PAPER

A Permutation Entropy analysis to determine significant daily intervals to improve risk stratification tasks from COVID patients

Biomedical Physics & Engineering Express

PAPER

A Permutation Entropy analysis to determine significant daily intervals to improve risk stratification tasks from COVID patients

RECEIVED
17 August 2023

REVISED
26 December 2023

ACCEPTED FOR PUBLICATION
10 January 2024

PUBLISHED
18 January 2024

Zayd Isaac Valdez[1,3], **Luz Alexandra Díaz**[1,3], **Miguel Vizcardo Cornejo**[1,3,*] and **Antonio G Ravelo-García**[2,3]

[1] Escuela de Física, Universidad Nacional de San Agustin de Arequipa, Av. Independencia s/n, Arequipa, 04002, Arequipa, Peru
[2] Departmento de Señales y Comunicaciones, Universidad Las Palmas de Gran Canaria, Campus Tafira, Las Palmas de Gran Canaria, 35017, Las Palmas, Spain
[3] These authors contributed equally to this work.
[*] Author to whom any correspondence should be addressed.

E-mail: zvaldez@unsa.edu.pe, ldiazpar@unsa.edu.pe, mvizcardoc@unsa.edu.pe and antonio.ravelo@ulpgc.es

Keywords: permutation entropy, risk stratification, Covid-19

Abstract

SARS-CoV-2 infection has a wide range of clinical manifestations making its diagnosis difficult, which is an important problem to solve. We evaluated heart rate data extracted from the Stanford University database. The data set considers heart rate and step records of 118 patients, where 90 correspond to healthy individuals and 28 patients with COVID. Each daily record was divided into 5-minute segments, providing 288 data per patient. The date of symptom onset was considered as a reference point to extract subsets of data whose variability was considerable, such as 30 days before the date and 30 days after it. Each of the 60 segments of 288 data per patient was treated using Permutation Entropy, Approximate Entropy, Spectral Entropy and Singular Value Decomposition Entropy. The average of the data from each group was used to construct the circadian profiles which were analyzed using the Mann-Whitney-Wilcoxon test, determining the most relevant 5-minute segments, whose p-value was less than 0.05. In this way, the Spectral Entropy was discarded as it did not show any significantly different segment. The efficiency of the method was reflected in the performance of a logistic model for binary classification proposed in this work, which reflected an accuracy of 94.12% in the PE case, 88% in the ApEn case and 94% in the SVDE case. The proposed analysis turns out to be highly efficient when detecting significant segments that allow improving the classification tasks carried out by Machine Learning models, which provides a basis for the study of statistics such as entropy to delimit databases and improve the performance of classifier models.

1. Introduction

1.1. Pandemic description

A new type of coronavirus, capable of causing mild to severe infections in humans, was identified as 2019-nCoV or COVID-19 and first appeared in Wuhan, China, in late 2019 [1, 2]. From its first appearance it spread rapidly throughout the world despite considerable efforts to contain it. After the large number of new cases and their expansion across different countries, on March 11, 2020, the World Health Organization (WHO) declared COVID-19 a global pandemic. Only until June 3, 2020, approximately 6 months after its outbreak, this virus would have caused more than 6,500,000 cases of infection in 210 countries and around 383,000 deaths [2]. This rapid expansion is due to the ease of contagion since its transmission is through common routes such as by contact and through the air, due to coughing, sneezing, inhalation of drops and contact with mucous membranes, being that the excretion viral is produced in the respiratory tract, saliva, feces and urine [3–5].

1.2. Challenges in diagnosis and use of AI/Smart devices

SARS-CoV-2 infection has a wide range of clinical manifestations ranging from asymptomatic to symptomatic, including respiratory symptoms, fever, shortness of breath, cough, dyspnea, and viral pneumonia and, in severe cases, pneumonia, severe acute respiratory syndrome, heart failure, kidney failure, and even death [6].

Due to how variable the manifestation of the disease is from patient to patient, its diagnosis is a challenge. As a result of this, apart from the routine clinical diagnosis that is usually applied, various diagnostic and classification tools have been developed through the use of artificial intelligence that analyzes different clinical variables, blood tests, computed tomography, x-rays, heart rate , among others [7–9] where our interest lies in those that do it in a non-invasive way such as data from smart watches whose use is currently common [9, 10], where although they show results promising when analyzing time sections, ways of improving them should be sought through more accurate and efficient data selection and processing, respectively.

This study proposes using different types of information entropy as: Permutation Entropy, Approximate Entropy, Spectral Entropy and Singular Value Decomposition Entropy to analyze heart rate variability in hourly blocks of patients without complications or cardiorespiratory compromise. The resulting complexity will be compared with that of patients diagnosed with COVID-19, helping to identify the specific hourly blocks with higher incidences of cardiorespiratory compromise. This information will lead to generating a logistic model for classifying these hourly blocks, ultimately producing a more precise and accurate diagnostic tool.

2. Database

In this work, the database of Stanford University was used. This database contains data on heart rate and number of steps, collected by volunteers through the use of smart watches of different brands. In addition, the university develop a smartphone application where volunteers were able to record their symptoms, date of onset of symptoms, date of diagnosis and recovery date. Only heart rate data was used in the development of this article. The total number of patients in this database is 118, where 90 correspond to healthy individuals and 28 patients with COVID [11].

3. Method

3.1. Permutation entropy (PE)

Permutation Entropy (PE) is a non-linear statistical characteristic that measures the complexity and regularity of a time series, based on the presence of patters in it. This concept was defined and explored by Band and Pompe [12].

$$H(n) = -\sum p(\pi)\log p(\pi) \tag{1}$$

PE algorithm considered all the possible permutation patterns π of consecutive values and accounts for their relative appearance in the time series $p(\pi)$. It is a fact that this statistical characteristic does contemplate the order of appearance of the values.

