Prediction of Ventricular Tachyarrhythmia Using Deep Learning

Dalila Barbosa[1], E. J. Solteiro Pires[1,2(✉)], Argentina Leite[1,2], and P. B. de Moura Oliveira[1,2]

[1] Escola de Ciências e Tecnologia Universidade de Trás-os-Montes e Alto Douro, Vila Real, Portugal
al71375@utad.eu, {epires,tinucha,oliveira}@utad.pt
[2] INESC TEC, Porto, Portugal

Abstract. Ventricular tachyarrhythmia (VTA), mainly ventricular tachycardia (VT) and ventricular fibrillation (VF) are the major causes of sudden cardiac death in the world. This work uses deep learning, more precisely, LSTM and biLSTM networks to predict VTA events. The Spontaneous Ventricular Tachyarrhythmia Database from PhysioNET was chosen, which contains 78 patients, 135 VTA signals, and 135 control rhythms. After the pre-processing of these signals and feature extraction, the classifiers were able to predict whether a patient was going to suffer a VTA event or not. A better result using a biLSTM was obtained, with a 5-fold-cross-validation, reaching an accuracy of 96.30%, 94.07% of precision, 98.45% of sensibility, and 96.17% of F1-Score.

Keywords: Ventricular Tachyarrhythmia · Deep learning · Long Short Term Memory

1 Introduction

Ventricular tachycardia and Fibrillation are the primary ventricular tachyarrhythmias that cause sudden death. Ventricular tachycardia (VT) is mostly caused by a faulty heart signal that triggers a fast heart rate in the lower heart chambers, i.e., the ventricles. The fast heart rate does not allow the ventricles to completely fill and squeeze to pump the blood needed for the body. Ventricular fibrillation (VF) is also a type of abnormal heart rhythm that affects the heart's signals and causes the lower heart chambers to twitch uselessly. This result in the heart not pumping blood to the rest of the body. One solution to this problem is implanting cardioverter defibrillators. Implantable cardioverter defibrillators (ICDs) are the cornerstone of sudden cardiac death prevention by terminating ventricular tachyarrhythmias, such as ventricular tachycardia and ventricular fibrillation. Over the past four decades, ICDs have been used to reduce the risk

A. Cunha et al. (Eds.): MobiHealth 2022, LNICST 484, pp. 51–60, 2023.
https://doi.org/10.1007/978-3-031-32029-3_5

of cardiac arrest [7] by detecting arrhythmia and delivering electrical shocks to restore the heart rhythm [11]. It can be incorporated into the ICD, a reliable predictor of an imminent ventricular tachyarrhythmia, episode capable of preventive therapy, which would have important clinical utilities [10].

Several studies were made, considering signal features used to develop detection algorithms for anomalous cardiac signals. Oliver Faust *et al.* [2] proposed a deep learning system to identify irregularities in the heart signals, partitioning them into windows of 100 beats and feeding a Recurrent Neural Network with a Long Short-Term Memory. They obtained an accuracy in the range 95.51%-99.77% accurate. Atiye Riasi *et al.* [10] presented an technique using morphological features of an electrocardiogram (ECG). Also, the classification of selected features done by a Support Vector Machine (SVM) identified hidden patterns in the ECG before the occurrence of the episode, achieving 88% and 100% of sensitivity and specificity, respectively. Segyeong Joo *et al.* [4] proposed a classifier that could predict events recurring to artificial neural networks (ANNs) considering time-domain and non-linear parameters. It achieved 82.9% of sensitivity and 71.4% of specificity. Taye *et al.* [14] proposed a convolutional neural network, considering one dimension (1-D CNN) with signal features. Furthermore, comparing the obtained CNN prediction to other machine learning performances, such as ANN, SVM, and KNN, the higher accuracy value of 84.6% was reached.

Analyzing the aforementioned, this work uses a long short-term memory (LSTM) network, analyzing different features, which can help to predict the occurrence of abnormal episodes.

The paper is organized into more three sections. Section 2 describes the database used, the signals pre-processing, the features extracted, and neural networks. Next, Sect. 3 presents and discusses the results. Finally, Sect. 4, draws the conclusions.

