

"© 2017 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works."

A CHF Detection Method based on Deep Learning with RR Intervals

Wenhui Chen, Guanzheng Liu, Steven Su, Qing Jiang*, *IEEE Fellow* and Hung Nguyen*, *Senior Member, IEEE*

Abstract—There are extensive studies investigating congestive heart failure (CHF) detection based on heart rate variability. Although a high level of accuracy has been achieved, its robustness under different conditions is not guaranteed. To improve the robustness, we applied sparse auto-encoder-based deep learning algorithm in CHF detection with RR intervals. A total data size of 30,592 (5-min RR interval) was obtained from 72 healthy persons and 44 CHF patients. The deep learning algorithm first extracts unsupervised features using a sparse auto-encoder from raw RR intervals, then constructs a deep neural network model with various hidden nodes combinations. Results showed that the model achieved 72.41% accuracy. This demonstrated that RR intervals have potential in CHF detection but cannot fully reflect dynamic change in 24-h.

I. INTRODUCTION

As a common chronic cardiovascular syndrome, congestive heart failure (CHF) has attracted many researchers' attention. CHF is linked to several complex problems, including end stage heart suffering [1], that cause chambers of the heart to fail. When symptoms suggest CHF, the diagnosis is usually confirmed by physical examination, patient history and various tests to detect abnormal function of the left ventricle and/or heart valves [2]. This diagnosis process often results in time delay and high costs.

Electrocardiogram (ECG) is a commonly used physiological signal in monitoring cardiovascular condition, as it's non-invasive, portable and easy to use. ECG shows both heart rate and rhythm (steady or irregular) [1].

For CHF detection, many researchers focused on ECG analysis, including QRS waves [3], QT intervals [4], RR variability [5], and so on. RR interval is one of the important components of ECG and has attracted great attention as its variability has great potential in CHF prognosis [5]. This variability is either described by RR variability (i.e. variation of RR intervals) or heart rate variability (HRV). HRV based CHF detection has achieved over 95% accuracy with linear/non-linear measures of short-term (5-min)/long-term (24-h) RR intervals data [5]. However, due to its sensitivity along with clinical condition changes [2], the robustness of the HRV-based approaches is still an issue to be addressed. In our previous work [5], we proposed dynamic HRV to describe

dynamic fluctuation of HRV in 24 hours. The analysis revealed that HRV is dynamically changed along with conditions, in accordance with RR intervals fluctuation. Thus, it is necessary to improve the robustness of CHF detection with RR intervals data.

Although there is a guideline [6], manual error and characteristic preprocessing of data are still used. Furthermore, all these former CHF detection methods with HRV are supervised learning. This may require a large amount of labelled data. In this paper we propose to construct an automatic CHF detection without manual extraction and supervision. This process is based on a sparse auto-encoder (SAE)-based deep learning (DL) algorithm.

Deep learning has been applied using unsupervised features [7] and achieved remarkable results in many fields like image recognition [8]. In this work, SAE was applied to learn features automatically from raw unsupervised RR intervals data. Then a deep learning neural network was trained to construct a model to discriminate CHF. In order to improve the neural network performance, various hidden nodes were searched in proportion to input data length. The features in our algorithm are constructed to reflect as much fluctuations in RR intervals.

This work contributes to CHF detection model construction based on SAE-based DL algorithm. It analyses potential mechanism behind between dynamically changed RR intervals. We intend to achieve an automatic method for CHF detection.

The organization of this paper is as follows: in Section II, procedure of model construction of SAE based DL is introduced. The analysis results as well as potential physiological mechanism are discussed in Section III. Finally, Section IV provides the conclusion.

II. METHODS

A. Data

In our work, we applied 116 subjects' RR interval data from PhysioNet [9,10]. All subjects provided informed written consent. The study was approved by the institutional Review Boards of Beth Israel Deaconess Medical Center (Boston, MA) and the Massachusetts Institute of Technology (Cambridge, MA). We downloaded the 24-h RR interval data of 72 healthy persons and 44 CHF patients. The healthy person data came from two databases: the MIT/BIH Normal Sinus Rhythm Database and the Normal Sinus Rhythm RR Interval Database [9]. The data of the CHF patients came from the Congestive Heart Failure RR Interval Database and BIDMC Congestive Heart Failure Database [10]. The data were manually reviewed and corrected by experts.

