

PAPER • OPEN ACCESS

Performance Analysis of ANN and SVM in ECG Based Arrhythmia Identification

To cite this article: V C C Roza *et al* 2018 *J. Phys.: Conf. Ser.* **1065** 132004

View the [article online](#) for updates and enhancements.



ECS The Electrochemical Society
Advancing solid state & electrochemical science & technology

239th ECS Meeting with IMCS18

DIGITAL MEETING • May 30-June 3, 2021

Live events daily • Free to register

Register now!

Performance Analysis of ANN and SVM in ECG Based Arrhythmia Identification

V C C Roza^{1,2}, A M Almeida¹, P M B Silva Girão^{3,4}, and O A Postolache^{1,5}

¹ Instituto Universitário de Lisboa, ISCTE-IUL/IT, Lisbon, Portugal

² Universidade Federal do Rio Grande do Norte, UFRN, RD Norte, Brazil

³ Instituto de Telecomunicações, Lisbon, Portugal

⁴ DEEC, Instituto Superior Técnico/UL, Lisbon, Portugal

⁵ Instituto de Telecomunicações, IT-IUL, Lisbon, Portugal

E-mails: valber_cesar_roza@iscte.pt, ana.almeida@iscte.pt,
psgrao@tecnico.ulisboa.pt, opostolache@lx.it.pt

Abstract. This paper presents a performance analysis of Artificial Neural Network (ANN) and Support Vector Machine (SVM) algorithms in arrhythmia identification task based on ECG signals. Six features are used for both algorithms: short signal 1-D wavelet energy (SS-WVE), short signal continuous wavelet transform mean (SS-CWTM), heart rate (HR), R-peaks root mean square (R-RMS), RR-peaks variance (RR-VAR) and QRS-complex standard deviation (QRS-SD). The identification methods use the MIT-BIH Dataset (Massachusetts Institute of Technology–Beth Israel Hospital) for training, validation and test phases. In this work, preliminary results shown that the classification obtained using SVM is marginally better than the one obtained with the ANN classifier for the same classification task (i.e. arrhythmia pattern identification).

1. Introduction

One of the most important and non-invasive techniques used in medicine is the Electrocardiography (ECG). It is used to observe the heart activities, i.e. variation and abnormalities over a period of time. This signal consists of several waves patterns such as U-waves, P-waves, T-waves and QRS-complexes and provides an important information about the electrical activity of the heart. Abnormal patterns in this electrical activity may represent heart diseases such as, atrial fibrillation (AF) [1], Brugada syndrome, long QT syndrome, short QT syndrome, and arrhythmia [2], that are defined by any structural cardiac defects, responsible for a large number of sudden, unexpected deaths. However, some of these diseases can't be visually distinguished easily by a medical specialist due to its similar appearance with other ECG signals [3]. A computational analysis may be used to detect these small abnormalities. Then, to allow an automatic detection of these abnormalities, several features may be extracted from ECG such as: heart rate variability (HRV) [4], morphological features using temporal and frequency-domain analysis [5, 6]. To identify it, Markov models, artificial neural networks (ANN) and support vector machine (SVM) [7-9] can be used.

In a general definition, arrhythmia represents an irregularity or rapidity of the heartbeat or an abnormal heart rhythm and can initiate or exacerbate acute systolic heart failure [1, 9, 10, 11]. Its analysis and patterns are explored in several works. Leren et al., investigated early markers of arrhythmic events and improved risk stratification in early arrhythmogenic right ventricular cardiomyopathy [2]. Lin, proposes a method for heartbeat identification from ECG using ANN and grey relational analysis (GRA) [1]. Shadmand and Mashoufi developed a new personalized ECG signal classification using an ANN variant named block-based neural network (BBNN) [3]. Pavinelli et al. proposed a novel nonlinear phase space based method to quickly and accurately identify life-threatening arrhythmias [12].

