Arrhythmia Classification Algorithm based on Sparse Autoencoder

Mengnan Liang*, Aimin Jiang*, Xiaofeng Liu*, Hon Keung Kwan† and Yanping Zhu‡
* Hohai University, Changzhou, China
E-mail: lmn@hhu.edu.cn, jiangam@hhuc.edu.cn, liuxf@hhuc.edu.cn
† University of Windsor, Windsor, Canada
E-mail: kwan1@uwindsor.ca
‡ Changzhou University, Changzhou, China
E-mail: zhuyanping@cczu.edu.cn

***Abstract*—Cardiovascular diseases are the leading cause of death from noncommunicable diseases worldwide, among which arrhythmias are a common manifestation. Feature extraction is an important part of arrhythmia classification algorithms. Most traditional classification algorithms rely on manual design and extraction of features. In order to improve the efficiency of feature extraction and reduce manual participation, this paper presents a novel and efficient feature extraction framework based on sparse autoencoder, which aims to extract high-dimensional and sparse features through two sparsity regularizers. Features obtained by the autoencoder can be exploited by different classifiers. Experimental results on the MIT-BIH database show that the classification performance of the proposed approach outperforms most of the state-of-the-arts.**

I. Introduction

Cardiovascular diseases are the leading cause of death from noncommunicable diseases worldwide, among which arrhythmias are a common manifestation. About 80% of sudden cardiac deaths are the result of ventricular arrhythmias, which can lead to sudden death or progressive heart failure [1]-[2]. Therefore, accurate and rapid determination of arrhythmias is of great importance. The diagnosis of heart rate arrhythmias mainly relies on electrocardiogram (ECG), and the automatic analysis and diagnosis system of ECG greatly reduces the workload of physicians.

The major steps of ECG-based arrhythmia classification include feature extraction and classification [3]. Regardless of supervised or unsupervised classification methods, feature extraction plays a crucial role. Various features (e.g., time domain features [4], frequency domain features, wavelet morphology features, Stockwell transform (ST) features [5], Hermitian coefficient features [6]) can be exploited to achieve effective features. For example, the RR interval, which is the time between two consecutive heartbeat peaks, is one of features commonly used in this classification task. Liet *et al*. adopted morphological features to identify ECG arrhythmias [7]. Many researchers have exploited multiple features in classification. Li *et al*. employed kernel independent component analysis (ICA) and discrete wavelet transform (DWT) to obtain multi-domain features [8].

Besides traditional features aforementioned, some novel techniques have also been introduced to analyze ECG signals. Kamath C. *et al*. applied the Teager energy operator (TEO) to describe the characteristics of nonlinear components in both time and frequency domains [9]. The main advantage of TEO is that it models the energy of the source that generates ECG signals, rather than that of signals themselves. In this way, any deviation from the regular rhythmic activity of a heart is reflected in the Teager energy function, which is useful when analyzing different classes of ECG signals. The QRS complex reflects the changes in left and right ventricular depolarization potential and timing. The first downward wave in an ECG signal is the Q wave and the upward wave is the R wave, followed by the downward wave as the S wave. Medical analysis of ECG waveforms have revealed that three parameters (i.e., position, width, and amplitude) may be sufficient to describe a QRS waveform population. Therefore, a location, width and magnitude model (LWM) was developed in [10] to synthesize original ECG waveforms. Then, parametric features of the synthesized heartbeats were further used in subsequent classification.

Based on the above extracted features, researchers have developed a series of supervised or unsupervised classification models and algorithms, such as neural networks (NNs) [11], k-nearest neighbor clustering [12], mixture-of-experts [13], classification and regression trees [14]-[15], support vector machines (SVMs) [16], probabilistic NN [17], recurrent NN (RNN) [18] and pathforest [19]. However, such algorithms are highly dependent on hand-crafted features or their combination. With the advent of deep learning [20]-[21], many researchers turn their focus to deep neural networks, that is capable of fulfilling end-to-end feature extraction and classification. X. Yin *et al*. employed one-dimensional convolutional neural network (1D-CNN) to extract complex features from ECG data, that are then fed to a bi-directional long short-term memory (BILSTM) for classification [1]. H. Dang *et al*. proposed a baseline network (network A) and a multi-scale fusion CNN architecture (network B) based on network A to automatically identify five different types of heartbeats [2]. It has been demonstrated in [2] that the multi-scale fusion CNN architecture (network B) is slightly better than network A due to the introduction of a convolution block consisting of three convolution layers, which aggregate features from

