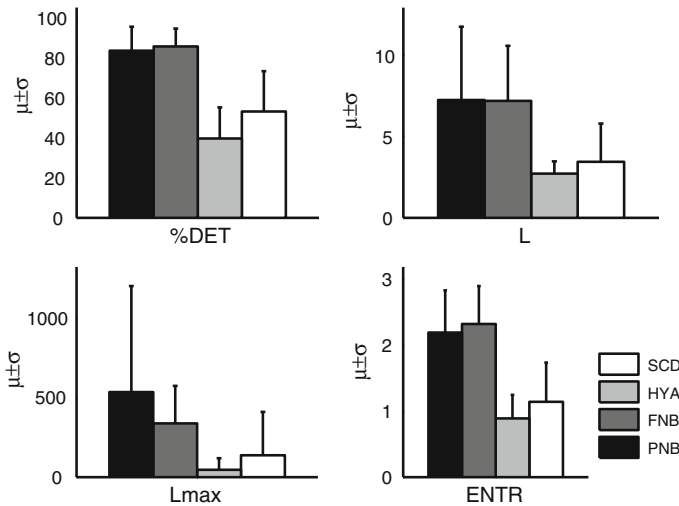


Fig. 7 Average values and standard deviation to RQA diagonal parameters for each group

(in milliseconds) the duration of a heartbeat, the values displayed in Fig. 9a are only correlated with the mean heart rate for the time series of each group. However we emphasize that the mean value of the NN interval, i.e., the heartbeat average is not enough to characterize the homeostasis of an individual, which is a dynamical process that is reflected in the heart rate variability. On the other hand, upon analyzing the

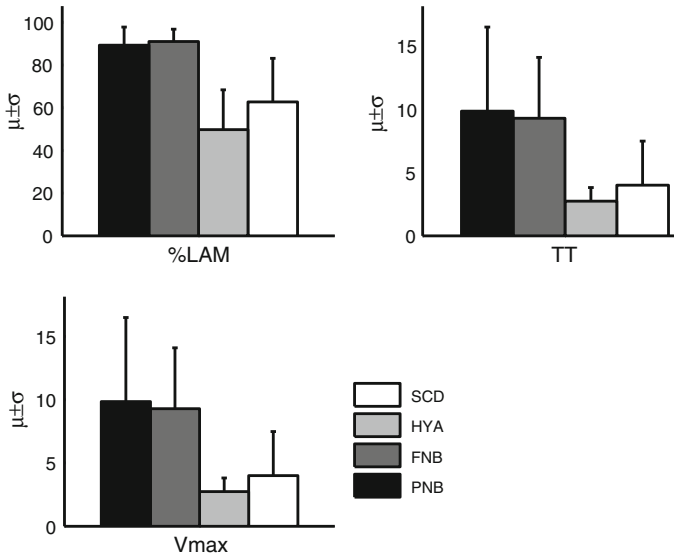


Fig. 8 Average values and standard deviations of vertical-based RQA measures for each group

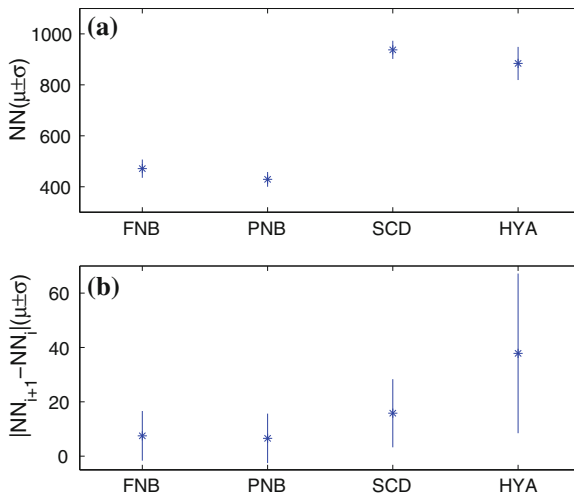


Fig. 9 Average values for the full set of HRV time series by taking for each series: **a** the NN interval and **b** the beat-to-beat NN interval variability

set of series in terms of beat-to-beat NN interval variability, the separation between groups is no longer possible as demonstrated in Fig. 9b.