This paper presents a comparison analysis between the performance ANN and SVM in arrhythmia identification based on two types of arrhythmias patterns from ECG signals, considering six different features and the MIT-BIH Arrhythmia Dataset.

2. Dataset description

This work uses the MIT-BIH (Massachusetts Institute of Technology–Beth Israel Hospital) Dataset to provide the ECG signals used during training, validation, and test of both classifiers.



Content from this work may be used under the terms of the [Creative Commons Attribution 3.0 licence](https://creativecommons.org/licenses/by/3.0/). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

Three ECG classes are considered according to Table 1. It is based on time-date in seconds, grid interval x-axis of 0.2 seconds, grid interval y-axis of 0.5 mV, and standard data format, consisting of 240 NSRD instances, 432 AD instances, and 252 SAD instances.

Table 1. Dataset classes and phases for identification task

MIT-BIH Dataset				
Classes (instances)	Training (50%)	Validation (25%)	Test (25%)	Total
Normal (NSRD)	120	52	68	240
Arrhythmias (AD)	216	117	99	432
Suprav. Arrhyt. (SAD)	126	62	64	252
Total instances	462	231	231	924

3. Processing and feature extraction

In the training, validation and test phases, 67 ECG signals (i.e. large-signals or complete signals) with 1 hour duration from the MIT-BIH Dataset are used. In the processing, these signals are divided in 12 equal parts (or short-signals SS) with duration according with the literature t_s (5-min) [2,7], amount of R-peaks N_p by short signal, and signal length L_p (from the first to the last R-peak).

3.1 Signal processing

Signal processing is used in this work to enable the ECG signals to be used in the classifier inputs, including the signal samplings, low pass Butterworth filtering, data smoothing by Savitzky–Golay filtering, FFT transform and baseline wander. These steps are applied to remove noises, to support the RR-peaks identification and to maintain the baseline signals during the signal processing executed before the feature extraction process, as shown Figure 1.

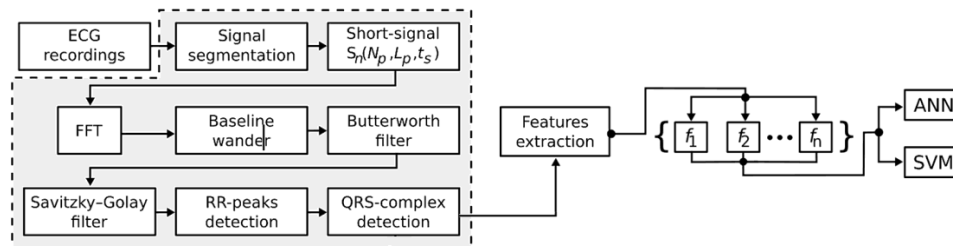


Figure 1. Signal processing (in gray) applied for each ECG signal from the dataset.

Similarly to other tests found in the literature [1], this work considers the variations of the power spectrum from 0 to 20Hz in frequency domain. Afterwards, the Fast Fourier Transform (FFT) is applied to determine the frequency spectrum with signal sampling frequency of 500Hz and average of 54,166 samples for each short-signal.

3.2 Features extraction

In this work, all extracted features from each short-signal, are applied to both identification methods. These features were chosen according to several bibliographies which presented satisfactory results. These features are: beats per minutes or heart rate (HR); R-peaks root mean square (R-RMS); QRS-complex standard deviation (QRS-SD); RR-peaks variance (RR-VAR); short signal 1-D wavelet energy (SS-WVE); short signal continuous wavelet transform mean (SS-CWTM).

4. Arrhythmia identification

The present arrhythmia pattern identification uses six features and its classification process is based on ANN and SVM methods. Thus, is considered three phases: training, using 50% of the dataset to train the classifier; validation, using 25% of the dataset to validate the training results; and test, that using the last 25% of the dataset to identify arrhythmias.