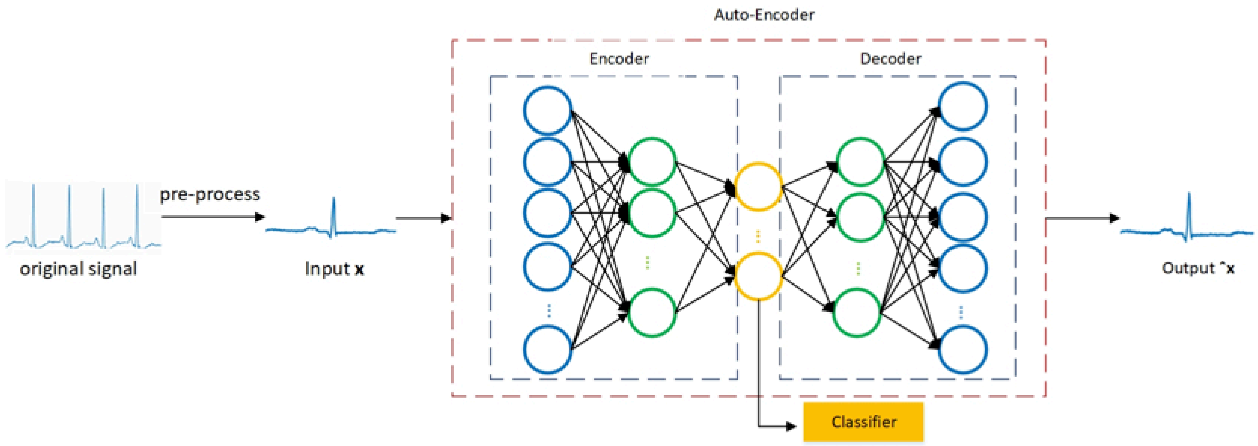

Fig. 1. Network architecture of the proposed AE

all the convolution branches. S. M. Jadhav *et al*. developed a multilayer perceptron (MLP) feedforward neural network model combined with a static back-propagation (BP) algorithm to classify arrhythmia into two categories, normal and abnormal, which guarantees the true estimation of the boundary of complex decision-making [3].

In this paper, we propose an autoencoder (AE) architecture, originally proposed for compressing data [22]. It is similar to but more powerful than principal component analysis (PCA) in terms of feature extraction. The encoder part of the AE is able to learn a low-dimensional representation of ECG signals, while its decoder part aims to reconstruct data by ignoring "noise". The differentiable representation of input ECG learnt in an unsupervised way can be flexibly combined with a variety of classifiers. To further enhance the sparsity of extracted features, L_1 regularization is also introduced into the proposed architecture of the AE.

The paper is organized as follows. In Section II, we first introduce the architecture of the proposed AE. Then, the sparse optimization algorithm is developed. Experimental results are presented in Section III to evaluate the effectiveness of the proposed algorithm. Section IV concludes the paper.

II. Sparse AE

A. Architecture

Fig.1 depicts the proposed network architecture, which consists of preprocessing, encoder and decoder parts. In the preprocessing, QRS waves are first located using the classical Pan-Tompkins algorithm [23]. Then, each heartbeat of fixed length is extracted, resulting in a number of input samples of ECG signals. Assume that there are m labelled samples in the training set $\{(\mathbf{x}_1, y_1), \cdots, (\mathbf{x}_m, y_m)\}$, where $\mathbf{x}_i$ $(i = 1, \ldots, m)$ represents a single QRS wave and $y_i \in \{1, \ldots, c\}$ is a given label corresponding to the heartbeat category. In this paper, we focus on five classes of arrhythmia, i.e., normal (N), left bundle branch block (L), right bundle branch block (R), premature ventricular contractions (V) and atrial premature complexes (A). ECG samples $\{\mathbf{x}_i\}_{i=1}^m$ are directly fed to the encoder network, generating representative features, which are further used in the decoder network to reconstruct segmented ECG signals. The whole network is trained by minimizing the reconstruction loss. Once the AE network training is completed, representative features are employed as inputs to train a classifier. More details regarding each part of the AE are given in the following subsections.

