International Journal of Scientific & Engineering Research Volume 10, Issue 9, September-2019
ISSN 2229-5518

Heart Rate Variability Prediction based on the combination of Wavelet Decomposition and LSTM Networks

Imane El Boujnouni, Abdelhak Tali

Abstract— Heart Rate Variability (HRV) time series analysis are used as a prognostic tool which can help to identify hidden patterns to make decisions on lifestyle changes in high risk patients and aid in providing early diagnosis and treatments for people who have heart failure and need immediate help.
In this paper, we use the wavelet decomposition levels of temporal heart rate variability obtained from the historical information in the MIT-BIH database, as input to BLSTM to enhance the model performance in predicting heart rate compared to direct use of Bidirectional Long Short Term Memory with original heart rate temporal sequence.

Index Terms— BLSM, Deep Learning, HRV, Wavelet Decomposition.

1 INTRODUCTION

Cardiovascular diseases (CVDs) take the lives of 17.9 million people every year, 31% of all global deaths [1]. Heart failure (HF) has become a frequent manifestation of cardiovascular disease (CVD) affecting more than 23 million patients worldwide [2]. HF is caused by an inadequate blood filling function of the ventricular pump. A condition in which the heart does not pump well can cause the heart's discharge volume to be insufficient to meet the needs of body's metabolism [2].

Heart rate variability presents the time interval between two consecutive QRS complexes lying adjacent in ECG. The variation in RR interval is represented by heart rate variability (HRV), and its variation can be interpreted as a current or upcoming disease.

There are many different mathematical methods to analyze HRV time series, the most common and optimum one is the wavelet transform, which provides a good localization in both time and frequency domain.

This paper consists of two major parts: Discrete Wavelet Transformation (DWT) and Bidirectional Long Short Term Memory model (BLSTM).

Firstly, the heart rate time series obtained from the MIT-BIH database are used to train the LSTM and used as the inputs to the LSTM for heart rate prediction. Then, the neural network is tested using the sum of all wavelets coefficient at each level obtained from discrete wavelet decomposition.

2 WAVELET THEORY

2.1 Wavelet

A wavelet function ψ(t) is a small wave, which must be oscillatory in some way to discriminate between different frequencies, Fig. 1 shows an example of a possible wavelet, called the Morlet wavelet.

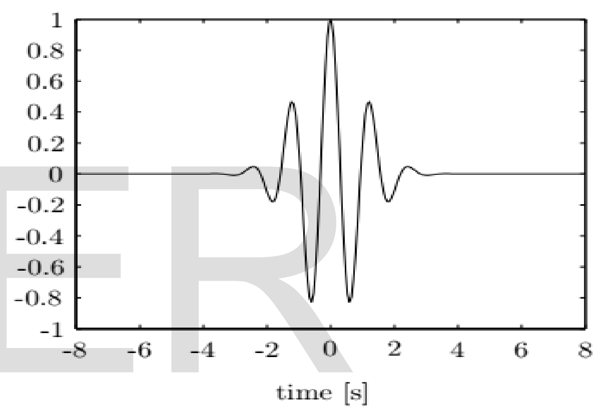

Fig.1 the Morlet wavelet

An analyzing function ψ(t) is classified as a wavelet if the following mathematical criteria are satisfied [4]:

1. A wavelet must have finite energy

$$E = \int_{-\infty}^{\infty} |\psi(t)|^2 dt < \infty \quad (1)$$

The energy E equals the integrated squared magnitude of the analyzing function $\psi(t)$ and must be less than infinity.

2. If Ψ(f) is the Fourier transform of the wavelet $\psi(t)$, the following condition must hold.

$$C_\psi = \int_0^{\infty} \frac{|\hat{\psi}(f)|^2}{f} df < \infty \quad (2)$$

This condition implies that the wavelet has no zero frequency component (Ψ(0) = 0), the mean of the wavelet $\psi(t)$ must equal zero. This condition is known as the admissibility constant. The value of $C\psi$ depends on the chosen wavelet.