3.2. Singular value decomposition entropy (SVDE)

The Singular Value Decomposition (SVD), according to [13], can be described as follows: Singular Value Decomposition (SVD) is the factorization of a matrix $\mathbf{A}$ (where $\mathbf{A}_{m,n} \in \mathbb{B}$ in our case, but SVD works for matrices of real numbers as well) into the form $\mathbf{U}\cdot\Sigma\cdot\mathbf{V}^T$. Where $\mathbf{U}$ is an $m\times m$ orthogonal matrix and $\mathbf{V}$ an $n\times n$ orthogonal matrix. The columns in these matrices are, respectively, the left and right-singular vectors of $\mathbf{A}$, were $\mathbf{U}=\mathbf{A}\mathbf{A}^T$ and $\mathbf{V}=\mathbf{A}^T\mathbf{A}$. Σ is a matrix that only contains non-negative σ values along its diagonal and all other entries are zero. Where $\sigma_i=\Sigma_{ii}$, which contains the singular values of $\mathbf{A}$. When the values of σ are arranged in descending order, the singular values (Σ) are unique, though the singular vectors ($\mathbf{U}$ and $\mathbf{V}$) may not be.

In a intuitive sense, the singular value $i(\sigma_i)$ quantifies the proportion of the dataset explained by each vector. Following [14] approach, the entropy of σ can be assessed. Elevated SVD entropy implies that all vectors bear equal significance, signifying that the ecological network's structure resists compression efficiently. Consequently, high entropy suggests heightened complexity in the ecological network [15].

3.3. Approximate entropy (ApEn)

Approximate Entropy (ApEn) is a technique that measures irregularity and complexity, according to the definition given by Steve Pincus [16], in which if the time-series data consists on N elements:

$$\phi^m(r) = \frac{1}{n}\sum_{i=1}^{n} log(C_i^m(r)) \tag{2}$$

$$ApEn(m,\, r,\, N)(u) = \phi^m(r) - \phi^{m+1}(r) \tag{3}$$

Where m is the embedding dimension, r is a threshold and A_i and B_i are the measures of proximity between embedding vectors in m and $m+1$ dimensions respectively.

3.4. Spectral entropy (SE)

According to [17]: It is the computation of spectral power distribution along with forecastability of time-series signal. This entropy is based on Shannon and information entropy in the information data. The spectral entropy of the signal is given by,

$$SE(F) = -\frac{1}{\log N_u \Sigma_u (P_u(F)\log_e P_u(F))} \tag{4}$$

$$SSH(F) = -\Sigma_u (P_h(F)\log_e P_h(F))) \tag{5}$$

Where, $P_u(F)$ represents the power spectral density function, $P_h(F)$ represents the estimation of the Shannon entropy ($SSH(F)$), and N_u signifies the entire frequencies.

It quantifies the level of flatness in the signal spectrum density, serving as an indicator of the regularity of the signal spectrum. Elevated SE values (approaching 1) are linked to a signal resembling white noise, while SE values near zero suggest organized signals akin to a

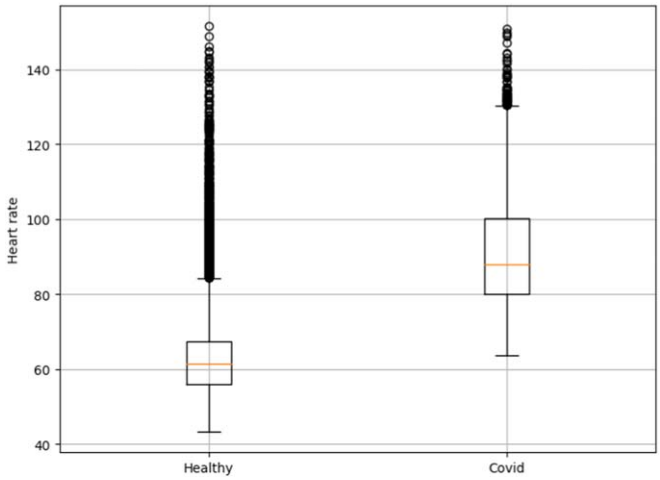

Figure 1. Box and whisker plot of heart-rate for patient belonging to the 'Control' or 'Healthy' and 'COVID' group.

sinusoidal pattern [18]. Consequently, SE evaluates the degree of disorder present in the signal.

3.5. Logistic classifier

Basic libraries whose implementation was made in Python 3.10 were used in order to propose the classification model.

$$p_k(X_i) = \frac{e^{X_i W_k + W_{0,k}}}{\sum_{l=0}^{K-1} e^{X_i W_l + W_{0,l}}} \tag{6}$$

The logistic classifier was implemented using the open source Scikit-Learn library, which was designed in order to contain multifunctional tools to machine learning projects. We considered a *Logistic Regression* (or also called *Max-entropy classifier*) model that is implemented in the Linear Models presented by the library. We considered the solver Limited-memory Broyden-Fletcher-Goldfarb-Shanno algorithm (*ibfgs*) that is based and can approximate the Broyden-Fletcher-Goldfarb-Shanno algorithm.