4.1. Artificial neural network method

The proposed ANN is based on a multilayer perceptron (MLP-ANN) with backpropagation algorithm, considering each ECG instance. Its *induced local field*, is represented by Equation 1, where x_i goes to input to neuron j and w_{ji} denotes a connection from neuron j to i [13].

$$v_j(n) = \sum_{i=1}^m w_{ji}(n)x_i(n), j \geq 1 \quad (1)$$

For the hidden and output layers, were applied the *sigmoidal activation function* and *softmax function*, respectively.

4.2. Support vector machine method

The SVM is the other used method, which for each ECG instance x_i , it returns an output y_i . Since the SVM is based on decision surfaces as a parameter of classification, the considered decision surfaces or hyperplanes are defined as below [13],

$$\sum_{j=1}^{\infty} w_j \phi_i(x) = 0 \quad (2)$$

where w represents the *weight vector*, and $\phi(x)$ represents the *feature vector*.

4.3. Output patterns

The output patterns are: “SAD” (supraventricular arrhythmia); “NSRD” (normal cardiac rhythm); and “AD” (arrhythmic pattern).

5. Results and discussion

Table 2, shows the results of the ANN according to several parameters such as, training with 500 iterations, amount of hidden neurons (h_n), amount of hidden layers (h_l) and static values for $\eta=0.3$ and $\alpha=0.2$. The best accuracy from the ANN test was reached with $h_l=1$ and $h_n=100$.

Table 3, shows the results of the SVM test, changing the complexity parameter (c). The best accuracy from the SVM test is reached with $c = 10^3$. The ANN classification reached a total accuracy of 82.68%, and partial accuracy (accuracy by classes) of 85.29% when identifying NSRD patterns, of 85.85% when identifying AD patterns and 75.00% for SAD patterns. When only arrhythmias are analysed, the ANN reached an accuracy of 92.39%, i.e. from 92 AD patterns, only 7 results were misclassified as SAD. Moreover, from 59 SAD patterns, only 11 were classified as AD, reaching an accuracy of 81.36%.

Table 2. Accuracy analysis of ANN using percentage split.

Hidden layers (h_l)	Hidden neurons \times accuracy				
	$h_n = 6$	$h_n = 30$	$h_n = 50$	$h_n = 70$	$h_n = 100$
1	75.75%	79.22%	80.51%	80.08%	82.68%
2	74.89%	71.86%	77.05%	77.05%	73.59%
3	51.94%	78.78%	78.35%	78.35%	73.59%

Table 3. Accuracy analysis of SVM using percentage split.

Kernel Function (k)	Complexity parameter				
	$c = 1$	$c = 10$	$c = 40$	$c = 10^2$	$c = 10^3$
RBF Kernel	42.85%	42.85%	47.61%	57.14%	55.84%
Polynomial	50.64%	56.71%	55.41%	55.41%	54.97%
Pearson Kernel	72.29%	74.89%	82.25%	83.98%	84.84%

The SVM classification reached a total accuracy of 84.84%, and partial accuracy (by classes) of 86.76% when identifying NSRD patterns, 86.86% when identifying AD patterns and 79.68% for SAD patterns. When only arrhythmias are analysed, the SVM reached an accuracy of 90.52%, i.e. from 95 AD patterns, only 9 results were misclassified as SAD. From 61 SAD patterns, 10 were classified as AD, reaching an accuracy of 83.60%.

6. Conclusion

This work presented a comparison analysis between ANN and SVM within the context of arrhythmia pattern identification using ECG from MIT-BIH dataset, and the features, HR, R-RMS, RR-VAR, QRS-SD, SS-WVE and SS-CWTM. the proposed SVM method produces better results the proposed ANN method. Was possible to note that the SVM reached better total accuracies than ANN classifier for the same classification task (i.e. arrhythmia pattern identification). Although, when the classification was based between arrhythmias, the ANN reached better results. Furthermore, the ANN tests shown that the amount of hidden layers does not always determine the goodness fit of classification.

