Deep Learning and Permutation Entropy in the Stratification of Patients with Chagas Disease

Diego Rodrigo Cornejo[†], Antonio Ravelo-García[§], Esteban Alvarez[‡], María Fernanda Rodríguez[†], Luz Alexandra Díaz[†], Victor Cabrera-Caso[†], Dante Condori-Merma[†], Miguel Vizcardo Cornejo[†]

[†]Escuela Profesional de Física, Universidad Nacional de San Agustín de Arequipa, Perú
[§]Instituto for Technological Development and Innovation in Communications, Universidad de Las Palmas de Gran Canaria, Spain
[‡]Escuela de Física, Universidad Central de Venezuela, Venezuela

Abstract

Chagas disease is a life threatening illness that in the last decades was becoming a public health problem because of the change in the epidemiological pattern. It may be silent and asymptomatic in the chronic phase. Hence the necessity of the development of early markers. To achieve this, we propose a deep neural network architecture in order to classify 292 patients into three groups: The Control group with 83 volunteers, the CH1 group with 102 patients with positive serology and no cardiac involvement and the CH2 group with 107 patients with positive serology and incipient heart failure. The used data comes from 24-hour ECG, the RR intervals from each subject was divided in 288 frames of 5 minutes each. Then it was preprocessed using permutation entropy obtaining the circadian profile for each patient. And by applying PCA each patient ended up represented by a vector of 144 entries. This was in turn used for the training of the proposed NN architecture. The classification performed with 91% accuracy and an average of 92% precision, consisting in a great work of classification validated by the AUC in each ROC curve. As this results were obtained with a limited quantity of data, this study can be improved provided with more samples, making this model a tool for analyzing ECG in order to try to do an early evaluation and diagnosis of a cardiac compromise related to the generally silent chronic phase.

1. Introduction

There is always a need for cheaper and faster ways to detect a disease. This need becomes more relevant if the disease it is meant to detect is spreading to new regions and faster. This is the case for the Chagas disease, caused by the parasite Tripanozoma Cruzi [1], which is mainly found in countries of Latin America with 6 to 7 million people infected and dozens more at risk of infection in these areas according to the Pan American Health Organization [2]. At first Chagas disease was an endemic disease from Latin America but in recent years it was spreading to other countries like Canada, the US, and some countries in Europe [3]. This increasing rate of infection as well as the change in the epidemiological pattern have its origins in various factors apart from the vector transmission, including blood transfusion and organ donation from undiagnosed people, infection during pregnancy, travellers returning from Latin America, population mobility and so on [4, 5]. The clinical evolution of the disease begins with an acute phase that can last up to two months. Then follows a chronic phase where 40% of the patients have cardiac compromise (most commonly congestive heart failure) [6–9]. Previous studies have used heart rate variability analysis in patients with congestive heart failure. The proposed heart rate variability feature to be analyzed in this work, the permutation entropy, showed good results distinguishing patients with Chagas disease from healthy people in particular [10]. Knowing that permutation entropy is a good way to find notorious differences between these groups we propose the use of a deep neural network. This has shown to be a useful method for classifying data based on pattern recognition [11–13]. Where medical data was classified remarkably well with the multilayer perceptron (MLP), a type of neural network [14–16]. The purpose of this work is then to implement an MLP architecture capable of correctly classify and distinguish healthy people from Chagasic patients reliably based on the permutation entropy.

2. Database

The database used was provided by the Instituto de Medicina Tropical of the Universidad Central de Venezuela. It consists on the 24 hour ECG records taken from people that underwent various evaluations including Machado-Guerreiro serological test, electrocardio-

gram, echocardiogram and chest X-ray. Based on the results of these evaluations each patient was classified into one of three groups. The first one is the control group which is conformed of 83 healthy people. The CH1 group consists on 102 patients with positive serology but normal on the rest. Finally there are 107 patients with positive serology and, in addition, cardiac compromise evident from one or more than one test (with incipient heart disease, first degree AV block, sinus bradycardia or RBBB), conforming the CH2 group.

3. Method

For the preprocessing we used the Pan-Tompkins algorithm [17] in order to obtain the QRS complexes from the ECG. The data was then split in a total of 288 segments of 5 minutes each. Next, the tachogram (R-R intervals) was generated for each five minute segment so that each patient is represented by a matrix with 288 rows containing the R-R intervals. Finally the data was processed with an adaptive filter [18].