3.6. Data processing

The main data consisted of heart rate (HR) values with an irregular sampling frequency, therefore it was decided to regularize this by dividing them into 5-minute segments, covering the daily record in 288 segments. Of course, it is possible to analyze time series with a great variety of existing methods in related works, and it is possible to describe that measures by using statistic parameters such us Permutation Entropy (PE), Approximate Entropy (ApEn), Spectral Entropy (SE) and Singular Value Decomposition Entropy (SVDE). Figure 1 shows the representation of the heart rate values for 2 patients belonging to the 'Control' and 'Covid' groups. It can be observed that the means have significantly different values, which could indicate the possibility of differentiating specific segments in the circadian cycle, accentuating said difference in the means. Likewise, the outliers show a consistent range with respect to the limit values for a person: In the case of the healthy patient, their outliers may mean physical activity or cardiovascular effort, while the boxplot for the 'COVID' patient includes these values. atypical within a lower quartile range, in addition to the extension of both extremes indicating the high variability of heart rate in this case.

Data related of each patient was categorized by labels according to their condition: CONTROL and COVID. Dates of symptoms for each patient were used to recognize the point of reference where the HR record will be able to express different values that can mean variability between two considered periods. These two periods were established as 30 days before and 30 days after the date of onset of symptoms, thus allowing to create matrices for each one in order to obtain vectors that can represent 288 5-minutes segments for 60 days. Then, we considered applying PE, ApEn, SE and SVDE algorithm for each of the 288 segments in order to obtain a characteristic vector for each example. This operation was applied to both the CONTROL and COVID group figures 2 and 3.

An average per column was taken from these matrices to create a PE, ApEn, SE and SVDE circadian profiles for each group and these characteristic vectors were analyzed using Mann-Whitney-Wilcoxon test to verify that there were significant difference between them. This method allowed to recognize which 5-minutes segments were relevant to differentiate one class from the other. In this way, the vectors of each patient were filtered in order to maintain segments where the entropy values had significant differences, reducing the number of segments.

The 118 examples (where CONTROL and COVID groups are included) were divided in order to consider only 56 examples (28 COVID examples and 28 CONTROL examples) for a reason related to the overfitting of the logistic mode and that could lead to a biased

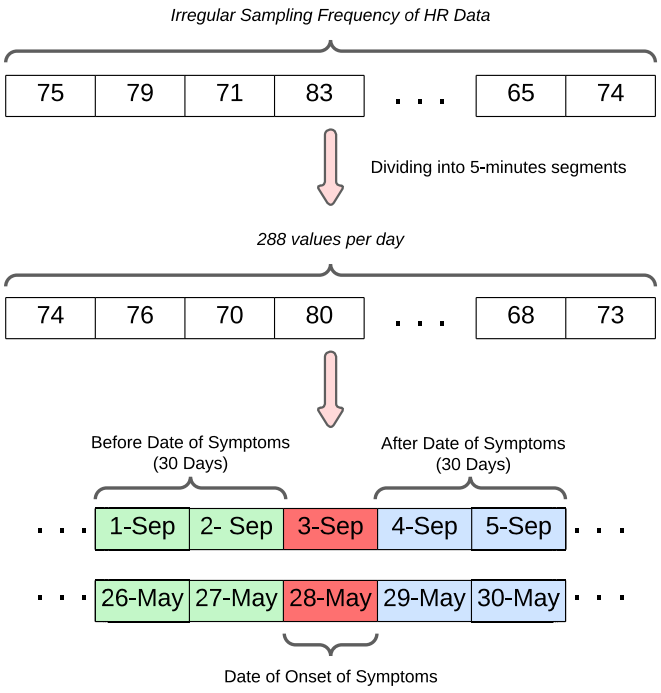

Figure 2. Division into 5-minute segments and creation of 60-day vectors.

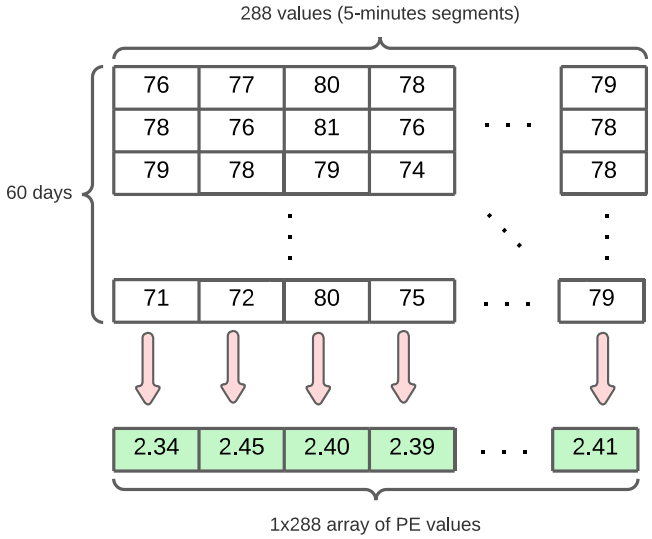

Figure 3. Obtaining PE values for each patient.

linear logistic fit due to the predominant CONTROL values in the original data.

According to the analysis carried out with the Mann-Whitney-Wilcoxon test regarding the comparison of the permutation entropy values related to each 5-minute segment present in a daily record, there were clear segments where the entropy values for each of the groups (CONTROL and COVID) was significantly different and referential for the approach of a classifying model. This fact led to the consideration of only the values related to these segments in question, allowing the model to train and classify using significant entropy values obtained from their corresponding daily segments. Those filtered examples were divided into two sets, the training and validation set and we used the random criterion presented in the Scikit-Learn library to obtain those groups. Finally, from the results of the test set we also obtain the confusion matrix that shows the correct and incorrect classification for each group and the ROC curve for the

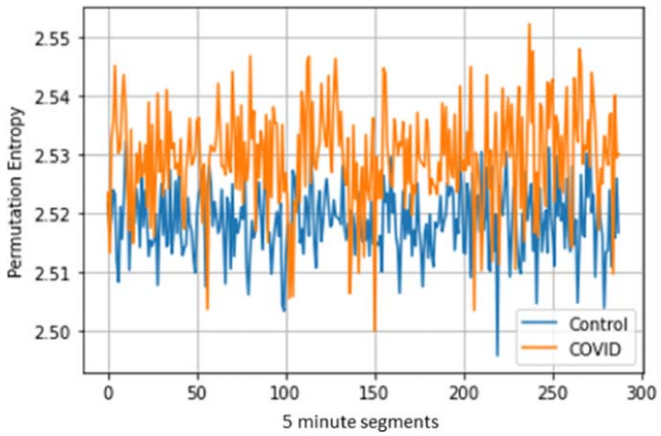

Figure 4. Circadian profiles of 24-hour PE mean values (288 frames), Control group (blue), COVID group (orange).