In future work, the same classification methods and tools will be applied to emotion recognition based on psychophysiological signals and speech.

Acknowledgement

The work now reported was supported by Instituto de Telecomunicações, IT-IUL and ISCTE-IUL.

References

- [1] D. Farwell, and M.H. Gollob, "Electrical heart disease: Genetic and molecular basis of cardiac arrhythmias in normal structural hearts," *Canadian Journal of Cardiology*, vol.23, pp. 16A-22A, 2007.
- [2] I.S. Leren, J. Saberniak, T.F. Haland, T. Edvardsen, and K.H. Haugaa, "Combination of ECG and Echocardiography for Identification of Arrhythmic Events in Early ARVC," vol. 10(5), pp. 505-513, 2017.
- [3] S. Shadmand, and B. Mashoufi, "A new personalized ECG signal classification algorithm using block-based neural network and particle swarm optimization," *Biomedical Signal Processing and Control*, vol. 25, pp. 12-23, 2016.
- [4] M. Malik, "Heart rate variability: standards of measurement, physiological interpretation, and clinical use," *European Heart Journal*, vol. 17, pp. 354-381, 1996.
- [5] W.-C. Peng, H. Wang, J. Bailey, V.S. Tseng, T.B. Ho, Z.-H. Zhou, and A.L. Chen, "Myocardial infarction classification by morphological feature extraction from big 12-lead ECG data," *Trends and Applications in Knowledge Discovery and Data Mining*, vol. 8643 of Lecture Notes in Computer Science, Springer International Publishing, 2014.
- [6] A. Kaveh, and W. Chung, "Temporal and spectral features of single lead ECG for human identification," *IEEE Workshop on Biomedical Measurements and Systems for Security and Medical Applications*, vol.1, pp. 17-21, 2013.
- [7] I. Mporas, V. Tsirka, E. Zacharaki, M. Koutroumanidis, and V. Megalooikonomou, "Evaluation of time and frequency domain features for seizure detection from combined EEG and ECGm signals," *Proceedings of 7th International Conference on Pervasive Technologies Related to Assistive Environments*, vol. 1, pp. 28:1-28:4, 2014.
- [8] S. Banerjee, R. Gupta, and M. Mitra, "Delineation of {ECG} characteristics features using multiresolution wavelet analysis method," *Measurement*, vol. 45(3), pp. 474-487, 2012.
- [9] P.P. San, S.H. Ling, and H.N. Nguyen, "Evolvable rough-block-based neural network and its biomedical application to hypoglycemia detection system," *IEEE Trans. Cybern*, vol. 44(8), pp. 1338-1349, 2014.
- [10] R. Gopinathannair, S.P. Etheridge, F.E. Marchlinski, F.G. Spinale, D. Lakkireddy, and B. Olshansky, "Arrhythmia-Induced Cardiomyopathies- Mechanisms, Recognition, and Management," *Journal of the American College of Cardiology*, vol. 66(15), pp. 1714-1728, 2015.
- [11] R. Kohno, Y. Oginosawa, and H. Abe, "Identifying atrial arrhythmias versus pacing-induced rhythm disorders with state-of-the-art cardiac implanted devices", *Journal of Arrhythmia*, vol. 30, pp. 82-87, 2014.
- [12] R.J. Povinelli, F.M. Roberts, M.T. Johnson, and K.M. Ropella, "Are nonlinear ventricular arrhythmia characteristics lost, as signal duration decreases?," *Computers in Cardiology*, vol. 1, pp. 221-224, 2002.
- [13] P.J.G. Nieto, J.M. Torres, M.A. Fernández and C.O. Galán, "Support vector machines and neural networks used to evaluate paper manufactured using *Eucalyptus globulus*," *Applied Mathematical Modelling*, vol. 36(12), pp. 6137-6145, 2012.