Transforming ECG into a Poincaré Plot-Based Image for Its Quality Assessment Through a Pre-trained Convolutional Neural Network

Álvaro Huerta[1], Óscar Ayo-Martín[2], José J Rieta[3], Raúl Alcaraz[1]

[1] Research Group in Electronic, Biomedical and Telecom. Eng., Univ. of Castilla-La Mancha, Spain
[2] Department of Neurology, University Hospital of Albacete, Spain
[3] BioMIT.org, Electronic Engineering Department, Universitat Politecnica de Valencia, Spain

Abstract

Wearable ECG acquisition systems present the potential to continuously monitor cardiac activity during patient's daily life. However, the acquired ECG signals are often disturbed by noise and their automated quality assessment (QA) is crucial. For this purpose, pre-trained 2-D convolutional neural networks (CNNs) have recently reported promising performance. Usually, linear time-frequency transformations have been only considered to turn the ECG signal into a 2-D image. Hence, this work aims to explore whether ECG transformation using a common approach specifically designed to visualize and feature nonlinear, complex, and chaotic dynamics in a time series, such as the Poincaré plot (PP), might be helpful in QA of the signal. Then, the pre-trained AlexNet model was inputted with the image obtained by the PP-based representation of the ECG. For comparison, this network was also fed with the common wavelet ECG-based scalograms. After processes of training and testing on separate datasets, no statistically significant differences in global terms of balanced accuracy and F_1*-score were noticed for the models obtained with the two ECG-based images. However, the PP-based model was about 3 times faster than that fed with scalograms. Thus, it could be concluded that the ECG transformation into a PP-based image might be a quick and useful approach for CNN-based QA of the signal.*

1. Introduction

The resting ECG is still the most widely used test for the diagnosis of many cardiovascular diseases [1]. However, this recording is extremely short lasting for about 10–20 seconds, and hence is not suitable for identification of intermittent cardiac diseases, such as paroxysmal atrial fibrillation (AF). Although conventional ambulatory systems are able to acquire the ECG signal with several days in length, it has been proven that the longer the duration of cardiac monitoring, the greater the possibility of detecting initial, brief episodes of the arrhythmia [2]. In this context, some emerging wearable systems with the ability to monitor ECG continuously for long weeks or months without disrupting the patient's daily life represent a great opportunity to improve early diagnosis of occult AF and restrict the current exponential increase of its prevalence [3].

However, integration of these devices into clinical practice has been limited to date, since the acquired ECG signal is often strongly contaminated by transient noise and environmental perturbations, leading to numerous false alarms of AF and requiring manual interpretation of extensive amounts of ECG data [4]. To overcome this problem, automated selection of ECG excerpts with sufficient quality for precise diagnosis has been widely proposed, and numerous algorithms for such ECG quality assessment can be found. Most are based on combining hand-crafted features derived from the ECG, such as mean RR interval, time consistency of PQRST waveforms, or coherence of QRS complexes, among others, via common machine learning classifiers [5]. These algorithms have reported successful performance on ECG signals acquired from healthy subjects, but their performance is notably reduced between 15% and 40% when ECG recordings from paroxysmal AF are analyzed [5]. This last case is a more challenging scenario because fibrillatory waves present waveforms and time–frequency features similar to the most common transient noise found during the ECG acquisition [5].

Contrarily, the recent analysis of pre-trained 2-D convolutional neural network (CNN) architectures has shown a promising ability to evaluate ECG quality even in the context of paroxysmal AF [6]. As in many other previous applications, these networks were fed with the ECG transformation into a wavelet scalogram using Continuous Wavelet Transform (CWT) [6]. This time-frequency transformation has proven to be effective to represent non-stationary and time-varying physiological signals and noises [7], but it only makes use of concepts of linear analysis. Hence, this work aims to explore whether ECG transformation using a common approach specifically designed to visualize

 ISSN: 2325-887X DOI: 10.22489/CinC.2024.374

and feature nonlinear, complex, and chaotic dynamics in a time series, such as the Poincaré Plot (PP), might be helpful in quality assessment of the signal.

2. Materials and methods

2.1. Databases

To obtain realistic and non-inflated classification results, two separate databases were used for training and testing of the generated models. Moreover, it should be noted that the two databases were acquired with different wearable recording systems and electrode configurations, thus involving a more challenging framework for a strict evaluation of the generalization ability of the algorithms.