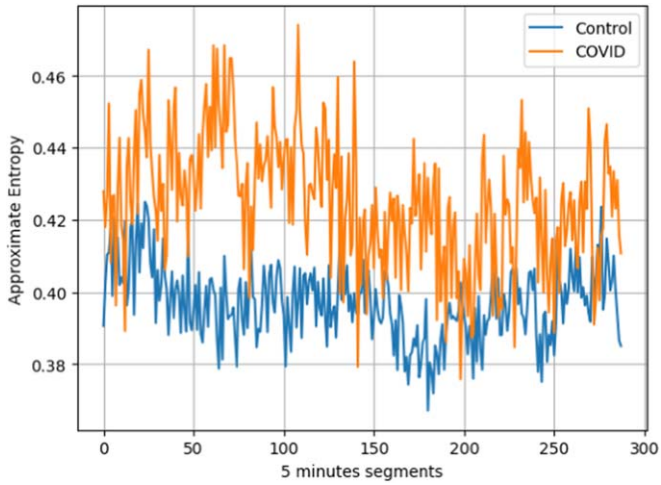

Figure 5. Circadian profiles of 24-hour ApEn mean values (288 frames), Control group (blue), COVID group (orange).

classifier from which we can obtain the respective area under the curve.

4. Results

Figures 4, 5, 6 and 7 shows the mean PE, ApEn, SE and SVDE values, respectively, for COVID patients and healthy volunteers.

In figure 4 it is possible to distinguish higher PE values in COVID patients throughout the day with certain exceptions. There are interferences between both profiles in the first 50 segments (around 12:00 am to 4:00 am), in segment 100 (approximately 8:00 am), between segments 140 to 150 (11:00 am to 12 :00 am approximately) and between segment 200 to 225 (about 4:00 pm to 7:00 pm). Figure 5 shows higher values of ApEn in COVID patients than in healthy patients. Interences exists in the first 25 frames (around 12:00 am to 2:00 am) and often between segments 175 to 288 (since around 2:00 pm to the end of the day).

The figure 6 also shows higher values for COVID patients than healthy patients. Some interferences could be seen between the frames 20 to 30 (near 1:00 am to 2:00 am), 120 to 130 (10:00 am to 11:00 am) and the last frames (around the end of the day).

The last image (figure 7), unlike the previous 3, shows higher values for healthy patients than for COVID patients. However, there are also fewer visually noticeable interferences, these being specifically at frames 170 and 270 (around 2:00 pm and 10:00 pm, respectively).

The previous observations are of a purely visual nature, for an exact determination of the differences between the PE values of both groups, the Mann-Whitney-Wilcoxon test was used. This test was applied between segments of both matrices, in order to determine in which segments there was a greater difference in values, plotting the *p*-values obtained in figure 8, considering only those with a value less than 0.05, which indicates a similarity between the samples of less than 5%, so it can be concluded that there are significant differences between them. The relevant segments refer to the daily intervals that have

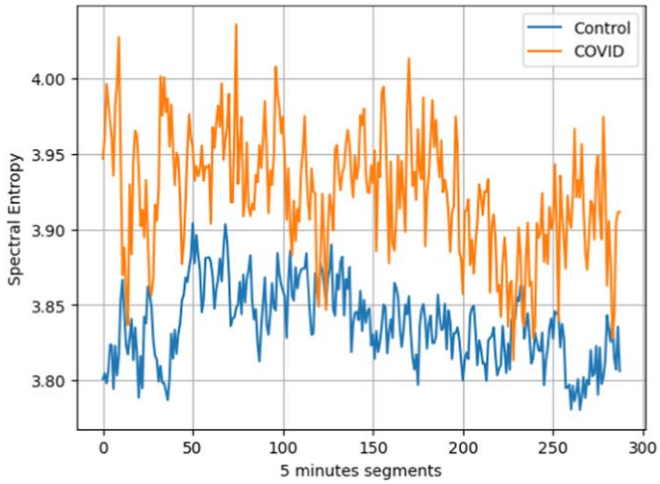

Figure 6. Circadian profiles of 24-hour Spectral-Entropy mean values (288 frames), Control group (blue), COVID group (orange).

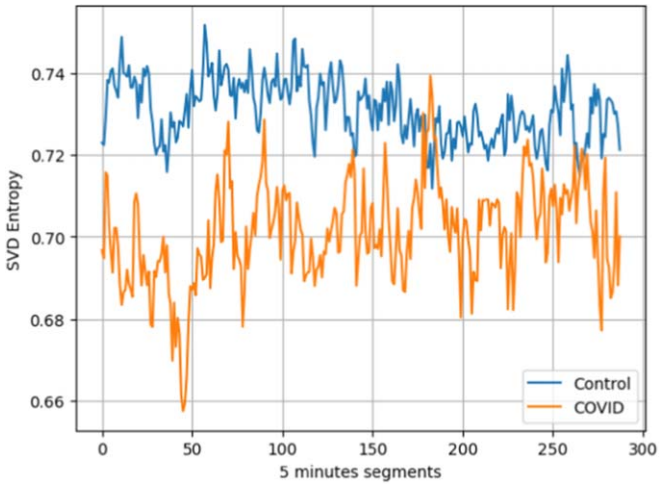

Figure 7. Circadian profiles of 24-hour SVD–Entropy mean values (288 frames), Control group (blue), COVID group (orange).