Suppose that the AE network is composed by n_l layers. In the coding phase, the output of the l-th layer is given by

$$\mathbf{h}^{(l)}(\mathbf{x}_i) = \sigma\left(\mathbf{W}^{(l)}\mathbf{h}^{(l-1)} + \mathbf{b}^{(l)}\right) \tag{1}$$

where $\sigma(\cdot)$ represents an activation function, $\mathbf{W}^{(l)} \in \mathbb{R}^{s_l \times s_{l-1}}$ and $\mathbf{b}^{(l)} \in \mathbb{R}^{s_l}$ denote, respectively, the connection weight matrix between two consecutive layers and the bias vector associated to each unit in the l-th layer, s_l represents the number of units in the l-th layer. On the input layer, $\mathbf{h}^{(0)} = \mathbf{x}_i$. In our architecture, $\sigma(\cdot)$ is always chosen as Sigmoid function. The decoding part performs the inverse conversion of the encoding process.

B. Sparse AE

To avoid trivial solutions, sparsity regularizers can be introduced to the reconstruction loss. In practice, two ways are under consideration.

1) KL divergence : The reconstruction loss of the AE network is first formulated as

$$J(\mathbf{W}, \mathbf{b}) = \frac{1}{2m}\sum_{i=1}^{m}\left\|\mathbf{h}^{(n_l)}(\mathbf{x}_i) - \mathbf{x}_i\right\|_2^2 + \frac{\lambda}{2}\sum_{l=1}^{n_l-1}\left\|\mathbf{W}^{(l)}\right\|_F^2, \tag{2}$$

where the first term represents the reconstruction error given each training sample $\mathbf{x}_i$, while the second one is the weight attenuation term, which is generally used to reduce the size of weights and prevent the over-fitting. The weight decay parameter λ controls the relative importance of these two terms. The optimization of the network is achieved generally by the stochastic gradient descent [24].

TABLE I
PERFORMANCE EVALUATION

Heartbeat category	AE(L_1)+SVM			AE(L_1)+BP		
	$Se\,(\%)$	$Sp\,(\%)$	$P+(\%)$	$Se\,(\%)$	$Sp\,(\%)$	$P+(\%)$
Nb	**97.35**	**97.44**	**92.11**	93.63	93.02	79.39
Lb	97.45	**97.41**	**91.70**	**98.00**	91.78	77.25
Vb	97.93	**97.28**	**90.80**	**98.74**	91.65	76.13
Rb	**95.74**	**97.89**	**92.73**	89.78	94.22	82.99
Ab	**100**	**97.12**	**80.20**	78.11	94.89	63.86
$Acc\,(\%)$		**97.42**			93.16	

It is believed that high-dimensional but sparse representation is useful in both compression and classification tasks. To this end, given input $\mathbf{x}_i$, we first compute the activation value $a_j^{(l)}(\mathbf{x}_i)$ of hidden unit j in the l-th layer. For simplicity, in practice, it can be set equal to Sigmoid function value, that is the j-th element of $\mathbf{h}^{(l)}(\mathbf{x}_i)$. Then, the average activation value is given by

$$\hat{\rho}_j^{(l)} = \frac{1}{m}\sum_{i=1}^{m} a_j^{(l)}(\mathbf{x}_i). \tag{3}$$

To make $\hat{\rho}_j$ as small as possible, a sparsity limitation ρ is further specified, which is close to 0 (e.g., $\rho = 0.05$ in our experiments). Then, the KL divergence is exploited to evaluate the sparsity of the whole network, yielding

$$J_{KL}(\mathbf{W}, \mathbf{b}) = J(\mathbf{W}, \mathbf{b}) + \beta \sum_{l=1}^{n_l-1}\sum_{j=1}^{s_l} KL\left(\rho \middle|\middle| \hat{\rho}_j^{(l)}\right), \tag{4}$$

where β controls the relative importance of the sparsity penalty term, and

$$KL\left(\rho \middle|\middle| \hat{\rho}_j^{(l)}\right) = \rho \log \frac{\rho}{\hat{\rho}_j^{(l)}} + (1-\rho)\log\frac{1-\rho}{1-\hat{\rho}_j^{(l)}}. \tag{5}$$