The Permutation entropy (PE) was the non linear feature selected to be the input of the neural network. This feature measures the complexity and regularity of a time series and was defined by Brand and Pompe as [19]:

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

What PE takes in consideration are all the possible permutation patterns (represented by π) of consecutive values and counts their relative apparition within the time series $p(\pi)$ and computes the entropy from that. In this case, the PE is computed for each row of R-R intervals so that each patient ends up with a vector of 288 values of PE over the day.

The libraries and modules specialized in machine learning like Tensorflow were used (and the GUI implemented within it, Keras), aside with Scikit-learn. In addition, basic modules like numpy, pandas, itertools and matplotlib were used.

As a first step, the PE data generated previously was imported alongside with the corresponding labels for each patient (Control, CH1 or CH2) which were hot encoded for future use in a neural network. A dimensionality reduction process was considered and for that the principal component analysis (PCA) was used. In this way it was possible to condense the information for each patient in shorter vectors with the advantage to conserve the majority of the information contained in the original PE vectors.

The data is then divided into two groups for training and test, made up of 80% and 20% of the total data respectively. Additionally, 25% of the training data was destined for the validation of the model.

The neural network architecture consists on an input layer of 144 neurons (product of the dimensionality reduction from the PCA), four hidden layers of 18, 14, 14 and 12 neurons respectively and the corresponding output layer with 3 neurons. With a random initialization of weights and biases.

The network was compiled using categorical crossentropy as the loss function, Adam as the optimization algorithm and categorical accuracy and the loss function itself as the evaluation metrics. During the training, a batchsize of 10 vectors per epoch was considered. And for the callbacks, early stopping along with learning rate scheduler were used with a reduction of the factor e**0.0001 that can be changed later. The number of epochs for which the network is trained is 20.

Finally, we obtain the training and validation graphs over the epochs. From the results on the test set we also get the confusion matrix showing the correct and incorrect classification for each group and the ROC curve for the classifier from which we can obtain the respective area under the curve.

4. Results

The first results obtained were the graphs of the loss function and categorical accuracy shown in figures 1 and 2 over 20 epochs. The loss function shows a decreasing behaviour in both train and validation sets. As the loss function serves to indicate the error made by the neural network this means that it is improving over the epochs. Although both curves are following the same trend, the two separate from each other.

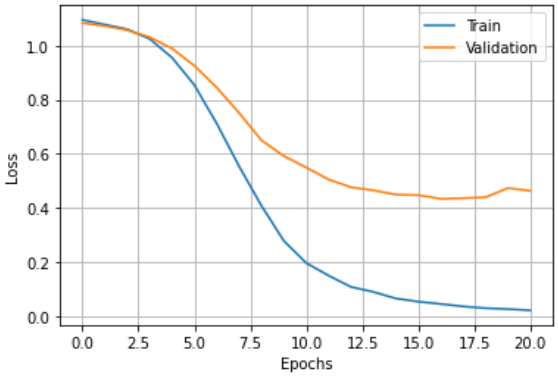

Figure 1. Loss function trough the epochs

The same can be said about the categorical accuracy, where both curves are increasing (which again, is a sign of the improvement of the network) but separating from each other, though not as much as in the case of the loss function. Both of these curves reach acceptable values (around 0.8 for the validation set and around 1 for the training set

by the end of the training of the neural network)

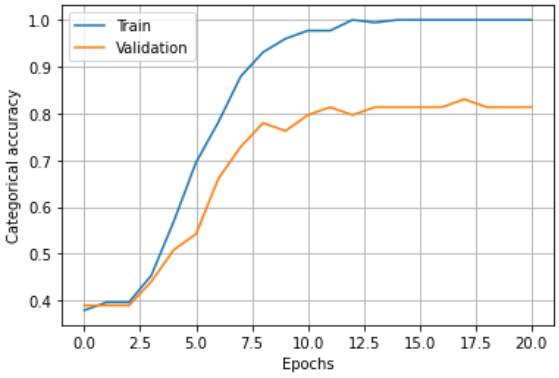

Figure 2. Categorical accuracy trough the epochs

The results for the classification on the test set is contained in the confusion matrix (figure 3). It can be seen from the matrix that there was an overall good classification, only with a maximum of two subjects classified incorrectly for each class. For instance, no subject belonging to CONTROL group was misclassified. Quantitatively the precision and recall can be computed for each group from the confusion matrix, and from these two metric the F1-score can also be computed. CONTROL group got 0.88 of precision, 1 of recall and thus the F1-score is 0.93. In the case of CH1 group the precision is 0.93, the recall is 0.87 and the F1-score is 0.90. Finally for the CH2 group the precision, recall and F1-score are 0.95, 0.86 and 0.90 respectively.