W. Chen is with the engineering faculty, Sun Yat-sen University, Guangzhou, Guangdong, China and the Center of Health Technologies, Faculty of Engineering and IT, University of Technology Sydney, Ultimo, NSW, Australia (e-mail: Wenhui.Chen@student.uts.edu.au).

G. Liu and Q. Jiang are with the engineering faculty, Sun Yat-sen University, Guangzhou, Guangdong, China (e-mail: liugzh3@163.com; 970747769@qq.com).

S. Su and H.T. Nguyen are with the Center of Health Technologies, Faculty of Engineering and IT, University of Technology Sydney, Ultimo, NSW, Australia (e-mail: Steven.Su@uts.edu.au; Hung.Nguyen@uts.edu.au).

*: dual senior authorship.

B. Deep Learning based CHF detection algorithm

In this work, we applied sparse auto-encoder-based deep learning (SAE-based DL) algorithm for CHF detection (shown in Fig.1).

Sparse auto encoder is a neural network that can learn automatically sparse features from data by minimizing reconstruction error [12]. As one of many famous unsupervised feature learning methods, SAE combines with DL to realize its effectiveness. The aim of this process is to reconstruct input data at the output layer by a sparse penalty term β [13].

Detailed steps of our algorithm are below:

1. Unsupervised feature automatically learning with SAE

Denote input data as $RR = \{rr(i)\}, i = 1, \dots, N, rr(i) \in R^M$. Here N is data sample number of one class, and M is data samples length. First, we input this unlabeled data set into SAE for automatic feature learning to constrain features. At hidden layer l , the feature was expressed as $F(rr(i), W, b), i = 1, \dots, N$, where W denotes the weights between two neighboring layers and b is the bias. This means that the required feature was controlled by input and connection parameters (W, b) .

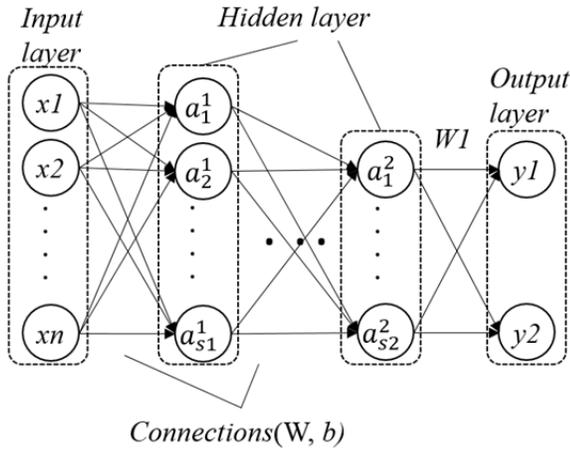


Figure 1. SAE-based DP structure. Each circle of hidden layers is a hidden node; input layer is learnt features with SAE.

The unlabelled segment data RR are first used to train unsupervised features automatically with SAE through the following steps:

- 1) Set up the sparsity penalty term as 3, decay parameter as 10^{-4} and desired average activation of the hidden units as 0.1, and initialize connections W and b randomly close to 0.
- 2) Use batch gradient descent to train the neural network in the forward propagation algorithm to compute the sparse cost function of each iteration of the layer as:

Cost(w, b) =

$$\left[\frac{1}{n} \sum_{i=1}^N \left(\frac{1}{2} \|F_{w,b}(rr(i))\|^2 \right) + \lambda \right] + \beta \sum_{j=1}^{s2} KL(\rho || \rho_j) \quad (1)$$

Here $\rho_j = \frac{1}{n} \sum_{i=1}^N a_j(rr(i))$ is the average activation of hidden unit j ; $KL(\rho || \rho_j) = \rho \log \frac{\rho}{\rho_j} + (1 - \rho) \log \frac{1 - \rho}{1 - \rho_j}$ is a standard measuring function of the difference of the two distributions; λ is the weight decay parameter; $s2$ is the hidden neuron numbers in the hidden layer.

- 3) Update the parameters W and b at each iteration of gradient descent updating:

$$W_{ij}(l) = W_{ij}(l) - \alpha \frac{\partial}{\partial W_{ij}(l)} \text{Cost}(W, b) \quad (2)$$

$$b_i(l) = b_i(l) - \alpha \frac{\partial}{\partial b_i(l)} \text{Cost}(W, b). \quad (3)$$

Here α is the learning rate/step size, optimized by linear search optimizer; l is hidden layer (here is 1).