3. For complex wavelets the Fourier transform Ψ (f) must be

both real and vanish for negative frequencies.

2.2 Continuous wavelet transform

The continuous wavelet transform (CWT) of the signal x*(t)* is defined as a convolution of the signal with a scaled and translated version of a base wavelet function . The wavelet transform (WT) of signal x*(t)* is defined as a combination of a set of basic functions, obtained by means of dilation *a* and translation *b* of a mother wavelet [5]:

$$W_a x(b) = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{+\infty} x(t)\psi\left(\frac{t-b}{a}\right) dt \tag{3}$$

Where ψ(t) is the mother wavelet, the ∗ indicates the complex conjugate. The mother wavelet is contracted and dilated by changing the scale parameter s.

2.3 Discrete wavelet transform

Discrete wavelet transform (DWT) is one of the wavelet transform development series. DWT works on two collections of functions called scaling functions and wavelet function that are each associated with a low-pass filter and a high-pass filter. The decomposition structure of wavelet transforms for level 3 is shown in Fig. 2 [5].

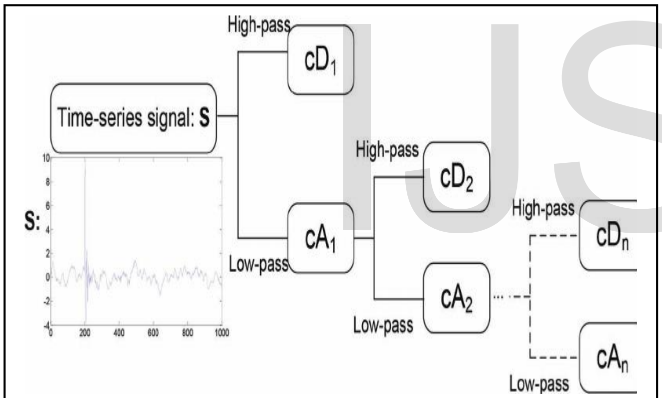

Fig. 2 The iterative decomposition process of discrete wavelet transforms

3 Deep LEARNING AND LSTM FOR TIME SERIES

3.1 Deep Learning

Deep Learning can be defined as special kind of neural networks composed of multiple layers. These networks have been successfully applied in various research fields. They are better than traditional neural network in persisting the information from previous event. Recurrent neural network (RNN) is one such machine that has a combination of networks in loop [5].

3.2 Recurrent NEURAL NETWORK

RNN has the ability to map target vectors from the whole history of the previous input .Thus RNN is more effective at modeling dynamics in sequential data when compared to traditional neural networks [5]. Sharing weights between hidden units across each time step, the RNN architecture is a natural way to model time series data, where each time step of the input depends on those in the past[6].

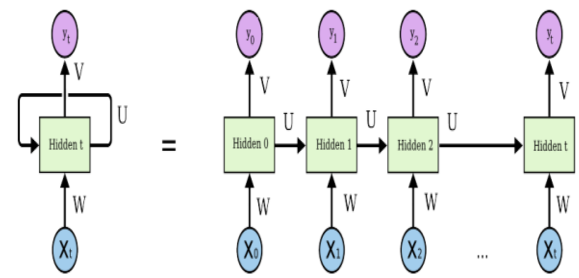

Fig. 3 Example of Unrolling a RNN over the input time steps. Weight sharing across time steps is encapsulated in the recurrent connection.

RNN have one major disadvantage called the vanishing gradient problem , in which the gradient of output error is based on previous inputs vanishes when time lags between inputs and errors increases [7].