2) L_1 regularization : L_1 regularizer employs L_1-norm to sparsify network weights, while the KL divergence aims to locate insignificant units that contribute much less to the final reconstruction than the other ones. Using L_1-norm, the cost function is defined as

$$J_{L_1}(\mathbf{W}, \mathbf{b}) = \frac{1}{2m}\sum_{i=1}^{m}\left\|\mathbf{h}^{(n_l)}(\mathbf{x}_i) - \mathbf{x}_i\right\|_2^2 + \frac{\lambda}{2}\sum_{l=1}^{n_l-1}\left\|\mathbf{W}^{(l)}\right\|_1, \tag{6}$$

where $\left\|\mathbf{W}^{(l)}\right\|_1 = \sum_{i=1}^{s_l}\sum_{j=1}^{s_{l-1}}\left|W_{ij}^{(l)}\right|$. Minimizing $J_{L_1}(\mathbf{W}, \mathbf{b})$ forces some weights equal to small values, that can be directly discarded.

The AE network is trained by the back-propagation algorithm [25], which involves the computation of gradients with respect to $\mathbf{W}$ and $\mathbf{b}$ using different sparse regularizers.

C. Classification

Once representative features of training samples $\{\mathbf{x}_i\}$ are obtained by the AE network, classifiers can then be trained. Two classes of classifiers are used in our experiments.

1) SVM : The basic idea of SVM is to find the best separating hyperplane on the feature space, that maximizes the gap of positive and negative sample margins. As suggested in [8], genetic algorithm is also employed to optimize parameters of the SVM classifier.

2) BP NN : In our experiments, a three-layer BP NN is also exploited to classify ECG samples. Sigmoid function is still chosen as activation functions in the network. The last layer outputs the predicted probability of each category.

III. EXPERIMENTAL RESULTS

A. Experimental Setup

In our experiments, we adopt the MIT-BIH Database [26] for performance evaluation. This database consists of 48 two-leads ECG recordings, each approximately half-hour long for each record with sampling frequency 360 Hz. In the preprocessing, a bandpass filter processed the signal to reduce interference, and dual-threshold processing was used to segment the ECG into single heartbeat [23]. In the training process, based on empirical values, λ and β are set equal to 0.004 and 4, respectively. The number of hidden layer units in the encoder is 50 and 25.

B. Performance Evaluation

The classification performance of the proposed algorithm are evaluated y four metrics: sensitivity Se, specifity Sp, positive predictivity $P+$ and accuracy Acc

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

$$Sp = \frac{TN}{TN+FP} \tag{8}$$

$$P+ = \frac{TP}{TP+FP} \tag{9}$$

$$Acc = \frac{TP+TN}{TP+TN+FP+FN} \tag{10}$$

where TP, TN, FP and FN denote, respectively, the True Positive, the True Negative, the False Positive and the False Negative.

Experimental results obtained by different classifiers are listed in Table I. It can be found that the sparse AE using SVM achieves better Sp and $P+$. For example, $P+$ can be improved, respectively, by 12.72%, 14.45%, 14.67%, 9.74% and 16.34% for five heartbeat categories when using the SVM classifier. The BP classifier achieves better Se in the Lb and Vb types of heartbeats. Overall speaking, the AE network using the SVM classifier performs better in terms of classification accuracy.

TABLE II
COMPARISON OF CLASSIFICATION ACCURACIES OF DIFFERENT ALGORITHMS

Heartbeat category	AE(L_1)+BP	AE(KL)+BP	ICA+BP	DWT+BP	LC-KSVD
Nb	**95.04**	93.44	88.14	58.24	87.45
Lb	**98.32**	96.72	76.01	66.13	92.71
Vb	95.78	**97.17**	73.75	61.57	93.37
Rb	**92.15**	92.01	76.28	65.19	83.35
Ab	79.33	81.18	59.45	80.03	**91.32**
Acc	**93.56**	93.23	77.67	63.04	89.04
Heartbeat category	AE(L_1)+SVM	AE(KL)+SVM	ICA+SVM	DWT+SVM	CNN
Nb	**98.38**	95.83	96.82	83.56	86.46
Lb	**98.09**	96.27	92.07	61.34	94.81
Vb	**97.64**	96.73	81.94	96.00	90.83
Rb	**96.51**	94.18	83.21	81.44	92.13
Ab	**100.0**	**100.0**	**100.0**	**100.0**	90.25
Acc	**97.96**	96.15	89.33	78.82	90.90