2 Materials and Methods

2.1 Dataset

The dataset considered was collected from the PhysioNet database, known as the Spontaneous Ventricular Tachyarrhythmia Database [3]. The dataset consists of 135 RR interval pairs, recorded by ICD, of 78 patients (RR is the time elapsing between two consecutive R waves in the electrocardiogram). The patient dataset has different numbers for VF and VF events. Of the 135 pairs of RR intervals, 29 were VF and 106 VT, and every one of them had its corresponding normal sinus rhythm.

2.2 Pre-processing

Every signal collected from the Spontaneous Ventricular Tachyarrhythmia Database was initially filtered to reduce and smooth out high-frequency noise. The length of each signal has approximately 1024 beats. From every signal, we only use a specific section. Out of the normal sinus rhythm, we used the first 256 beats and the last 256 beats before the abnormal episode occurred.

2.3 Feature Extraction

Several methods can analyze variations in heart signals. The data analysis can be performed directly on the signals or indirectly, extracting characteristics in the first step and then analyzing these features. This work extracted some features from the time regions, the first 256 beats from the normal signal and the last 256 from the abnormal signal. Eight features were considered, three in the time domain, two in the frequency domain, and the last three are nonlinear features.

Time-Domain. The time-domain measures are the easiest to determine. In this category, the features used were the mean RR intervals (meanRR), the NN intervals standard deviation (SDNN), which comprises the cyclic components liable for variability in the recording period [1], and the proportion of successive RR interval differences greater than 50 ms (pNN50).

Frequency Domain. The frequency domain characteristics are obtained with the parametric power spectrum estimation from the RR intervals. The AutoRegressive (AR) model is the most widely used in this domain to characterize the data. [1]. This model eases the determination of the parameters by solving linear equations. Right below, it is represented the AR model in order of p, $AR(p)$:

$$x_t = \sum_{k=0}^{p} \alpha_k x_{t-k} + \varepsilon_t \tag{1}$$

with α_k representing the AR coefficients, ε_t the white noise with zero mean and σ_ε^2 variance. There is also the AR model a spectrum equation:

$$P_{AR}(\omega) = \frac{\sigma_\varepsilon^2}{|1 - \sum_{k=1}^{p} \alpha_k e^{(-iwk)}|^2} \tag{2}$$

Information quantification relative to different frequency bands is possible because of spectral analysis. In this case, the low frequency (LF) and high frequency (HF) bands are in between (0.04–0.15 Hz), and (0.15–0.4 Hz), respectively. Typically, HF characterizes parasympathetic nervous system activity, while LF is associated with sympathetic and parasympathetic systems. The evaluation between these two components generally is made from the determination of LF and HF areas [1].

Nonlinear-Domain. The nonlinear features were obtained through nonlinear methodologies, such as, Detrended Fluctuation Analysis (DFA) [8] and Approximate Entropy (ApEn) [9]. DFA can understand unique insights into the neural organization [12], identifying long-range correlations embedded in a nonstationary time series. The first step of this method has to do the first integration, using Eq. (3):

$$y(i) = \sum_{t=1}^{i} [u(t) - \overline{u}] \tag{3}$$

Then, it divides the signal into k segments, where the linear local trend $y_k(i)$ is calculated by a linear regression. DFA is summarized by Eq. (4), which gives the root-mean-square between a segment of the first integrated signal and $y_k(i)$.

$$DFA(k) = \sqrt{\frac{1}{N}\sum_{i=1}^{N}[y(i) - y_k(i)]^2} \tag{4}$$

Equation (4) is repeated for several segments of length k. This work considers two scaling exponents: $\alpha 1$ for $4 \leq k \leq 11$ and $\alpha 2$ for $12 \leq k \leq 32$, corresponding to the short memory correlation. For non correlated data, the scaling exponent adopted is $\alpha = 0.5$. On the other hand, values of $\alpha > 0.5$ for large scales k indicate data with long-range correlations [8].

