Computers in Biology and Medicine 45 (2014) 1–7

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: www.elsevier.com/locate/cbm

# Application of machine learning techniques to analyse the effects of physical exercise in ventricular fibrillation

Juan Caravaca [a,*], Emilio Soria-Olivas [b], Manuel Bataller [a], Antonio J. Serrano [b], Luis Such-Miquel [c], Joan Vila-Francés [b], Juan F. Guerrero [a]

[a] Digital Signal Processing Group, ETSE, Universitat de València, Avda Universitat S/N, Burjassot, València 46100, Spain
[b] Intelligent Data Analysis Laboratory, ETSE, Universitat de València, Avda Universitat S/N, Burjassot, València 46100, Spain
[c] Department of Physiotherapy, Universitat de València and INCLIVA, València, Spain

ARTICLE INFO

*Article history:*
Received 13 February 2013
Accepted 18 November 2013

*Keywords:*
Machine learning
Classification
Knowledge extraction
Logistic regression
Multilayer perceptron
Extreme learning machine
Ventricular fibrillation
Physical exercise

ABSTRACT

This work presents the application of machine learning techniques to analyse the influence of physical exercise in the physiological properties of the heart, during ventricular fibrillation. To this end, different kinds of classifiers (linear and neural models) are used to classify between trained and sedentary rabbit hearts. The use of those classifiers in combination with a wrapper feature selection algorithm allows to extract knowledge about the most relevant features in the problem. The obtained results show that neural models outperform linear classifiers (better performance indices and a better dimensionality reduction). The most relevant features to describe the benefits of physical exercise are those related to myocardial heterogeneity, mean activation rate and activation complexity.



## 1. Introduction

Several authors have proposed that physical exercise (PE) modifies the sympathetic-vagal balance of autonomic nervous system, producing an increase of parasympathetic activity that manifests in a decrease of cardiac frequency [1]. Besides, this vagal activity modification could have protective effects against the appearance of cardiac arrhythmias and death [2]. Other effects of PE are based on changes in the physiological properties of the hearth, and are called intrinsic modifications. The effects of such intrinsic modifications caused by PE have already been reported [3].

Other previous studies have shown that physical exercise modifies ventricular fibrillation (VF) response by intrinsic mechanisms [4]. Those modifications can be found in several parameters derived from frequency and time domains [4], and its spatial distributions [5]. There is a high amount of parameters to describe VF signals, and not all of them might be appropriate to describe the effects of PE in the same way [6].

The goal of this work is to find the most suitable parameters to explain the intrinsic modifications produced in VF by PE. Recordings acquired from two groups of isolated rabbit hearts (trained with PE and untrained) were analysed using a wrapper feature selection algorithm [7].

Finding the best features to describe the benefits of PE involves two main advantages. On one hand, it helps to understand the underlying mechanisms that PE may activate, and thus suggesting which VF features are modified. Those features can be transferred to the heart physiological characteristics. On the other hand, it allows a better analysis of these intrinsic modifications, improving future works by the use of the most relevant features.

Feature selection is a widespread application field of machine learning. There are many applications of feature selection in different areas of biomedical engineering [8]. Regarding applications of feature selection methods in VF analysis, the most common applications are related to arrhythmia discrimination [9], and classification of ECG signals [10]. This paper uses a wrapper feature selection algorithm based on the analysis of the performance of some classifiers [7]. This feature selection will identify the best features to classify between two groups of VF records and, therefore, it will find the most relevant features to characterise the differences between both groups.

Next section presents data acquisition and explains the features used with the classifiers. Afterwards, the following classifiers are explained: logistic regression (classical method in statistics) [11]; multilayer perceptron (the most extended neural model) [12] and,

* Correspondence to: Avenida de la Universidad S/N, 46100, Burjassot, Valencia, Spain. Tel./fax: +34 963544143.
*E-mail addresses:* caravaca.juan@gmail.com, juan.caravaca@uv.es (J. Caravaca).

0010-4825/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.compbiomed.2013.11.008

finally, the extreme learning machine [13] (a recently proposed algorithm to find the parameters in multilayer perceptron with one hidden layer). Section 3 shows the obtained results and finally conclusions are shown in Section 4.

## 2. Methods

The proposed study consists in four stages: data acquisition, data processing, classification and knowledge extraction. Electrograms were acquired from two groups of rabbit hearts. Next, these electrograms were processed measuring four parameters from time and frequency domains. Using these parameters, 18 features were calculated and used to perform a classification between the groups in the experiment. Finally, these classifiers were analysed with a wrapper feature selection algorithm to perform knowledge extraction, analysing the relevance of the different features and performing subset selections. Fig. 1 shows a diagram of the proposed study.

Next subsection, data acquisition, explains the first stage of the study. Further subsections explain data processing, classification and knowledge extraction stages, respectively.