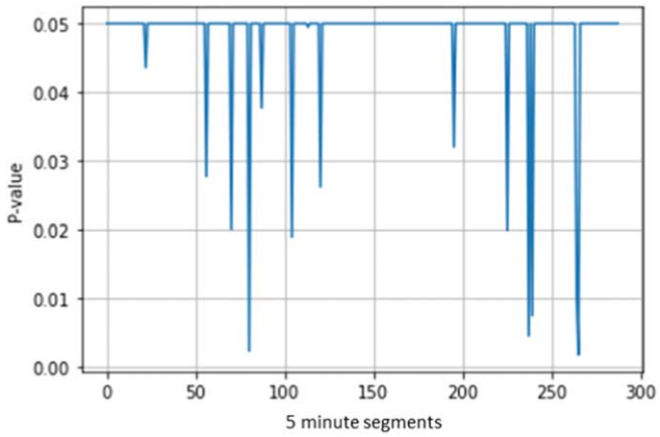

Figure 8. Mann-Whitney-Wilcoxon test of PE of the 288 frames.

significant effects on heart rate. These intervals can help improve patient monitoring and assist in the search for patterns and signs to better classify the two groups.

In figure 8 it is possible to see this p-value in the case of permutation entropy. Thus, values of the parameter p less than 0.05 occur approximately in segment

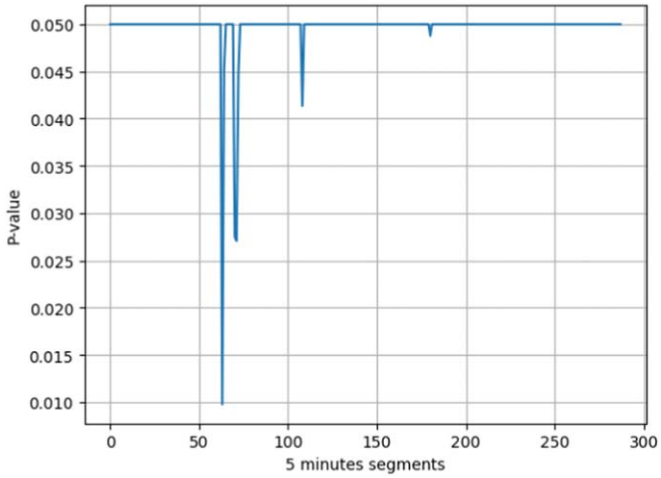

Figure 9. Mann-Whitney-Wilcoxon test of ApEn of the 288 frames.

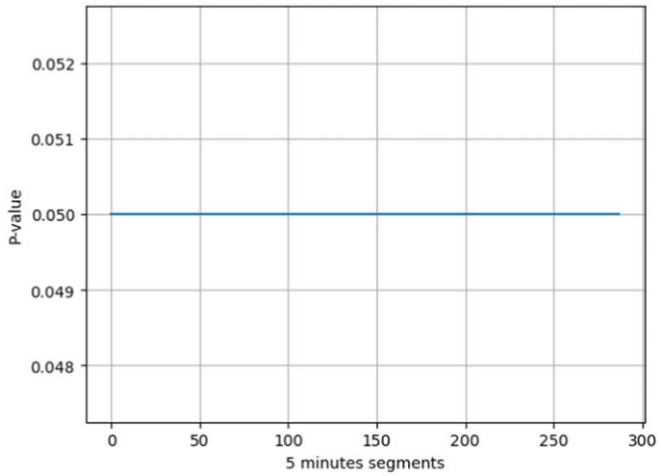

Figure 10. Mann-Whitney-Wilcoxon test of SE of the 288 frames.

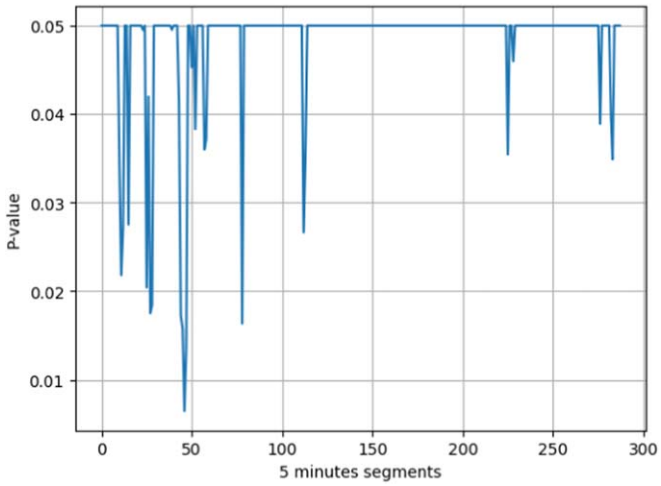

Figure 11. Mann-Whitney-Wilcoxon test of SVDE of the 288 frames.