C. Performance Comparison of Existing Algorithms

In this set of experiments, classification accuracies Acc of the proposed algorithm are compared with those of four state-of-the-arts, including ICA [8], DWT [8], Label Consistent K-Singular Value Decomposition (LC-KSVD) algorithm [27] and CNN [2]. Except the LC-KSVD and CNN, the other feature extraction methods are evaluated using both BP and SVM. In our experiments, 10×10-fold cross validation is employed.

Table II lists classification accuracies of five different types of heartbeats (i.e., Nb, Lb, Vb, Rb, Ab) using aforementioned algorithms. For different types of heartbeats, the best results are shown in bold. It can be seen that the average accuracy of the proposed algorithm is better than the other approaches, especially when the AE network is equipped with SVM as the classifier. The average accuracy increases by at least 1.81%. Experimental results also indicate that features extracted by the AE using L_1 regularizer are more effective in the classification task.

IV. CONCLUSIONS

In this paper, an effective algorithm using sparse AE for feature extraction has been proposed for arrhythmia classification. Compared with the traditional methods of manually designing and extracting features, it can fulfill automatic feature extraction, which is capable of achieve higher classification accuracy. Two sparsity regularizers have been introduced into the architecture of the AE network. Experimental results have demonstrated the superior performance of the proposed algorithm over the state-of-the-arts.

ACKNOWLEDGMENT

This work was supported in part by the National Nature Science Foundation of China under grant 61801055, and the Fundamental Research Funds for the Central Universities under grant 2018B23014.

REFERENCES

[1] X. Yin and C. Hang, "Arrhythmia classification based on CNN_BILSTM,"in *Proc. 6th Int. Conf. Syst. Informat. (ICSAI)*, Nov. 2019, pp. 1105-1109, 2019.

[2] H. Dang, M. Sun, G. Zhang, X. Zhou, Q. Chang and X. Xu, "A novel deep convolutional neural network for arrhythmia classification," in *Proc. Int. Conf. Adv. Mech. Syst. (ICAMechS)*, Aug. 2019, pp. 7-11, 2019.

[3] S. M. Jadhav, S. L. Nalbalwar and A. Ghatol, "Artificial neural network based cardiac arrhythmia classification using ECG signal data," in *Proc. Int. Conf. Electron. Inf. Eng.*, vol. 1, 2010, pp. V1-228–V1-231.

[4] C. Ye, B. V . K. V . Kumar, and M. T. Coimbra, "Heartbeat classification using morphological and dynamic features of ECG signals," *IEEE Trans. Biomed. Eng.*, vol. 59, no. 10, pp. 2930–2941, Oct. 2012.

[5] M. K. Das and S. Ari, "Patient-specific ECG beat classification technique," *Healthcare Technol. Lett.*, vol. 1, no. 3, pp. 98–103, 2014.

[6] T. H. Linh, S. Osowski, and M. Stodolski, "Online heartbeat recognition using Hermite polynomials and neurofuzzy network," *IEEE Trans. Instrum. Meas.*, vol. 52, no. 4, pp. 1224–1231, Aug. 2003.

[7] P . Liet al., "High-performance personalized heartbeat classification model for long-term ECG signal," *IEEE Trans. Biomed. Eng.*, vol. 64, pp. 78–86, Jan. 2017.

[8] H. Li, D. Yuan, Y . Wang, D. Cui, and L. Cao, "Arrhythmia classification based on multi-domain feature extraction for an ECG recognition system," *Sensors*, vol. 16, no. 10, pp. 1744, 2016.

[9] Kamath C ," ECG beat classification using features extracted from Teager energy functions in time and frequency domains," *IET Signal Proc.*, vol. 5, no. 6, pp. 575-581, 2011.