### 2.1. Data acquisition

Twenty-one male New Zealand white rabbits (*Oryctolagus cuniculus*) were used in the present study. Animals were divided into two experimental groups: an untrained group (G1, with a sample size of 10) and a trained group (G2, with a sample size of 11). Animals in the untrained group were housed in the animal quarters for 46 days, and rabbits in the trained group were submitted to a physical exercise program. After familiarization with treadmill running for four days, animals in the trained group ran five days/week for 6 weeks at 0.33 m $s^{-1}$. Each training session was divided into six periods of 4 min of running and 1 min of rest [14]. The correct execution of treadmill exercise was constantly supervised, and those animals that did not adequately run on the treadmill because they either stopped frequently or ran irregularly were excluded from the study. Housing conditions and experimental procedures were in accordance with the European Union regulation on the use of animals for scientific purposes (*2003/65/CE*) and as promulgated by Spanish legislation (*RD 1201/2005*). Besides, the *University of Valencia Animal Care and Use Committee* approved all the procedures used in this study.

In order to analyse the intrinsic modifications of cardiac response in VF, isolated hearts were used to make them independent of vagal influence. Perfusion was maintained with a *Langendorff* system in order to avoid metabolic deterioration [15].

Cardiac mapping recordings were acquired with a commercial 256-channel system (*MAPTECH, Waalre, The Netherlands*). An electrode array of 240 electrodes (inter-electrode distance of 1 mm) was localised on the left ventricle. Each recording had 5 min of duration, acquired at a sampling rate of 1 kHz. VF was induced by pacing with increasing frequencies using an electrode placed in the ventricle, outside of the array capturing electrode.

### 2.2. Data processing

The procedure undergone to analyse the recordings involved a pre-processing stage, a frequency domain analysis and a time domain analysis. These analyses measured four parameters from which 18 features were calculated.

1. *Pre-processing stage.* Recordings were processed in consecutive segments of four seconds. The signal quality of each 4 seconds-segment was analysed, discarding the signals of those electrodes in the array with low amplitude or a high presence of noise [4].
2. *Frequency domain analysis. Welch* spectrum was obtained for all recording electrodes in each segment, using a *Hanning* window (2 non-overlapped sections and zero padding until 4096 samples). The Dominant Frequency (DF) and the Normalized Energy (NE) were calculated [16]. The DF was determined as the frequency with maximum spectral energy. The NE was defined as the spectral energy in a window placed at DF $\pm$ 1 Hz, and normalised with spectral energy in the interest band (5–35 Hz).
3. *Time domain analysis.* In order to analyse VF regularity and organization, two parameters were calculated: Regularity Index (RI) and Number of Occurrences (NO). The algorithm used for the RI computation [4] was a modification of the original one [17], in order to adapt it to the electrophysiological characteristics of the used cardiac model. More precisely, the local activation wave duration was increased up to 50 ms. The number of occurrences (NO) was also calculated as the ratio of samples which amplitude was inside a zero centred window respect to the total number of samples [18].

With this procedure, DF, NE, RI and NO parameters were sequentially calculated for each electrode and temporal segment, obtaining one map for each parameter and time segment. The first eight features were obtained as the mean value (*mDF,mNE,mRI, mNO*) and standard deviation (*sDF,sNE,sRI,sNO*) of each parameter map. The variation coefficient of the Number of Occurrences maps (*νcNO*) was also computed.

The authors developed an algorithm to study the regions of interest (ROI) [5], previously used in the VF analysis of the dominant frequency [19]. To obtain the ROI, a threshold was applied to each parameter map. Later on, a ROI membership label was assigned to each electrode, according to the threshold criteria and its neighbourhood with electrodes that also passed the threshold. From this ROI analysis, three features were obtained for each DF, NE and RI parameter map [5]:

*ROI spatial number (ROIsnDF, ROIsnNE, ROIsnRI):* the number of ROI detected in a map, a measure of spatial fragmentation.

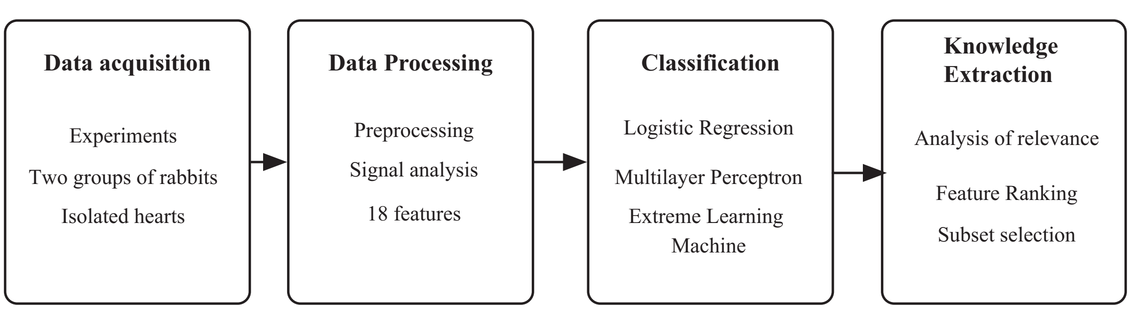


**Fig. 1.** Proposed study workflow.