ApEn is commonly used in short and noisy data. Being m and r an integer and real positive number, respectively. Given a signal $u(1), u(2), ..., u(N)$, form a sequence of vectors $x(1), x(2), ..., x(N-m+1)$ in $\mathbb{R}^m$, defined by $x(i) = [u(i), u(i+1), \ldots, u(i+m-1)]$ [9].

Next, it is defined for each $i, 1 \leq i \leq N-m+1$,

$$C_i^m(r) = \frac{\text{number of } j : d[x(i), x(j)] \leq r}{N-m+1},$$

where $d[x(i), x(j)] = max(|u(i+k-1) - u(j+k-1)|)$, $k = 1, 2, \ldots, m$

Therefore, the *ApEn* feature is calculated by

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

where

$$\phi^m(r) = (N-m+1)^{-1}\sum_{i=1}^{N-m+1}\log(C_i^m(r)).$$

In the present work, the *ApEn* uses $m = 2$ and $r = 0.2\sigma$, where σ is the data standard as recommended by [9]. Features were extracted from the RR intervals, and in Fig. 1, it can be seen the values of each class, the normal sinus rhythm, and the precedent of the abnormal episode.

Table 1 presents the mean values of the several features extracted. On the left are the values of the normal sinus rhythm, and on the right, we have the ones from the abnormal signal.

2.4 Neural Network

Artificial neural networks (ANNs) are a machine learning subset considered the heart of deep learning models. ANNs' structure is inspired by the human brain, imitating how biological neurons' signals spread between them. In the network, the neurons nodes are connected with links, enabling them to communicate and exchange information, conceiving them to store information.

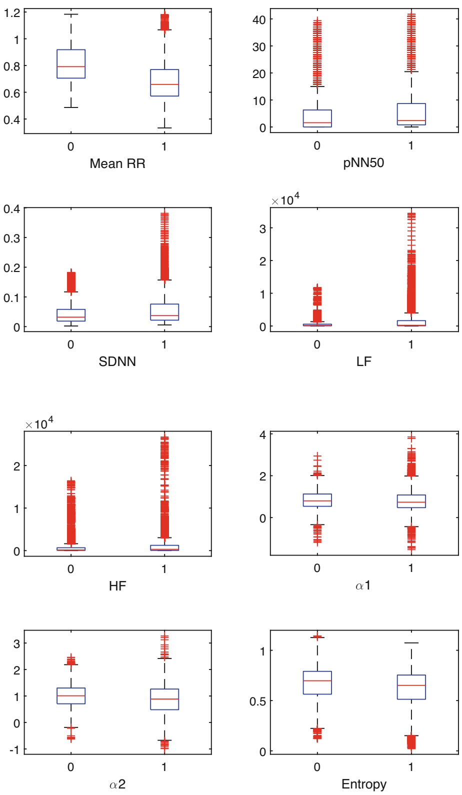

Fig. 1. Boxplot of HRV features for two groups: Normal Sinus Rhythm (0) and Signal before the VTA episode (1).

This work considers a long short-term memory (LSTM) and a biLSTM, which are a type of recurrent neural network (RNN). There is an architectural difference between RNN and LSTM, where the LSTM is a specialized form of RNN architecture. In practice, simple RNNs have limitations related to the capacity to learn longer-term dependencies. RNNs are commonly trained through backpropagation, in which they may experience either a 'vanishing', or 'exploding' gradient problem, causing the network to become very small or very large, affecting the effectiveness in applications using a network. That is where LSTMs come in. To overcome this problem by using additional gates to control the information exported from the hidden cells as output and to the next hidden state.