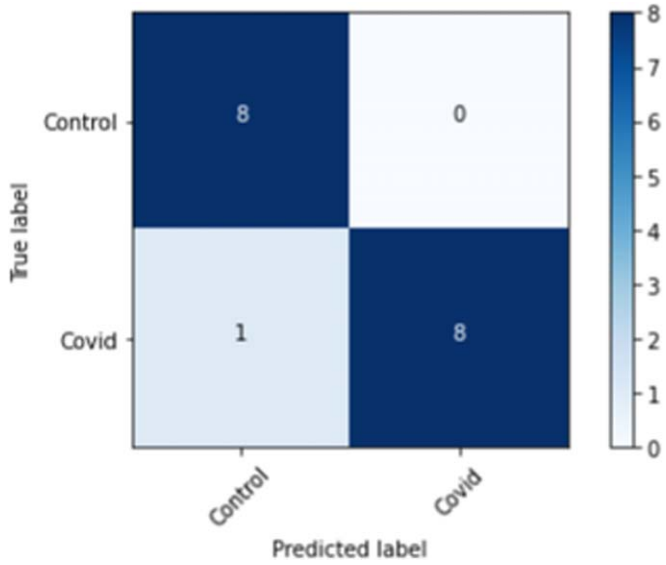

Figure 12. Confusion matrix of the logistic classifier using PE.

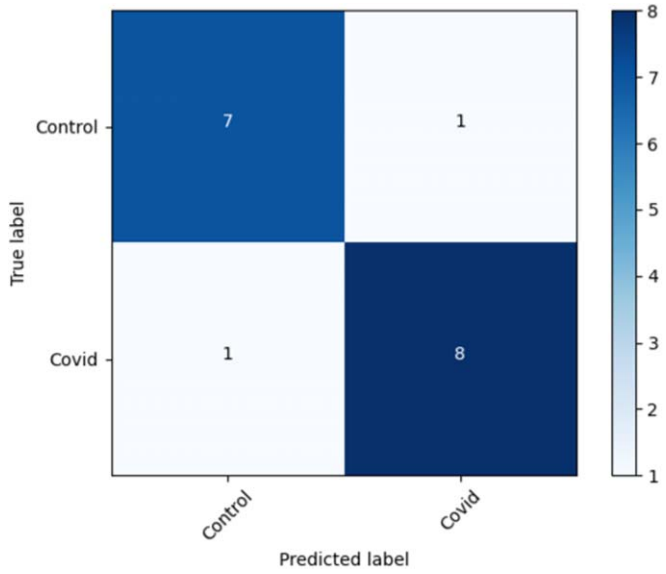

Figure 13. Confusion matrix of the logistic classifier using ApEn.

22 (around 1:00 am), between 50 and 120 (from 4:00 a. m. to 10:00 a.m.), around segment 200 (4:00 p.m.), from 230 to 240 (7:00 p.m. to 8:00 p.m.) and around the 260 segment (9:00 p.m.). Figure 9 shows the significant intervals in the case of ApEn. This intervals are approximately in segments 60 to 70 (from 5:00 am to 6:00 am), 110 (9:00 am) and 170 (2:00 pm). Figure 10 shows the p-value results for spectral entropy. It is possible to see a solid line at 0.05, that means that all comparisons obtained a p value greater than 0.05, so in this case no significant intervals could be extracted. Therefore, continuing the analysis with this entropy is ruled out.

Finally, figure 11 shows the p-value for the SVDE case. There are several significant segments there mainly from the first segments to segment 125 (from 12:00 to 10:am), near segment 225 (6:00 pm) and the last segments (near the end of the day).

Subsequently, the significant segments of each entropy were selected to develop the logistic classifier in the manner described in the method. From this, the confusion matrices that can be seen in figures 12, 13, 14 were obtained.

From the matrix in figure 12 it is possible to calculate the overall accuracy, which is 94.12%. And the precision per class, obtaining 100% when classifying healthy patients and 89% accuracy when classifying COVID patients. The recall was 89% and 100%, respectively. Finally, the F-1 score was 94% in both cases.

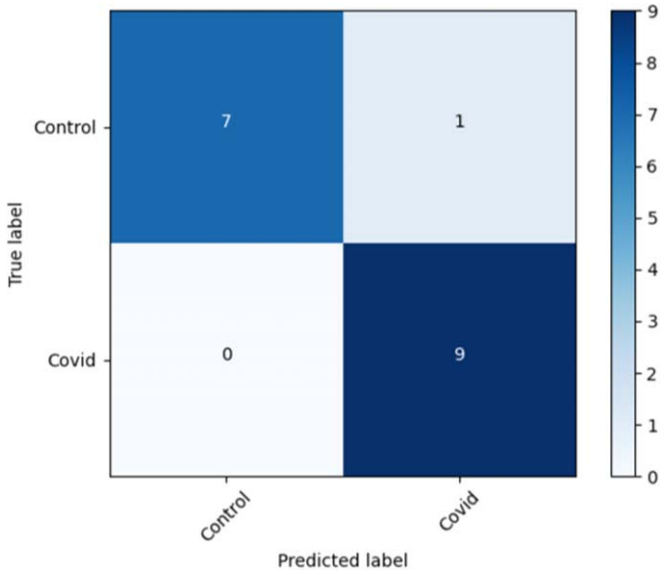

Figure 14. Confusion matrix of the logistic classifier using SVDE.

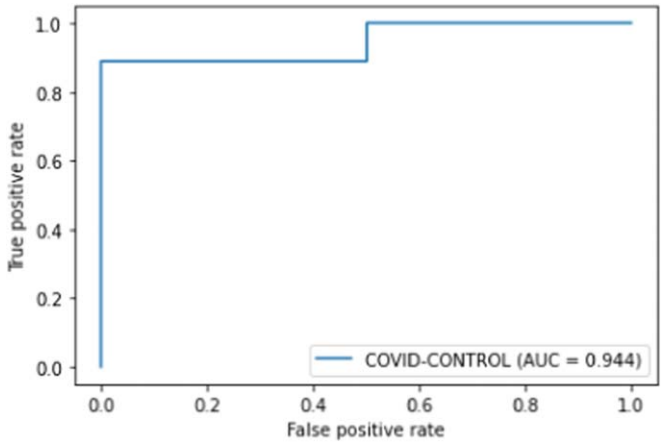

Figure 15. ROC curve for logistic classifier using PE.