The average accuracy values of RQA indexes obtained from SVM through comparison between groups of different ages are reported in Fig. 10. It is seen that RQA indices are better at distinguishing groups the larger is the age difference. In fact, for

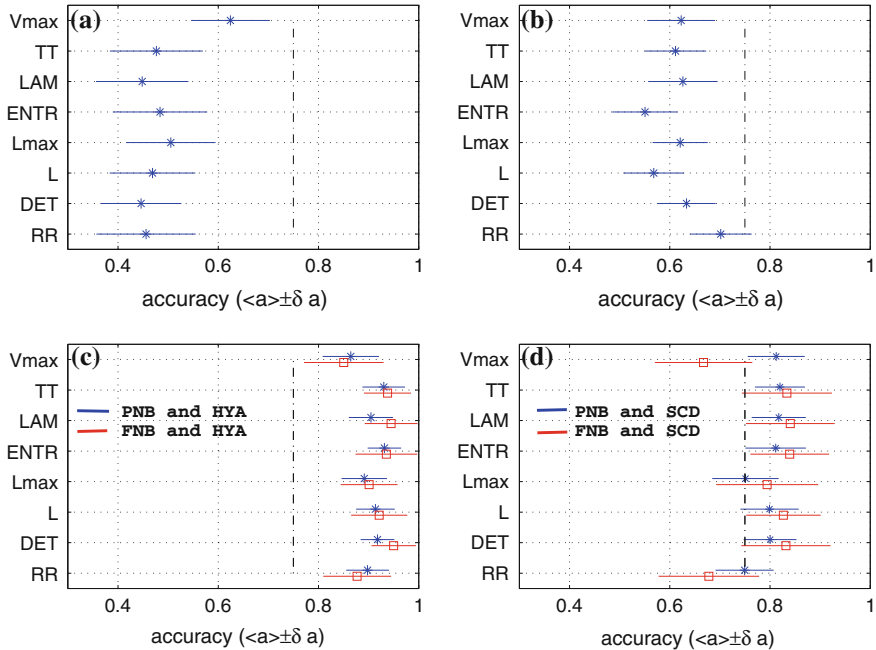


Fig. 10 Average accuracy and standard deviation ($\langle a \rangle \pm \delta a$) obtained by SVM from the comparison between two NN intervals time series groups to RQA indexes. **a** Full-term newborn (FNB) and premature newborn (PNB), **b** healthy young adult (HYA) and adult with severe coronary disease (SCD), **c** premature newborn and healthy young adult (red line); full-term newborn and healthy young adult (blue line), **d** premature newborn and adult with severe coronary disease (blue line); full-term newborn and adult with severe coronary disease (red line)

close age difference as in Fig. 10a, b the average accuracies are restricted to 50 and 60%, respectively. Nevertheless, this might indicate that age difference between the HYA and SCD groups is more significant than in groups FNB and PNB. In support to this conclusion, we see in Fig. 5 that the recurrence plots for the groups FNB and PNB look more similar than the RPs for the HYA and SCD groups (Fig. 6).

In addition, comparison of newborns with older individuals yields higher accuracy, namely, 80% as demonstrated in Fig. 10d and 90% in Fig. 10c. It is to be mentioned, however, that a larger age difference does not necessarily imply a larger accuracy, i.e., the larger accuracy in Fig. 10c is related to an age difference smaller than that in Fig. 10d.

5 Conclusion

The present study was concerned with recurrence quantification analysis (RQA) of HRV time series for groups of individuals with different ages. RQA was proven to be a powerful discriminatory tool to detect the degree of determinism of the systems

examined. Among the four groups studied, all the RQA measures (Figs. 7 and 8) were lower in the healthy young adults (HYA). Low TT, Lam, and Vmax, for instance, imply high complexity in the system's dynamics. This result is in line with the concept that high complexity is a general feature of healthy dynamics compared to pathological conditions.

We also verified that RQA measures were able to differentiate groups, with the results demonstrating that better discrimination is achieved the higher the age difference is. It was noted in Fig. 10c, d, however, that a higher age difference does not imply a higher discriminatory accuracy. The closeness of the comparison of the SCD group with the newborns (PNB and FNB) and the higher degree of dissimilarity between the HYA group and the newborns reflect the fact that the comparisons were quantified in terms of HRV, which is age dependent. This result shows that the HRV decreases with age as described in [3, 4].

Given that HRV time series reflects the complex interactions of different control loops of the cardiovascular control system, the results obtained here provide important information on the autonomic control of circulation in normal and diseased conditions. In addition, the approach discussed here permits an automatic analysis of a large number of time series, thus making the method useful in clinical sets and in epidemiological studies to analyze HRV series or other biomedical signals.