An LSTM unity has three gates: input, output, and forget. These gates control the information flow within the cell. The input gate updates the unit status and decides the relevant information to the current step. The output gate determines the next hidden state value. The forget gate has better control over the gradient flow than RNN and determines the important information from those that should be ignored. The choice to preserve or delete information is defined

Table 1. HRV features: Control and VTA dataset

Features	Control Sinus Rhythm	VTA
MeanRR (s)	0.81 ± 0.14	0.69 ± 0.16
pNN50 (%)	5.03 ± 7.71	5.92 ± 8.20
SDNN (s)	0.04 ± 0.03	0.05 ± 0.04
LF (s^2)	535.80 ± 1046.8	1455.20 ± 2759.60
HF (s^2)	766.21 ± 1684.1	1124.80 ± 2202.30
$\alpha 1$	0.84 ± 0.39	0.80 ± 0.44
$\alpha 2$	1.00 ± 0.42	0.89 ± 0.51
Entropy	0.67 ± 0.16	0.62 ± 0.20

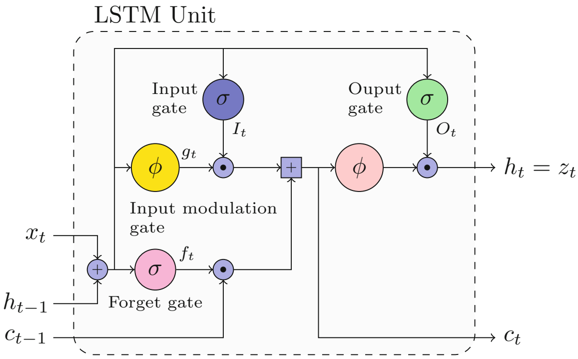

Fig. 2. Long short-term memory unit [6].

by weights determined during the training phase [13]. The advantage these gates bring is that they allow learning long-term relationships more effectively.

The Bidirectional LSTM (BiLSTM) consists of two hidden layers one after another. The additional LSTM layer allows the information flow to be routed in both directions, *i.e.*, the input sequence flows backward in the additional LSTM layer. Therefore, the network is able to learn long terms easier [13]. Figures 2 and 3 illustrates the LSTM unit and biLSTM architecture.

2.5 Classifier Performance

The classifiers performance is evaluated using four metrics. These are accuracy ACC (5) which determines how close a measurement is to the true or accepted value of being true, precision (PRE), referring to how close measurements are to each other, and sensibility (SEN) to evaluate how the parameters and states of the model influence the model output, and F1-Score (F1), measuring the model's accuracy on our dataset. These metrics were evaluated from the equations below:

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \tag{5}$$

$$PRE = \frac{TP}{TP + FP} \tag{6}$$

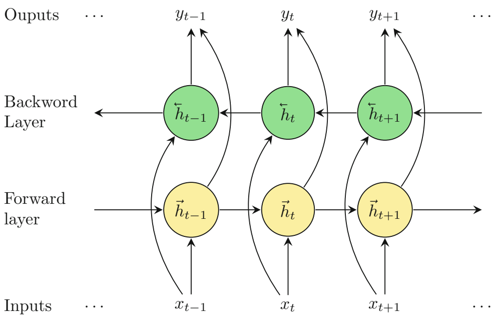

Fig. 3. Bidirectional long short-term memory [6].

$$SEN = \frac{TP}{TP + FN} \tag{7}$$

$$F1 = \frac{2PRE.SEN}{PRE + SEN} \tag{8}$$

The network performance evaluation is measured using four groups of results: TP, TN, FP, and FN. Where TP represents the true positive data, $i.e$, the correctly classified patients as positive; TN represents the true negative, also classified correctly but this time as negative; FP represents the false positive, classified patients improperly as positive; And lastly, FN represents the false negative data, this is, the incorrectly classified data as negative.

The technique used to evaluate the performance of the model, in particular the generalization ability, is k-fold cross-validation. This technique consists of dividing the data (signals) into k groups. Each group is separated to test the estimation obtained by training the other groups [5]. In this study, 5-fold cross-validation was used, meaning the signals are divided into five groups executing five iterations, one for each group.

3 Results and Discussion

In this study, two networks were trained and tested. We tried different quantity combinations of hidden units, epochs, and mini-batch size for better performance. The number of hidden units in an LSTM refers to the dimensionality of the 'hidden state' of the LSTM. The hyperparameter *Epochs* defines the number of times the entire training dataset is processed, and mini-batch size is used for a subset of the dataset to take another step in the learning process. The LSTM model used 30 hidden units, 30 epochs, and 20 mini-batch size were used. The biLSTM model used the same numbers except for the epochs, instead of 30, 20 was used.