On the one hand, a proprietary database was used for training. In this case, ECG recordings were acquired using a textile wearable Holter placed on the thorax, with a sampling frequency of 250 Hz and a resolution of 12 bits over a dynamic range of ±5 mV. The detection of paroxysmal AF episodes in these ECG signals was automatically addressed [8], but under visual supervision of two expert physicians. Additionally, they also manually categorized the ECG segments into noisy or clean based on their ability to detect R-peaks. Segments where R-peaks were clearly identifiable were labeled as high-quality, while the remaining segments were considered low-quality. The reviewers also identified arrhythmias other than AF, as well as premature ventricular and atrial contractions, within the high-quality ECG intervals, labeling them as other rhythms (OR). From the labeled segments, 10,000 noisy and 10,000 high-quality 5-second ECG excerpts were randomly selected to create our final database.

On the other hand, the public PhysioNet/CinC Challenge 2017 database [9] was used for testing. This comprises 8,528 ECG recordings, ranging in duration from 30 to 60 seconds. They were obtained using a portable device connected to a smartphone, which operated with a sampling frequency of 300 Hz and 16-bit resolution. Annotations from several experts related to four classes, i.e., sinus rhythm, AF, OR and noisy excerpts, are also provided. Subsequently, the recordings were segmented into 5-second excerpts, resulting in 47,439 high-quality and 1,168 low-quality ECG intervals.

2.2. PP-based representation of the ECG

Before transforming the ECG signal to a PP, it was preprocessed. Thus, since different sampling rates were used to acquire the signals in each dataset, if needed, the ECG excerpt was firstly resampled at 250 Hz. Next, it was normalized through a min-max scaling and its values were restricted to the range -1 and 1. Then, the well-known Takens' delay embedding theorem was used to represent the ECG into a PP using an embedded dimension m of 2 samples and a time delay of 1 sample. In brief, this approach is based on plotting the original time series versus its time-delayed copies [10] and, for every normalized ECG segment $x(n)$, these time-lagged vectors were obtained as

$$\mathbf{X(i)} = \{\mathbf{x(i)}, \mathbf{x(i+1)}, \mathbf{x(i+2)}, \ldots, \mathbf{x(i+m-1)}\}, \quad (1)$$

i ranging from 1 to $N-(m-1)$, and N being the length of the ECG segment. Hence, the PP-based representation of the ECG signal was obtained by plotting $\mathbf{X(i)}$ versus $\mathbf{X(i+1)}$. Finally, to convert this plot into an image, it was divided into a grid of 227 $\times$ 227 cells, and a Jet colormap with 128 colors was applied to the range of values obtained by computing the logarithm of the number of points in each grid cell. As an example, Figure 1(a), shows the PP-based representation of a typical 5 second-length high-quality ECG excerpt, whereas Figure 1(c) depicts a PP-based graph of a common 5 second-length low-quality ECG segment.

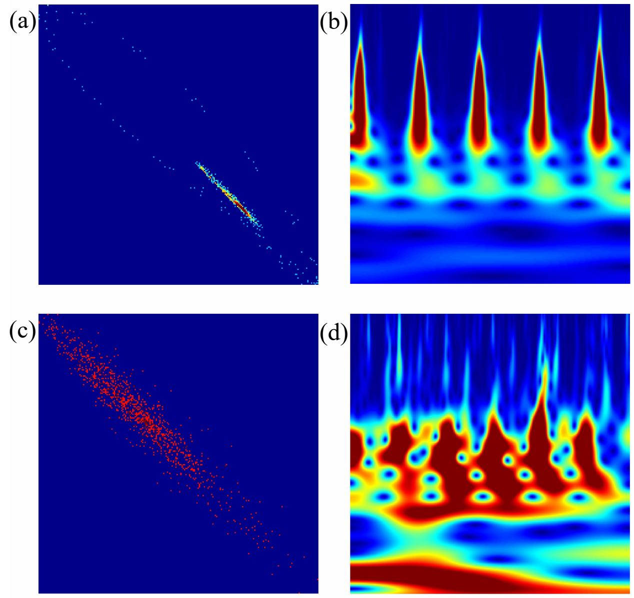

Figure 1. Two typical 5 second-length ECG segments, on the top, the ECG excerpt belonging to a high-quality class converted to (a) PP-based and (b) CWT-based images, on the bottom, the ECG interval belonging to a low-quality class, (c) PP-based and (d) CWT-based images.