Acknowledgments The authors thank CAPES/Brazil (process 8954 – 11 – 9) and CNPq/Brazil (process 151597/2013 – 8) for financial support. E.E.N.M. would like to thanks CNPq and FAPESP (process 2011/50151 – 0).

References

1. Wijngaarden, M.A., Pijl, H., van Dijk, K.W., Klaassen, E.S., Burggraaf, J.: *Clin. Endocrinol.* **79**, 648 (2013)
2. Kuusela, T.: Heart rate variability (HRV) signal analysis. In: Kamath, M.V., Morillo, C., Upton, A. (eds.) *Methodological Aspects of Heart Rate Variability Analysis*, pp. 9–40. CRC Press, Boca Raton (2013)
3. Yukishita, T., Lee, K., Kim, S., Ando Y.Y., Kobayashi, A., Shirasawa, T., Kobayashi, H.: *Anti-Aging Med.* **7**(8), 94 (2010)
4. Moodithaya, S., Avadhany, S.T.: *J. Aging Res.* **2012**(679345): 1–7 (2012)
5. dos Santos, L., Barroso, J.J., Macau, E.E.N., de Godoy, M.F.: *Med. Eng. Phys.* **35**, 1778 (2013)
6. Webber Jr, C.L., Zbilut, J.P.: *J. Appl. Physiol.* **76**, 965 (1994)
7. Marwan, N., Romano, M., Thiel, M., Kurths, J.: *Phys. Rep.* **438**(5–6), 237 (2007)
8. Ngamga, E., Senthilkumar, D., Prasad, A., Parmananda, P., Marwan, N., Kurths, J.: *Phys. Rev. E* **85**, 026217 (2012)
9. Wessel, N., Marwan, N., Meyerfeldt, U., Schirdewan, A., Kurths, J.: *Lecture Notes in Computer Science* **2199**(2199), 295 (2001)
10. Peng, Y., Sun, Z.: *Med. Biol. Eng. Comput.* **49**(1), 25 (2011)
11. Ramírez Ávila, G., Gapelyuk, A., Marwan, N., Stepan, H., Kurths, J., Walther, T., Wessel, N.: *Autonomic Neuroscience: Basic and Clinical* (2013)
12. Mesin, L., Monaco, A., Cattaneo, R.: *BioMed Res. Int.* **2013**, 420–509 (2013)
13. Zbilut, J.P., Webber Jr, C.L.: *Phys. Lett. A* **171**, 199 (1992)

14. Javorka, M., Trunkvalterova, Z., Tonhajzerova, I., Lazarova, Z., Javorkova, J.: *Clin. Physiol. Funct. Imaging* **28**(5), 326 (2008)
15. Selig, F.A., Tonolli, E.R., Godoy, M.F., da Silva, E.V.C.M.: *Arq. Bras. Cardiol.* **96**(6), 443 (2011)
16. Leal, J.C., Petruccic, O., de Godoy, M.F., Braile, D.M.: *Interact. CardioVasc. Thorac. Surg.* **14**, 22 (2012)
17. Gamelin, F.X., Berthoin, S., Bosquet, L.: *Med. Sci. Sports Exerc.* **38**(5), 887 (2006)
18. Vanderlei, L.C., Silva, R.A., Pastre, C.M., Azevedo, F.M., Godoy, M.F.: *Braz. J. Med. Biol. Res.* **41**(10), 854 (2008)
19. Nunan, D., Donovan, G., Jakovljevic, D.G., Hodges, L.D., Sandercock, G.R., Brodie, D.A.: *Med. Sci. Sports Exerc.* **41**(1), 243 (2009)
20. Task Force of the European Society of Cardiology the North American Society of Pacing Electrophysiology. *Circulation* **93**, 1043 (1996)
21. Marwan, N., Kurths, J.: *Phys. Lett. A* **302**, 299 (2002)
22. Cortes, C., Vapnik, V.: *Mach. Learn.* **20**, 273 (1995)
23. National Taiwan University, Taiwan, LIBSVM: A Library for Support Vector Machines (2012). <http://www.csie.ntu.edu.tw/~cjlin/libsvm/>
24. Kennel, M.B., Brown, R., Abarbanel, H.D.I.: *Phys. Rev. A* **45**(6) (1992)
25. Fraser, A.M., Swinney, H.L.: *Phys. Rev. A* **33**(2) (1986)
26. Zbilut, J.P., Thomasson, N., Webber, C.L.: *Med. Eng. Phys.* **24**(43) (2002)