Table 2. Training Set LSTM (Hidden Units - 30, maxEpochs - 30, miniBatchsize - 20)

It.	TP	TN	FP	FN	ACC (%)	PRE (%)	SEN (%)	$F1$ (%)
1.º	108	108	0	0	100.00	100.00	100.00	100.00
2.º	107	108	0	1	99.54	100.00	99.07	99.53
3.º	107	108	0	1	99.54	100.00	99.07	99.53
4.º	107	108	0	1	99.54	100.00	99.07	99.53
5.º	107	108	0	1	99.54	100.00	99.07	99.53
mean					99.63	100.00	99.26	99.63

Table 3. Testing Set LSTM (Hidden Units - 30, maxEpochs - 30, miniBatchsize - 20)

It.	TP	TN	FP	FN	ACC (%)	PRE (%)	SEN (%)	F1-Score (%)
1.º	25	26	1	2	94.44	96.15	92.59	94.34
2.º	26	26	1	1	96.30	96.30	96.30	96.30
3.º	27	27	0	0	100.00	100.00	100.00	100.00
4.º	26	25	2	1	94.44	92.86	96.30	94.55
5.º	25	26	1	2	94.44	96.15	92.59	94.34
mean					95.93	96.29	95.56	95.90

Table 4. Training Set biLSTM (Hidden Units - 30, maxEpochs - 20, miniBatchsize - 20)

It.	TP	TN	FP	FN	ACC (%)	PRE (%)	SEN (%)	F1-Score (%)
1.º	107	108	0	1	99.54	100.00	99.07	99.53
2.º	107	108	0	1	99.54	100.00	99.07	99.53
3.º	107	108	0	1	99.54	100.00	99.07	99.53
4.º	107	108	0	1	99.54	100.00	99.07	99.53
5.º	108	108	0	0	100.00	100.00	100.00	100.00
mean					99.63	100.00	99.26	99.63

We applied a 5-fold cross-validation using the 8 features extracted from the signals, such as meanRR, SDNN, pNN50, LF, HF, α_1, α_2, and entropy. The 270 signals were divided into 80% of the data for train, and the last 20% for the test classifier, then 216 signals train the model and the last 54 test it. From the 270 signals, 135 of them are from a normal sinus signal, more precisely, the first 256 beats from the normal signal. The other 135 signals are from a signal that precedes a ventricular tachyarrhythmia episode, and it was only used the last 256 beats before the abnormal episode happens.

Tables 2, 3, 4 and 5 presents the obtained results. The best result are achieved with the biLSTM with 30 hidden units, 20 epochs, and 20 mini-batch size, getting an accuracy of 95.93%, 96.29% of precision, 95.56% of sensibility, and lastly, 95.90% of F1-score. On the other hand, with the LSTM, the results were very similar, but slightly lower than the biLSTM. It used 30 hidden units and 20 mini-batch size, changing only the number of epochs to 30. It reached an accuracy of 96.30%, 94.07% of precision, 98.45% of sensibility and 96.17% of F1-Score.

Table 5. Testing Set biLSTM (Hidden Units - 30, maxEpochs - 20, miniBatchsize - 20)

It.	TP	TN	FP	FN	ACC (%)	PRE (%)	SEN (%)	F1-Score (%)
1.º	26	27	0	1	98.15	100.00	96.30	98.11
2.º	24	26	1	3	92.59	96.00	88.89	92.31
3.º	27	27	0	0	100.00	100.00	100.00	100.00
4.º	26	26	1	1	96.30	96.30	96.30	96.30
5.º	24	27	0	3	94.44	100.00	88.89	94.12
mean					96.30	94.07	98.46	96.17

4 Conclusion and Future Work

This work considered LSTM and BiLSTM networks to predict humans from having a ventricular tachyarrhythmia episode, *i.e.*, Ventricular Fibrillation or Ventricular Tachycardia, using a dataset retrieved from PhysioNET. Features were extracted from the RR intervals considering 270 signals from the RR intervals retrieved from the dataset. The data was separated in 80% and 20% for training and testing, respectively. The results indicate that using deep learning with features extraction from the RR intervals, it is possible to prevent with high accuracy a VTA event.