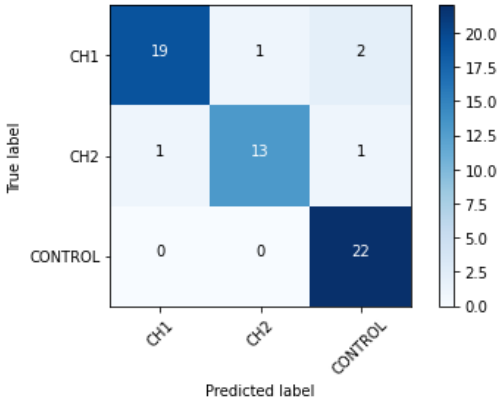

Figure 3. Confusion matrix for the test group

For further analysis, the different ROC curves are plotted in figure 4 where we used a one-vs-all approach in the curves referencing individual groups. In an ROC plot the x-axis is the false positive rate or 1-specificity, a quantity that we want to minimize. The true positive rate or sensitivity is plotted in the y-axis which we want to maximize. So the more the ROC curve approaches the top left corner of the graph the better the classification is. There is also a referential diagonal line that behaves like a classifier that labels each individual randomly. The good classification of the neural network can be clearly seen from all the curves in the figure.

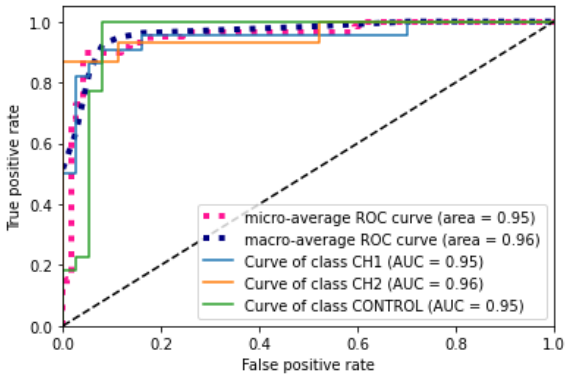

Figure 4. Individual, macro and micro average ROC curves

One last metric worth analysing from figure 4 is the AUC wich stands for area under the curve. As (0,1) is the optimal point in an ROC, 1 is the maximum AUC value as well as the best for a classifier. In the figure each ROC has its own AUC, and all values for AUC vary between 0.95 and 0.96 which is great and a sign of good classification work.

5. Discussion and conclusions

The neural network used with the specific proposed architecture performed good as the results are mainly positive in the sense that the metrics chosen (precision, recall, F1 score for each group, as well as the accuracy of the model and the AUC from all the distinct ROC) have all high values.

However, it is not perfect and one of the main problems one can see from this classifier is the fact that the curves for training and validation are not always close for both loss function and categorical accuracy. This suggests that there is a problem of overfitting that has to be revised. This problem can be overcome in multiple ways. For example by selecting different proportions in the way the data was split for the training of the network. Also the addition of new data is always helpful. And one more thing we can do is to combine the PE values with other significant features like ApEn.

Considering the limitations, this neural network did well and can be improved as discussed. Hence constituting a promising way to diagnose and stratify patients with Chagas disease and even expand its use to other important con-

ditions related to the heart.

Acknowledgements

Universidad Nacional de San Agustin de Arequipa

References

[1] Briceño-León R.(2009). La enfermedad de Chagas en las Américas: una perspectiva de ecosalud. *Cadernos de saúde pública, 25*,S71-S82.

[2] Pan American Health Organization(2017). *Enfermedad de Chagas en las Américas. Hoja informativa para los trabajadores de salud*.