[10] Zhu J , He L , Gao Z . "Feature extraction from a novel ECG model for arrhythmia diagnosis," *Bio-Med. Mater. Eng.*, vol. 24, no.6, pp. 2883-91, 2014.

[11] P . Ghorbanian, A. Ghaffari, A. Jalali, and C. Nataraj, "Heart arrhythmia detection using continuous wavelet transform and principal component analysis with neural network classifier," *Proc. Comput. Cardiology.*, pp. 669–672, Sep. 2010.

[12] S. Kiranyaz, T. Ince, J. Pulkkinen, and M. Gabbouj, "Personalized long-term ECG classification: A systematic approach," *Expert Syst. Appl.*, vol. 38, no. 4, pp. 3220–3226, Apr. 2011.

[13] Y . H. Hu, S. Palreddy, and W. J. Tompkins, "A patient-adaptable ECG beat classifier using a mixture of experts approach," *IEEE Trans. Biomed. Eng.*, vol. 44, no. 9, pp. 891–900, Sep. 1997.

[14] J. Fayn, "A classification tree approach for cardiac ischemia detection using spatiotemporal information from three standard ECG leads," *IEEE Trans. Biomed. Eng.*, vol. 58, no. 1, pp. 95–102, Jan. 2011.

[15] L. Pecchia, P . Melillo, and M. Bracale, "Remote health monitoring of heart failure with data mining via CART method on HRV features," *IEEE Trans. Biomed. Eng.*, vol. 58, no. 3, pp. 800–804, Mar. 2011.

[16] S. Osowski, T. Hoai, and T. Markiewicz, "Support vector machine-based expert system for reliable heartbeat recognition," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 4, pp. 582–589, Apr. 2004.

[17] J.-S. Wang, W.-C. Chiang, Y.-L. Hsu, and Y.-T. C. Yang, "ECG arrhythmia classification using a probabilistic neural network with a feature reduction method," *Neurocomputing*, vol. 116, pp. 38–45, Sep. 2013.

[18] E. D. Ibeyli, "Recurrent neural networks employing Lyapunov exponents for analysis of ECG signals," *Expert Syst. Appl.*, vol. 37, no. 2, pp. 1192–1199, Mar. 2010.

[19] E. J. da S. Luz, T. M. Nunes, V . H. C. De Albuquerque, J. P . Papa, and D. Menotti, "ECG arrhythmia classification based on optimum-path forest," *Expert Syst. Appl.*, vol. 40, no. 9, pp. 3561–3573, Jul. 2013.

[20] Y. Yu, M. Li, L. Liu, Y. Li and J. Wang, "Clinical big data and deep learning: Applications, challenges, and future outlooks," *Big Data Mining Analytics,* vol. 2, no. 4, pp. 288–305, 2019.

[21] X. Chen and X. Lin,“Big data deep learning: Challenges and perspectives,” *IEEE Access*, vol. 2, pp. 514–525, May 2014.

[22] J. Liu, C. Li and W. Yang, “Supervised learning via unsupervised sparse autoencoder,” *IEEE Access*, vol. 6, pp. 73802–73814, 2018.

[23] J. Pan and W. J. Tompkins, “A real-time QRS detection algorithm,” *IEEE Trans. Biomed. Eng.*, vol. BME-32, no. 3, pp. 230-236, March 1985.

[24] A. Moussavi-Khalkhali and M. Jamshidi, “Constructing a deep regression model utilizing cascaded sparse autoencoders and stochastic gradient descent,” in *Proc. 15th IEEE Int. Conf. Mach. Learn. Appl.*, Dec. 2016, pp. 559-564.

[25] R. Jiao, X. Huang, X. Ma, L. Han and W. Tian, “A model combining stacked autoencoder and back propagation algorithm for short-term wind power forecasting,” *IEEE Access*, vol. 6, pp. 17851-17858, 2018.

[26] G. B. Moody and R. G. Mark, “The impact of the MIT-BIH arrhythmia database,” *IEEE Eng. Med. Biol. Mag.*, vol. 20, no. 3, pp. 45-50, May-June 2001.

[27] S. M. Mathews, L. F. Polanía and K. E. Barner, “Leveraging a discriminative dictionary learning algorithm for single-lead ECG classification,” in *Proc. Biomed. Eng. Conf.*, 2015, pp. 1–2.