3.3 Long Short Term Memory

LSTMs can overcome this problem; it will allow error to propagate throughout the entire network. The long short term memory (LSTM) network called memory cell is capable of learning long-term dependencies, as well as forgetting unnecessary information based on the data at hand[8].
A LSTM network is formed exactly like a simple RNN, except that the nonlinear units in the hidden layers are replaced by memory blocks[9].
Three gates control the information flow into and out of the neuron's memory cell: the input, output, and forget gate.
Each gate in the LSTM gets the same input as the input neuron.1 in LSTM cell h(t) can be considered as a short-term state, and c(t) can be considered as a long-term state . When c(t − 1) point enters into cell, it first goes through a forget gate to drop some memory; then, some new memories are added to it via an input gate; finally, a new output y(t) that is filtered by the output gate is obtained [10].
The structure of LSTM cell is shown in Fig.4, and the operations are calculated by the following equations:

$$i = \sigma(w[x_t, h_{t-1}, C_{t-1}] + b_t) \tag{4}$$

$$C_t = f_t . C_{t-1} + i_{t.}\tanh(w[x_t, h_{t-1}, C_{t-1}] + b_c) \tag{5}$$

$$O_t = \sigma(w[x_t, h_{t-1}, C_t] + b_o) \tag{6}$$

$$h\,(t) = \tanh(C_t). O_t. \tag{7}$$

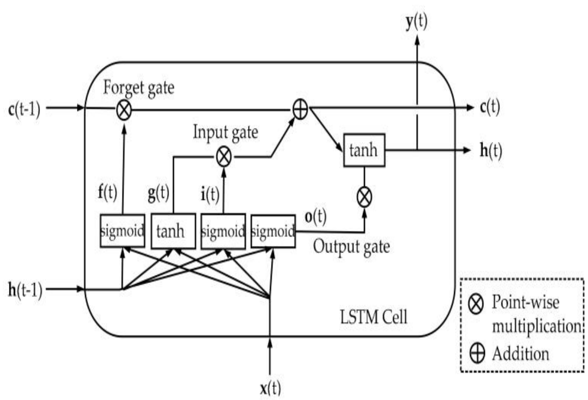

Fig. 4 The structure of Long Short-Term Memory Cell

3.4 Bidirectional LSTMs

BLSTM is an extension of the unidirectional LSTM, which designed to capture the past and the future information of sequential data, and it has two parallel layers propagating in two directions, the forward and backward pass of each layer are carried out in similar way of regular neural networks [11]. Because of its structure, BLSTM can improve LSTM model performance in classification and forecasting processes[12].
Fig. 3 shows a basic BLSTM structure used for running on sequential inputs, which has two different LSTM networks connecting to a common activation layer to produce outputs[12].

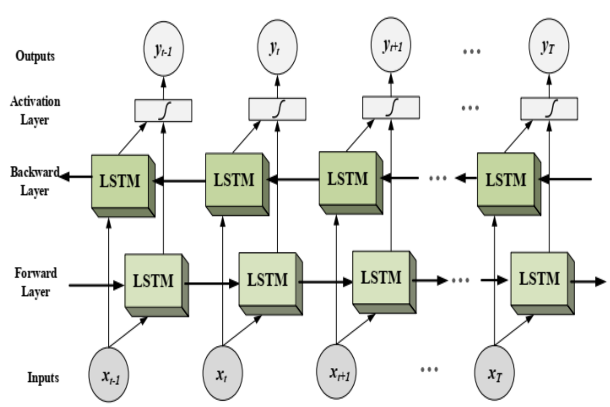

Fig.5 General structure of BLSTM network

4 CASE STUDY HEART RATE VARIABILITY

In this study, heart rate time series is obtained from the MIT-BIH Database that consists of 1800 evenly-spaced measurements of instantaneous heart rate from a single subject [5].
Keras Deep Learning library is *used to build* the BLSTM *model*; it includes one input data, a hidden layer with three LSTM blocks and an output layer that produces a single value prediction. The activation function for LSTM blocks is rectified linear unit (*ReLU*). The network is trained for 50 epochs and a batch size of 32 is used. For preventing Over fitting, a dropout layer was used in this paper, and we set the rate = 0.1. We also used Adam (Adaptive Moment Estimation) as the optimizer and mean square error (MSE) as a loss function to evaluate the prediction accuracy of the proposed model BLSTM based. We split 80% of the data as the training set and the 20% as test set. After training the BLSTM for 50 epochs, we test the performance of our algorithm by predicting the value of testing data and compare them with the actual data. In the output, the orange line represents the actual heart rate time series, while the blue line represents the predicted heart rates. A mean squared error (MSE) of 6.7 is obtained.