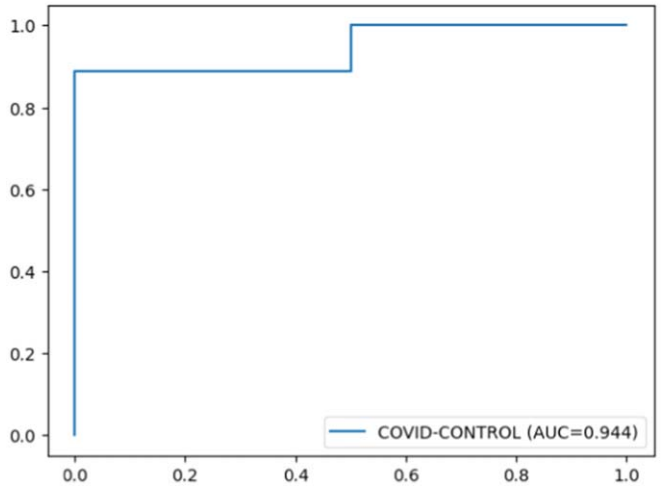

Figure 16. ROC curve for logistic classifier using ApEn.

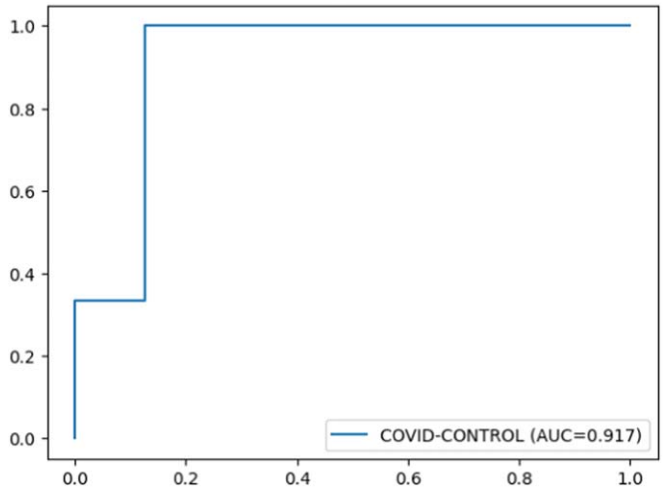

Figure 17. ROC curve for logistic classifier using SVDE.

In the case of figure 13, an overall accuracy of 88% is obtained. The precision per class is 88% when classifying healthy patients and 89% when classifying COVID patients. Regarding recall and F-1 score, the same percentages were obtained, respectively.

Regarding figure 14, a global accuracy of 94% was obtained. Precision per class of 100% when classifying healthy patients and 90% when doing so with COVID patients. The recall is 88% and 100%, respectively, and the F-1 score is 93% and 95%.

Finally, from this result it is possible to draw the ROC (Receiver Operating Characteristic Curve) curve related to the classification. In a ROC graph, the rate of false positives or 1-specificity is established on the x-axis, and the true positives or sensitivity on the y-axis, with the point in the upper left corner that minimizes false positives and maximizes true positives being an ideal point.

The ROC curves for each case are shown in figures 15, 16 and 17. In all three cases it is possible to see that the points are close to the upper left corner, indicating a good classification job, validated by the parameter area under the curve or AUC, which has a value of 0.944 for PE and ApEn and 0.917 for SVDE, all of them quite close to 1, thus indicating that the logistic classifier satisfactorily differentiates both groups of patients.

5. Discussion and conclusions

Permutation entropy, approximate entropy and SVD entropy revealed to be a powerful tool for the treatment of biological data; since, based on its application to cardiac variability data, the present study managed to characterize patients with COVID and healthy volunteers, clearly differentiating both. In the other In this way, it was observed that COVID patients present higher values of permutation and approximate entropy throughout the day compared to healthy patients. This result is comparable to other studies whose severely ill patients generally presented higher entropy values compared to sick patients but not in such an advanced state of the disease and healthy volunteers [19]. These types of results may also indicate cardiac compromise on the part of COVID patients. In the case of SVD entropy, an opposite behavior was observed, where healthy patients are observed with higher values during the day compared to COVID patients. Despite this, these 3 entropies are useful to differentiate the two groups and be able to build a classifier.

The Mann-Whitney-Wilcoxon test allowed us to isolate the most significantly different segments during the day in the case of each entropy. In this way, spectral entropy turned out to be not a good method to find significant segments between the two groups analyzed, so it was discarded. Regarding the other 3 entropies, their good work was verified through the results provided by the logistic classifier for the case of PE, ApEn and SVDE, obtaining an overall precision of 94%, 88% and 94%, respectively. ; Likewise, precision values of 89% and 100%, 88% and 89% and 100% and 90%, for the classification of COVID and Control Groups, in each case. These results on the classifier suggest that with a more sophisticated method and a larger amount of data, notable performance would be possible.

In this manner, the application of a single statistical parameter, such as permutation entropy, approximate entropy or SVDE, even their joint application, enable a robust characterization of patients. By selecting the most relevant daily segments, we achieved a highly accurate classification using a simple model, such as the logistic classifier. Thus, this work introduces a straightforward yet powerful tool with significant potential for the medical field.