[3] World Health Organization (2010). First WHO report on neglected tropical diseases: working to overcome the global impact of neglected tropical diseases. In *First WHO report on neglected tropical diseases: Working to overcome the global impact of neglected tropical diseases* (p.172)

[4] Schmunis, G. A. (2007). Epidemiology of Chagas disease in non endemic countries: the role of international migration. *Memorias do Instituto Oswaldo Cruz*, 102 ´ , 75-86

[5] *Chagas disease control and prevention in Europe. Report of a WHO Informal Consultation (jointly organized by WHO headquarters and the WHO Regional Office for Europe)*. Geneva, Switzerland, 17–18 December 2009. Geneva, World

[6] Mendoza, I., Moleiro, F., Marques, J., & Britto, I. M. (2008). Arritmias y muerte súbita en la enfermedad de Chagas. *Publ. del Instituto de Medic. Tropical.*

[7] Moleiro, F. (1980). Miocardiopatía crónica chagásica: Un estudio epidemiológico utilizando métodos electrofiosiológicos de exploración clínica

[8] Hagar, J. M., & Rahimtoola, S. H. (1991). Chagas' heart disease in the United States. *New England Journal of Medicine*, 325(11), 763-768.

[9] Di Lorenzo Oliveira, C., Nunes, M. C. P., Colosimo, E. A., de Lima, E. M., Cardoso, C. S., Ferreira, A. M., ... & Ribeiro, A. L. P. (2020). Risk Score for Predicting 2-Year Mortality in Patients With Chagas Cardiomyopathy From Endemic Areas: SaMi-Trop Cohort Study. *Journal of the American Heart Association, 9*(6), e014176.

[10] Cornejo, D. & Rodríguez, M. & Díaz, L. & Alvarez, E. & Vizcardo, M. (2020,September). Application of Permutation Entropy in the stratification of patients with chagas disease. In *Computing in Cardiology* (pp. 1-4). IEEE.

[11] Oh, S. L., Hagiwara, Y., Raghavendra, U., Yuvaraj, R., Arunkumar, N., Murugappan, M., & Acharya, U. R. (2018). A deep learning approach for Parkinson's disease diagnosis from EEG signals. *Neural Computing and Applications*, 1-7.

[12] Driss, S. B., Soua, M., Kachouri, R., & Akil, M. (2017, May). A comparison study between MLP and convolutional neural network models for character recognition. In *Real-Time Image and Video Processing 2017* (Vol. 10223, p. 1022306). International Society for Optics and Photonics.

[13] Bogu, Gireesh & Snyder, Michael. (2021). Deep learning-based detection of COVID-19 using wearables data. *medRxiv* 10.1101/2021.01.08.21249474.

[14] E. Haselsteiner & G. Pfurtscheller, Using time-dependent neural networks for EEG classification. *IEEE Transactions on Rehabilitation Engineering, 8*(4), pp. 457-463. doi: 10.1109/86.895948.

[15] Chatterjee, R., & Bandyopadhyay, T. (2016, January). EEG based Motor Imagery Classification using SVM and MLP. In *2016 2nd International Conference on Computational Intelligence and Networks (CINE)* (pp. 84-89). IEEE.

[16] Raad, A., Kalakech, A., & Ayache, M. (2012). Breast cancer classification using neural network approach: MLP and RBF. *networks, 7*(8), 9.

[17] Pan, J., & Tompkins, W. J. (1985). A Real-Time QRS Detection Algorithm. IEEE Transactions on Biomedical Engineering, BME-32(3), 230–236.

[18] Wessel, N., Voss, A., Kurths, J., Saparin, P., Witt, A., Kleiner, H. J., & Dietz, R. (1994, September). Renormalised entropy: a new method of non-linear dynamics for the analysis of heart rate variability. In *Computers in Cardiology 1994* (pp. 137-140). IEEE.

[19] Bandt, Christoph & Pompe, Bernd. (2002). Permutation Entropy: A natural complexity measure for time series. *Physical review letters, 88*(17), 174102.

[20] Akaike, H. (1969). Fitting Autoregressive for Prediction Models. *Statist Math, 21*, 243-247.

Correspondence

Miguel Vizcardo Cornejo
mvizcardoc@unsa.edu.pe
Calle Santa Catalina Nº 117 CP 04000. Arequipa Perú.