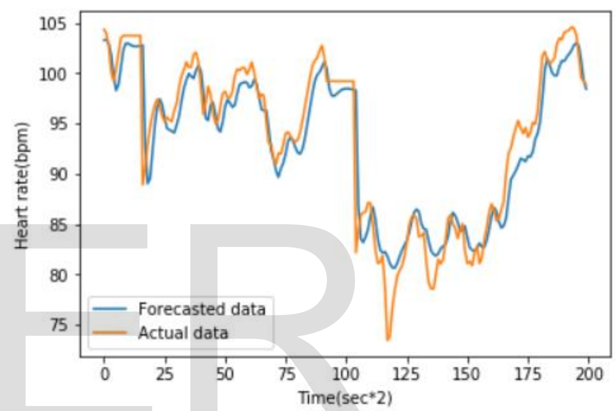

Fig. 6 Actual and forecasted heart rate time series using direct BLSTM

The next process is *Wavelet* transform decomposition which decomposes the original series into one approximation set and several detail sets. We apply the *Daubechies* D8 transform which decomposes the original series into one approximation set and seven detail sets.
Fig.7 shows the output result, the blue line represents the seven levels of wavelet decomposition of the actual heart rate time series, and the red line represents wavelet decomposition of predicted heart rates.

ISSN 2229-5518

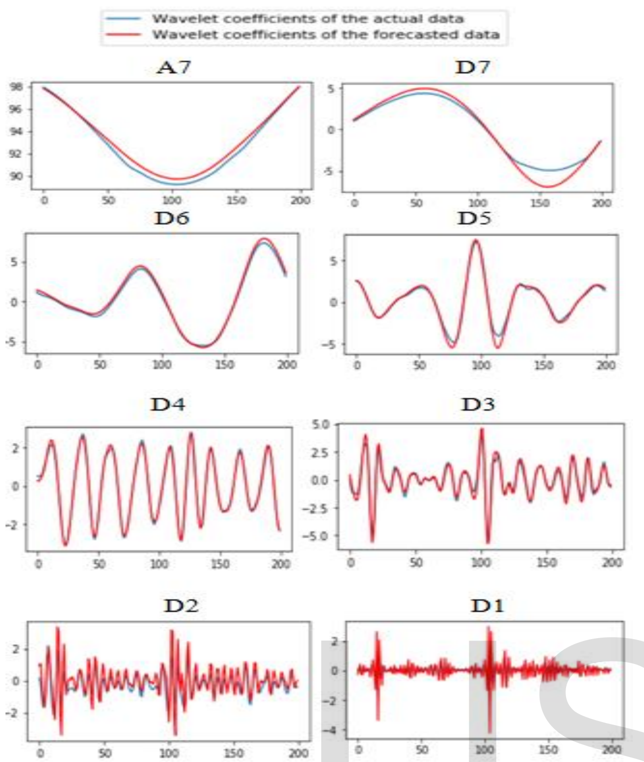

Fig. 7 Seven levels wavelet decomposition of actual and predicted heart rates

Fig. 8 shows the original test data in red and the sum of all coefficients of wavelet decomposition from the previous Fig.7 in blue.

A mean squared error (MSE) of 1.32 is obtained.

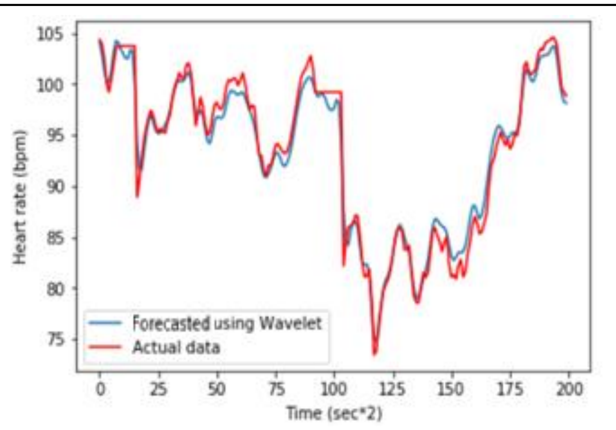

Fig. 8 Actual and forecasted heart rate time series using combination of BLSTM with wavelet decomposition

Discussion

We compared the results of the Fig.7 and Fig.8, as can be seen that prediction using the BLSTM model with wavelet decomposition provides higher performance. It has been able to capture the overall trend compared to direct use of LSTM with actual heart rate time series.