- 4) Repeat step 2-4 for the second hidden layer for each iteration until iteration is done (set to 300); or we reached the minimum cost value.

2. Deep neural network construction with DL

The feature set F was extracted and fed into DL for classification.

- 1) Select the values of the parameters (W, b) to initialize the DL neural network.
- 2) Set up the training parameters, and conduct the forward propagation algorithm to construct CHF detection model based on softmax classifier.
- 3) Compute the mean square error for the cost function of the DL using Eq. (1).
- 4) Conduct the back-propagation algorithm to update the connections and fine-tune the entire network for each iteration until iteration is done (set to 300); or we reached the minimum cost value.

Thus, a deep neural network was conducted for CHF classification. Finally, the test data set is used to verify the effectiveness of the presented SAE-based DL classifier. Classifier performance is assessed by

$$\text{Accuracy} = \frac{N_{TP} + N_{TN}}{N_{TP} + N_{TN} + N_{FP} + N_{FN}} \quad (4)$$

where N_{TP} is defined as the number of true positive, N_{FN} is the number of false negative, N_{FP} is the number of false positive, and N_{TN} is the number of true negative.

All above steps were done with different hidden node combination to find a suitable network construction. Here, we tested different hidden nodes combination in proportion with input length at range of [10, 200]. Train set and test set were randomly selected with ratio of 1:1. Performance of train set and test set were all verified to avoid over-training.

TABLE I. PERFORMANCE OF TRAIN AND TEST WITH DIFFERENT HIDDEN NODES COMBINATION

TABLE I.1

Hidden nodes of layer 1		200	200	200	200	200	200	150	150	150	150
Hidden nodes of layer 2		200	150	100	50	30	10	150	100	50	30
Train	Accuracy (%)	72.36	72.21	72.48	72.02	72.45	71.27	72.18	71.98	72.58	72.05
	Sensitivity (%)	48.78	49.32	49.41	49.38	42.13	39.83	48.6	48.38	46.64	40.95
	Specificity (%)	85.72	85.18	85.55	84.84	89.62	89.07	85.53	85.35	87.27	89.66
Test	Accuracy (%)	72.07	72.49	72.21	72.44	72.45	71.27	72.13	71.95	72.37	71.39
	Sensitivity (%)	48.83	49.61	48.43	50.39	42.42	39.01	48.18	47.58	46.33	39.97
	Specificity (%)	85.24	85.44	85.68	84.93	89.45	89.54	85.69	85.75	87.12	89.97

Table 1.2

Hidden nodes of layer 1		150	100	100	100	100	50	50	50	30	30
Hidden nodes of layer 2		10	100	50	30	10	50	30	10	30	10
Train	Accuracy (%)	71.18	72.16	71.8	72.79	71.22	72.22	71.7	71.41	70.65	71.37
	Sensitivity (%)	41.19	48.7	41.57	46.55	41.19	48.6	48.58	41.44	47.58	41.12
	Specificity (%)	88.16	85.44	88.92	87.66	89.38	85.6	84.8	88.37	83.72	88.51
Test	Accuracy (%)	71.03	72.41	72.43	72.23	71.28	72.86	72.4	71.23	71.47	71.26
	Sensitivity (%)	41.41	49.09	42.2	45.64	40.01	49.09	49.28	41.67	48.9	40.75
	Specificity (%)	87.81	85.62	89.57	87.29	89.52	86.33	85.49	87.97	84.25	88.53

III. RESULTS AND DISCUSSION

To construct SAE-based deep learning neural network in CHF detection, we downloaded 116 24-h RR interval records. In our work, a neural network was built with hidden node optimization.

Before model construction, all the 24-h data were preprocessed:

- 1) deleting the first and the last interval;
- 2) excluding RR intervals longer than 3 seconds [11];
- 3) dividing 24-h data into multiple 5-min segments as a circle reflecting cardiovascular condition [2] and saved in sequence.

The first two steps were in case of unstable measurement conditions and artificial error; the third step was preparing for CHF detection. After preprocessing, 30592 segments were accessed. After preprocessing, all segments were interpolated into equal length (to biggest segment length 779) according to algorithm requirement.

Then, 5-min RR intervals segments were applied for CHF detection which were reconstructed by SAE algorithm. This reconstructed feature learning happens automatically.