COVID-19, due to its diverse range of symptoms, which can vary among individuals, necessitates invasive external tests for diagnosis. In contrast, the classifier presented here only requires data from a commonly used mobile device (smartwatch) to

determine a patient's health status, making it a convenient and accessible diagnostic tool for a large portion of the population.

The presentation of relevant day segments can have a positive influence on various research endeavors. It can be employed to enhance the current results of risk classification and stratification studies, as authors can reduce their dataset to specific segments. This allows them to eliminate portions of the day that do not provide significant information and may instead introduce bias into the training phase of the neural network [8, 9, 11]. Similarly, this is relevant for future research projects, especially those in the process of building their database, as substantial resources can be conserved by concentrating data collection efforts on specific segments of the day, rather than capturing data for the entire day.

Finally, this work has placed a strong emphasis on maintaining simplicity in the analysis and classification, achieving excellent results. Further enhancements can be attained by considering additional statistical parameters that are also valuable in the analysis of chaotic time series. Additionally, exploring more advanced classifiers, such as deep learning-based models, is a viable avenue for improvement.

Acknowledgments

The authors would like to acknowledge Stanford University for granting access to the database utilized in this study, which was solely intended for academic purposes and did not compromise the privacy of the participants and Universidad Nacional San Agustin de Arequipa.

Data availability statement

The data that support the findings of this study are openly available at the following URL/DOI http://www.nature.com/articles/s41551-020-00640-6#data-availability.

Declarations

- Conflict of interest/Competing interests The authors declare they have no financial interests.
- Ethics approval This article does not contain any studies with human or animal subjects.
- Consent to participate The database used belongs to Stanford University.
- Consent for publication The database used belongs to Stanford University.
- Availability of data and materials The database used belongs to Stanford University.
- Code availability The codes used are available.
- Authors' contributions Authors contributed equally to this work.

ORCID iDs

Miguel Vizcardo Cornejo https://orcid.org/0000-0001-9989-8946
Antonio G Ravelo-García https://orcid.org/0000-0002-8512-965X

References

[1] Rai P, Kumar B K, Deekshit V K, Karunasagar I and Karunasagar I 2021 Detection technologies and recent developments in the diagnosis of COVID-19 infection *Appl. Microbiol. Biotechnol.* **105** 441–55
[2] Hadi A G, Kadhom M, Hairunisa N, Yousif E and Mohammed S A 2020 A review on COVID-19: origin, spread, symptoms, treatment, and prevention *Biointerface Research in Applied Chemistry* **10** 7234–42
[3] Li Q, Guan X, Wu P, Wang X, Zhou L, Tong Y and Feng Z 2020 Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia *New Engl. J. Med.* **382** 1199–207
[4] Chen N, Zhou M, Dong X, Qu J, Gong F, Han Y and Zhang L 2020 Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study *The Lancet* **395** 507–13
[5] Tang B, Bragazzi N L, Li Q, Tang S, Xiao Y and Wu J 2020 An updated estimation of the risk of transmission of the novel coronavirus (2019-nCov) *Infectious Disease Modelling* **5** 248–55
[6] Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y and Cao B 2020 Clinical features of patients infected with 2019 Novel coronavirus in Wuhan, China *The lancet* **395** 497–506
[7] Pathan S, Siddalingaswamy P and Ali T 2021 Automated detection of Covid-19 from Chest x-ray scans using an optimized CNN architecture *Appl Soft Comput.* **104** 107238
[8] Li W *et al* 2020 Using machine learning of clinical data to diagnose COVID-19: a systematic review and meta-analysis *BMC Med Inform Decis Mak* **20** 247
[9] Díaz L A, Ravelo-García A, Alvarez E, Rodríguez M F, Cornejo D R, Cabrera-Caso V and Cornejo M V 2022 *Densely connected neural network and permutation entropy in the early diagnostic in COVID patients 2022 Computing in Cardiology (CinC)* 498 (IEEE) pp 1–4
[10] Gadaleta M *et al* 2021 Passive detection of COVID-19 with wearable sensors and explainable machine learning algorithms *Npj Digit. Med.* **4** 166
[11] Mishra T *et al* 2020 Pre-symptomatic detection of COVID-19 from smartwatch data *Nat. Biomed. Eng.* **4** 1208–20
[12] Bandt C and Pompe B 2002 Permutation Entropy: a natural complexity measure for time series *Phys. Rev. Lett.* **88** 174102
[13] Strydom T, Dalla R, Giulio V and Poisot T 2021 SVD entropy reveals the high complexity of ecological networks *Frontiers in Ecology and Evolution* **9** 623141
[14] Shannon C E 1948 A mathematical theory of communication *The Bell System Technical Journal* **27** 379–423
[15] Gu R and Shao Y 2016 How long the singular value decomposed entropy predicts the stock market?Evidence from the Dow Jones Industrial Average Index *Physica A: Statistical Mechanics and Its Applications* vol 453 (Elsevier) 150–61
[16] Pincus S M 1991 Approximate entropy as a measure of system complexity *Proc. Natl Acad. Sci.* **88** 2297–301
[17] Devi D, Sophia S and Prabhu S R B 2021 Deep learning-based cognitive state prediction analysis using brain wave signal *Cognitive Computing for Human-Robot Interaction* (Elsevier) pp 69–84
[18] Gomez-Pilar J, Poza J, Bachiller A, Gómez V, Molina J and Hornero R 2015 *Entropy* **17** 5241–525
[19] Cornejo D, Rodriguez M, Diaz L, Alvarez E and Vizcardo M 2020 Application of permutation entropy in the stratification of patients with chagas disease *2020 Computing in Cardiology* (IEEE) pp 1–4