It can be concluded that the wavelet decomposition influence on forecasting performance in BLSTM model by giving a better result.

4 Conclusion

In this paper, a new algorithm for heart rate time series prediction, which is based on a combination of wavelet decomposition and BLSTM neural network, is proposed.

BLSTM neural network which is derived from LSTM network, has advantages in memorizing information for long periods in forward pass and backward pass, has provided recently a significant achievement in the field of machine learning, especially for time series prediction; it was employed in this study to predict the future heart rate variability

The wavelet is used to decompose HRV time series into components of high and low frequency, which provide a good time frequency resolution and more specific information about autonomic activity.

The discussion shows that the combination of Wavelet decomposition using Daubechies D8 and BLSM model gives better result by giving the lower mean squared error.

In future work, we will focus more on the impact of various parameters on forecast accuracy, which related to the Wavelet decomposition including selection of wavelets functions and the decomposition levels. Also we aim to study the hyper-parameters of BLSM model, for example sequence windows size, number of hidden layers and the number of training epochs. So with the help of Wavelets decomposition, we will try to find the optimal parameters of BLSTM to achieve the best performance of prediction.

References

[1] "WHO | World Heart Day," *WHO*, 2018.

[2] M. Elgendi, B. Eskofier, and D. Abbott, "Fast T wave detection calibrated by clinical knowledge with annotation of P and T waves," *Sensors (Switzerland)*, vol. 15, no. 7, pp. 17693–17714, 2015.

[3] G. Kheder, A. Kachouri, M. Ben Massoued, and M. Samet, "Heart Rate Variability Analysis Using Threshold of Wavelet Package Coefficients," vol. 1, no. 3, pp. 131–136, 2009..

[4] R. D. Wallen, "The illustrated wavelet transform handbook," *Biomed. Instrum. Technol.*, vol. 38, no. 4, p. 298, 2004

[5] G. Kheder, A. Kachouri, M. Ben Massoued, and M. Samet, "Heart Rate Variability Analysis Using Threshold of Wavelet Package Coefficients," vol. 1, no. 3, pp. 131–136, 2009.

[6] A. S. Jaffe, "Long short-term memory recurrent neural networks for classification of acute hypotensive episodes," 2017.

[7] Q. Gao, "Stock Market Forecasting Using Recurrent Neural Network," *A Thesis Present. to Fac. Grad. Sch. Univ. Missouri-Columbia*, no. May, 2016.

[8] L. Finsveen, “Time-series predictions with Recurrent Neural Networks Lene Finsveen,” no. June, 2018.

[9] Z. Yu *et al.*, “Using Bidirectional LSTM Recurrent Neural Networks to Learn,” *2015 IEEE Work. Autom. Speech Recognit. Understanding, ASRU 2015 - Proc.*, pp. 338–345, 2016.

[10] D. Wang, Y. Si, W. Yang, G. Zhang, and T. Liu, “A Novel Heart Rate Robust Method for Short-Term Electrocardiogram Biometric Identification,” *Appl. Sci.*, vol. 9, no. 1, p. 201, 2019.

[11] Y. Yao and Z. Huang, “Bi-directional LSTM recurrent neural network for chinese word segmentation,” *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 9950 LNCS, pp. 345–353, 2016

[12] Ö. Yildirim, “A novel wavelet sequences based on deep bidirectional LSTM network model for ECG signal classification,” *Comput. Biol. Med.*, vol. 96, no. March, pp. 189–202, 2018.