Since the architecture of SAE-based DL, including the size and parameter setting, influences its performance, it is critical to carefully select hidden node numbers and layers. In this work, the hidden layer was chosen as two while hidden nodes of each layer were tested proportionally as shown in Table I. Classification results shown in Table I were around 72% among various hidden node combinations. The train and test

accuracy was calculated as in Table I, which indicates that 72.44% testing accuracy is achievable even without an explicit supervised feature detection stage. The result also shows the variation of testing conditions in a whole day influenced CHF detection. This is consistent with the existing research. In addition, the accuracy of the train and that of the test are very close, indicating that there is no over-training.

Considering trade-off between classifier performance and algorithm effectiveness, a two-layer deep neural network with hidden nodes setting (200,50) was selected as optimal for this study. With this combination structure, the accuracy of train and test were 72.02% and 72.44%, respectively. Performance lost along with descent of nodes. It should be noted that considering the features are automatically extracted in an implicit way, the achieved accuracy is still acceptable.

As many works have shown good results in CHF detection with HRV (described in Section I), RR intervals do maintain useful information for differentiating CHF patients. Though traditional algorithms based on RR interval measurement have been applied in CHF detection and achieved accuracy over 90%, we noticed that differences existed while using long-term (24-h) HRV and short-term (5-min) HRV about its reliability [14]. To the best of our knowledge, there has not been any research which achieves stable CHF detection with only RR variability. This might demonstrate that RR variability is not quite stable as it is sensitive to test conditions, e.g., exercise condition, sleep condition, psychological status and so on [15]. Existing research development is mainly focused on feature extraction methods and statistical analysis [2]. Although these works did reveal potential linkages between disease risk and RR interval based features, it

involves a high level of manual operation, which leads to potential human error and unstable results. Thus, there is an urgent need for a fully automated CHF detection method without manual feature extraction.

In this paper, we applied SAE-based deep learning method into CHF detection. Essentially, features were extracted by the algorithm itself, also called auto-learning. All segments were trained to extract 50 features judged by Equation (1). These 50 learnt features were inputted into a softmax classifier, which is based on a two-layer neural network.

It has been reported that RR variability has potential in CHF detection, even prognosis. In our former work, we introduced dynamic measures of RR variability into CHF detection [5] and concluded that different conditions influence RR intervals classification performance, which might be one of the reasons that the detection accuracy in this study is difficult to improve further. There is no special technology recommended for optimal parameter selection. Thus, parameter setting selection will be the next task in the future to further improve the proposed approach.

It is certain that RR intervals contain information about cardiovascular condition. Actually, when analyzing classification results of segments of each person, it showed that less than 30% segments of 90.28% (65/72) healthy persons has been classified as negative, and higher than 30% segments of 77.27% (34/44) CHF patients has classified as negative. This supports that RR intervals can be used as CHF detection indicator, but this is sensitive to test condition variations. In the future, deep analysis based on our method might help in home monitoring for early detection of CHF.

IV. CONCLUSION

In this paper, a two-layer deep neural network model was constructed for CHF detection based on SAE-based DL algorithm. The results indicated that RR intervals have potential for CHF detection but is sensitive to body condition. Though the classification accuracy for the overall segments only reached 72.41%, further analysis demonstrated the efficiency of the proposed method and interpreted the rationality of the achieved classification results in CHF detection. This also demonstrated the dynamic change with body condition may not fully reflect by only analyzing short/long term RR variability. More analysis on RR intervals is needed. Thus, a dynamic assessment method is required for reliable CHF detection and prognosis. For future work, we will explore the dynamic analysis of the underlying mechanisms.

ACKNOWLEDGMENT

This work was supported by the Joint Institute for Health Technology and Innovation established by Sun Yat-sen University (SYSU) and University of Technology Sydney (UTS).

REFERENCES

[1] T. Kishi, "Heart failure as an autonomic nervous system dysfunction," *Journal of Cardiology*, vol. 59, no.2, pp. 117–122, 2012.