The BiLSTM network normally has higher accuracy than the LSTM network outperforming it by 0.4% in the mean accuracy. Thus by using this method it is possible to detect an episode of ventricular tachycardia or fibrillation and allow the patient to ask for quick assistance when needed. The best result of the BiLSTM was obtained using RR intervals features extraction obtaining 96.30% accuracy, 94.07% precision, 98.45% sensibility, and 96.17% F1-Score. Although the results between the two neural networks were really close, biLSTM got a slightly better outcome. Future work can be developed, focusing on trying a different group of features that could raise the accuracy.

References

1. Electrophysiology, task force of the European society of cardiology the north American society of pacing: heart rate variability: standards of measurement, physiological interpretation, and clinical use. Circulation **93**(5), 1043–1065 (1996)
2. Faust, O., Shenfield, A., Kareem, M., San, T.R., Fujita, H., Acharya, U.R.: Automated detection of atrial fibrillation using long short-term memory network with RR interval signals. Comput. Biol. Med. **102**, 327–335 (2018)
3. Goldberger, A.L., et al.: PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. Circulation **101**(23), e215–e220 (2000)
4. Joo, S., Choi, K.J., Huh, S.J.: Prediction of ventricular tachycardia by a neural network using parameters of heart rate variability. In: 2010 Computing in Cardiology, pp. 585–588. IEEE (2010)
5. Karal, Ö.: Performance comparison of different kernel functions in SVM for different k value in k-fold cross-validation. In: 2020 Innovations in Intelligent Systems and Applications Conference (ASYU), pp. 1–5. IEEE (2020)

6. Monteiro, S., Leite, A., Solteiro Pires, E.J.: Deep learning on automatic fall detection. In: 2021 IEEE Latin American Conference on Computational Intelligence (LA-CCI), pp. 1–6 (2021). https://doi.org/10.1109/LA-CCI48322.2021.9769783
7. Parsi, A., O'Loughlin, D., Glavin, M., Jones, E.: Prediction of sudden cardiac death in implantable cardioverter defibrillators: a review and comparative study of heart rate variability features. IEEE Rev. Biomed. Eng. **13**, 5–16 (2019)
8. Peng, C.K., Havlin, S., Stanley, H.E., Goldberger, A.L.: Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series. Chaos: Interdisc. J. Nonlin. Sci. **5**(1), 82–87 (1995)
9. Pincus, S.M.: Approximate entropy as a measure of system complexity. Proc. Natl. Acad. Sci. **88**(6), 2297–2301 (1991)
10. Riasi, A., Mohebbi, M.: Prediction of ventricular tachycardia using morphological features of ECG signal. In: 2015 The International Symposium on Artificial Intelligence and Signal Processing (AISP), pp. 170–175. IEEE (2015)
11. Sanghera, R., Sanders, R., Husby, M., Bentsen, J.G.: Development of the subcutaneous implantable cardioverter-defibrillator for reducing sudden cardiac death. Ann. N. Y. Acad. Sci. **1329**(1), 1–17 (2014)
12. Sengupta, S., et al.: Emotion specification from musical stimuli: an EEG study with AFA and DFA. In: 2017 4th International Conference on Signal Processing and Integrated Networks (SPIN), pp. 596–600. IEEE (2017)
13. Siami-Namini, S., Tavakoli, N., Namin, A.S.: The performance of LSTM and BiLSTM in forecasting time series. In: 2019 IEEE International Conference on Big Data (Big Data), pp. 3285–3292. IEEE (2019)
14. Taye, G.T., Hwang, H.J., Lim, K.M.: Application of a convolutional neural network for predicting the occurrence of ventricular tachyarrhythmia using heart rate variability features. Sci. Rep. **10**(1), 1–7 (2020)