2.3. CWT-based image of the ECG

To serve as a reference, the ECG signal was also converted into a 2-D image using CWT. This transform has been widely used in previous works dealing with ECG quality assessment because it is able to jointly exploit information from time and frequency domains, thus achieving deep abstractions of the high- and low-quality ECG excerpts [6]. Shortly, CWT decomposes the ECG signal into several time scales, each one representing a certain frequency range in the time-frequency plane. This results in a 2-D matrix containing a detailed map of wavelet coefficients whose location is based on their size and position. Applying a colormap to this matrix, an image known as a wavelet scalogram is obtained. To be inputted to AlexNet and comparable to the previous PP-based representation of the ECG, the scalogram was resized to 227×227 points. Figure1(b), shows the same high-quality ECG excerpt as before transformed into a CWT-based image, as well as Figure 1(d) depicts the same low-quality ECG interval as before transformed into a CWT-based image.

2.4. Experimental setup

AlexNet is a very popular CNN architecture, which has been used in a broad variety of applications [11]. It is composed of eight layers with the ability to learn, where five are convolutional and three are fully-connected. Moreover, two pooling layers are also included in the model to reduce the spatial dimension of the feature map, while the most relevant information is preserved. Based on this network, two models were generated for the classification of high- and low-quality ECG segments, one inputted with the PP-based images and another with the CWT-based ones.

The transfer learning paradigm was used in both cases by initializing all the layers of AlexNet with the knowledge learned from ImageNet [12]. Any layer was frozen during the fine-tuning with the proprietary database previously described, but the last fully-connected layer was modified for a two-class classification. For this process, a batch size of 512 images, 100 epochs of training, and a stochastic gradient descent algorithm with a constant learning rate of 0.01 were selected. Nonetheless, to minimize overfitting, an early stopping approach was considered. Indeed, the training process was stopped when the validation accuracy did not improve for 5 consecutive epochs.

To quantify the performance of both models on the testing dataset, common metrics like sensitivity (Se) and specificity (Sp) were computed. Nonetheless, given that the PhysioNet/CinC Challenge 2017 database presented a strong unbalanced number of high- and low-quality ECG excerpts, alternative indices have also been proposed to provide more truthful scores, including Balanced Accuracy ($BAcc$) and F_1-score. Moreover, the average time required by the two models to classify a 5 second-length ECG sample in the testing stage was also measured as an estimation of computational cost.

3. Results

Table 1 summarizes the main outcomes reported by the two performed classification models. As can be seen, no larger differences were noticed between the performance metrics, since values of Se, Sp, and $BAcc$ approximately ranged between 80% and 90%, and values F_1-score were greater than 91%. Nonetheless, whereas a McNemar test reported no statistically significant differences in global terms of $BAcc$ and $F1$-score for both models, a statistical significance lower than 0.05 was obtained both between values of Se and values of Sp. Precisely, the model inputted with PP-based images presented significantly larger values of Sp, and the another one (fed with CWT-based images) notably larger values of Se. Regarding the average time spent by the two algorithms for classification of every testing 5 second-length ECG segment, the first model was about 3 times faster than the second one, requiring 14 ms per image versus 42 ms per image.

4. Discussion

As Figure 1 shows for typical examples of high- and low-quality excerpts, the two ECG transformations generate very different 2-D distributions of points. In general terms, whereas the CWT-based image presents the information in a more evenly distributed and uniform manner throughout the plot, the PP-based image concentrates most points around the diagonal of the graph. Nonetheless, the global classification performance of AlexNet in both cases was very similar, with values of $BAcc$ and F_1-score differing by about 1% and 3.5%, respectively. Interestingly, although larger differences were noticed by both models between values of Se and values of Sp, their performance was complementary. Indeed, the PP-based images were able to generalize better for the detection of low-quality ECG excerpts, while the CWT-based images for identification of high-quality ones. Hence, a potential combination of both classifications should be explored in search of better quality assessment of wearable, single-lead ECG recordings.

Regarding computational load, the model inputted with PP-based images was notably faster than the other one. Since the same CNN architecture was used in both cases, the fact that PP-based reconstruction of the ECG does not require mathematical computation could explain this result. Additionally, in a recent work using the same validation framework [13], this ECG transformation has also reported a much lower computation time (between 8 and 10 times lower) when a lightweight CNN, with less than 160,000 parameters and trained from scratch, was used instead of AlexNet. In this case, the classification of high- and low-quality ECG excerpts on the testing dataset still remained very similar, with values of Se, Sp, $BAcc$,

Table 1. Performance metrics reported by the two generated models based on AlexNet.