[2] J. J. V. McMurray, S. Adamopoulos, S. D. Anker, A. Auricchio, M. Böhm, K. Dickstein, V. Falk, G. Filippatos, C. Fonseca, M. A. Gomez-Sanchez, T. Jaarsma, L. Køber, G. Y. H. Lip, A. P. Maggioni, A. Parkhomenko, B. M. Pieske, B. A. Popescu, P. K. Rønnevik, F. H. Rutten, J. Schwiter, P. Seferovic, J. Stepinska, P. T. Trindade, A. A. Voors, F. Zannad and A. Zeiher (Task Force Members), "ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure 2012," *European Heart Journal*, vol. 33, no. 14, pp. 1787–1847, Jul 2012.

[3] R. Dhirra, M. J. Pencina, T. J. Wang, B. H. Nam, E. J. Benjamin, D. Levy, M. G. Larson, W. B. Kannel, R. B. D'Agostino and R. S. Vasan, "Electrocardiographic QRS duration and the risk of congestive heart failure - The Framingham heart study," *Hypertension*, vol. 47, no. 5, pp. 861–867, 2006.

[4] L. G. Tereshchenko, I. Cygankiewicz, S. McNitt, R. Vazquez, A. Bayes-Genis, L. Han, S. Sur, J. P. Couderc, R. D. Berger, A. B. de Luna and W. Zareba, "Predictive Value of Beat-to-Beat QT Variability Index Across the Continuum of Left Ventricular Dysfunction Competing Risks of Noncardiac or Cardiovascular Death and Sudden or Nonsudden Cardiac Death," *Circulation-Arrhythmia and Electrophysiology*, vol. 5, no. 4, pp. 719–727, 2012.

[5] W. Chen, L. Zheng, K. Li, Q. Wang, G. Liu and Q. Jiang, "A Novel and Effective Method for Congestive Heart Failure Detection and Quantification Using Dynamic Heart Rate Variability Measurement," *PLoS ONE*, vol. 11, no. 11, pp. e0165304, 2016.

[6] J. J. V. McMurray, S. Adamopoulos, S. D. Anker, A. Auricchio, M. Böhm and K. Dickstein, "ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure 2012 The Task Force for the Diagnosis and Treatment of Acute and Chronic Heart Failure 2012 of the European Society of Cardiology. Developed in collaboration with the Heart Failure Association (HFA) of the ESC," *European Heart Journal*, vol. 33, no. 14, pp. 1787–847, 2012.

[7] A. M. Cheriyyadat, "Unsupervised feature learning for aerial scene classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, pp. 439–451, 2014.

[8] A. Krizhevsky, I. Sutskever and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Adv. Neural Inf. Process. Syst.*, vol. 25, pp. 1106–1114, 2012.

[9] A. L. Goldberger, L.A. Amaral, L. Glass, J. M. Hausdorff, P.C. Ivanov and R. G. Mark, "PhysioBank, PhysioToolkit, and PhysioNet—components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, no. 23, pp. 215–20, 2000.

[10] D. S. Baim, W. S. Colucci, E. S. Monrad, H. S. Smith, R. F. Wright and A. Lanoue, "Survival of patients with severe congestive heart failure treated with oral milrinone," *Journal of the American College of Cardiology*, vol. 7, no. 3, pp. 661–70, 1986.

[11] European Society of Cardiology Task Force and North American Society of Pacing and Electrophysiology, "Heart rate variability: Standards of measurement, physiological interpretation, and clinical use," *Circulation*, vol. 93, no. 5, pp. 1043–65, 1996.

[12] B. A. Olshausen and D. J. Field, "Emergence of simple-cell receptive field properties by learning a sparse code for natural images," *Nature*, vol. 381, pp. 607–609, 1996

[13] W. Sun, S. Shao, R. Zhao, R. Yan, X. Zhang and X. Chen, "A sparse auto-encoder-based deep neural network approach for induction motor faults classification," *Measurement*, Vol. 89, pp. 171–178, July 2016.

[14] F. Tekiner, K. Gemici, B. Emreçan, E. Tekiner and J. Jordan, "The efficacy and prognostic value of heart rate variability in 24-hour and short time recordings for determining cardiac autonomic dysfunction in congestive heart failure," *Anadolu Kardiyoloji Dergisi-the Anatolian Journal of Cardiology*, vol. 7, no. 2, pp. 118–123, 2007.

[15] T. L. Jong, B. Chang and C. D. Kuo, "Optimal Timing in Screening Patients with Congestive Heart Failure and Healthy Subjects During Circadian Observation," *Annals of Biomedical Engineering*, vol. 39, no. 2, pp. 835–849, 2011.