ECG image	Se	Sp	BAcc	F_1-score	Time/ECG interval (ms)
PP-based	0.839	0.897	0.868	0.911	14
CWT-based	0.904	0.810	0.857	0.947	42

and F_1-score about 85%, 83%, 85%, and 91%, respectively [13]. Hence, these findings suggest that this PP-based representation of the ECG might also be useful in developing new methods for quality assessment of the signal based on tiny CNN architectures and minimal training.

5. Conclusions

The utilization of a pre-trained CNN scheme fed with ECG transformation into a PP-based image has proven to be a quick and useful approach for quality assessment of the signal. Moreover, this ECG representation seems to provide different and complementary information to the common CWT-based scalogram, so that combination of the classifications provided by both approaches could improve quality assessment of wearable, single-lead ECG recordings acquired from paroxysmal AF. Nonetheless, this will have to be explored in the future.

Acknowledgments

This research has received financial support from the company Daiichi Sankyo S.A.U. and from public grants PID2021-128525OB-I00, PID2021-123804OB-I00, and TED2021-130935B-I00 funded by MICIU/AEI/10.13039/501100011033, FEDER EU, and NextGenerationEU /PRTR, grant SBPLY/21/180501/000186 funded by Junta de Comunidades de Castilla-La Mancha, and grand AICO/2021/ 286 funded by Generalitat Valenciana.

References

[1] Niu J, Deng C, Zheng R, Xu M, Lu J, et. al. The association and predictive ability of ECG abnormalities with cardiovascular diseases: A prospective analysis. Glob Heart Sep 2020;15(1):59.

[2] Kalarus Z, Mairesse GH, Sokal A, Boriani G, Średniawa B, et. al. Searching for atrial fibrillation: looking harder, looking longer, and in increasingly sophisticated ways. An EHRA position paper. Europace Feb 2023;25(1):185–198.

[3] Lippi G, Sanchis-Gomar F, Cervellin G. Global epidemiology of atrial fibrillation: An increasing epidemic and public health challenge. Int J Stroke Feb 2021;16(2):217–221.

[4] Ding EY, Marcus GM, McManus DD. Emerging technologies for identifying atrial fibrillation. Circ Res Jun 2020; 127(1):128–142.

[5] Satija U, Ramkumar B, Manikandan MS. An automated ECG signal quality assessment method for unsupervised diagnostic systems. Biocybern Biomed Eng 2018;38:54–70.

[6] Huerta A, Martinez A, Carneiro D, Bertomeu-Gonzalez V, Rieta JJ, Alcaraz R. Comparison of supervised learning algorithms for quality assessment of wearable electrocardiograms with paroxysmal atrial fibrillation. IEEE Access 2023;11:106126–106140.

[7] Pradhan BK, Neelappu BC, Sivaraman J, Kim D, Pal K. A review on the applications of time-frequency methods in ECG analysis. J Healthc Eng 2023;2023(1):3145483.

[8] Ródenas J, García M, Alcaraz R, Rieta JJ. Combined nonlinear analysis of atrial and ventricular series for automated screening of atrial fibrillation. Complexity 2017; 2017(1):2163610.

[9] Clifford GD, Liu C, Moody B, Lehman LWH, Silva I, Li Q, Johnson AE, Mark RG. AF classification from a short single lead ECG recording: The PhysioNet/Computing in Cardiology Challenge 2017. Comput Cardiol 2017;44.

[10] de Pedro-Carracedo J, Fuentes-Jimenez D, Ugena AM, Gonzalez-Marcos AP. Phase space reconstruction from a biological time series: A photoplethysmographic signal case study. App Sci 2020;10(4).

[11] Patil T, Pandey S, Visrani K. A review on basic deep learning technologies and applications. Data Science and Intelligent Applications Proceedings of ICDSIA 2020 2020; 565–573.

[12] Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. In Advances in Neural Information Processing Systems. 2012; 1097–1105.

[13] Huerta Á, Martinez-Rodrigo A, Bertomeu-González V, Ayo-Martin Ó, Rieta JJ, Alcaraz R. Single-lead electrocardiogram quality assessment in the context of paroxysmal atrial fibrillation through phase space plots. Biomed Sign Proc Control 2024;91:105920.

Address for correspondence:

Álvaro Huerta Herraiz
ITAV, Campus Universitario s/n, 16071, Cuenca, Spain
Phone: +34-969-179-100
e-mail: alvaro.huerta@uclm.es