Feature Learning Using Convolutional Neural Network for Cardiac Arrest Detection

Minh Tuan Nguyen
School of Electrical Engineering and Computer Science
Gwangju Institute of Science and Technology
Gwangju, South Korea
nguyenminhtuan@gist.ac.kr

Kim Kiseon
School of Electrical Engineering and Computer Science
Gwangju Institute of Science and Technology
Gwangju, South Korea
kskim@gist.ac.kr

***Abstract*—Arrhythmias including ventricular fibrillation and ventricular tachycardia, which are known as shockable rhythms, are the mainly cause of sudden cardiac arrests (SCA). In this paper, we propose a feature learning scheme applied for detection of SCA on electrocardiogram signal with the modified variational mode decomposition technique. The subsequent SAA consists of a convolutional neural network as a feature extractor (CNNE) and a support vector machine classifier. The features extracted by selected CNNE are then validated using 5-folds CV procedure on the evaluation data, and enable the accuracy of 99.02 %, sensitivity of 95.21 %, and specificity of 99.31 %.**

Keywords—Sudden cardiac arrest (SCA), Machine learning (ML), Deep learning (DL), Convolutional Neural Network (CNN).

I. Introduction

Sudden cardiac arrest (SCA) is heart ailment and automated external defibrillator (AED) helps to recover the normal sinus rhythms of the heart from the cardiac electrical activity [1]. The shock advice algorithms (SAA) applied for the AED in the literature to detect the SCA have counted on conventional feature extraction (FE) schemes. The input features are widely extracted from the stand-alone electrocardiogram (ECG) [1], [2], alternative signals [3], [4] based on clinical expertise. Subsequently, the performance of the classification models is significantly dependent on the quality of the extracted input feature space and the final feature combinations. This is accomplished by adopting the feature selection (FS) algorithms to eliminate the irrelevant features from the input feature space and improve the learning process of the ML classifiers [5], [6].

Deep learning, especially convolutional neural network (CNN), has been widespread applied for biomedical signal processing and application design problems due to its strong feature learning capabilities, no requirement of expertise-based FE and FS algorithms [7]. Indeed, the first research adopting deep learning technique for SCA detection is proposed in [7] in which the SAA is designed as the eleven-layer deep CNN. However, the detection performance of this SAA does not meet the American Heart Association recommendations for the AED. In general, previous works only pay attention to a specific CNN structure using the stand-alone ECG signal, which leads possibly to omit other effective structures of the CNN. Moreover, the learned features extracted by the CNN as a feature extractor have not been investigated for the input of the ML classifiers with respect to improvement of SCA detection performance.

In this paper, a novel FE and SAA is proposed for detection of SCA, which uses the feature set extracted by the CNN extractor (CNNE) as the input of the SVM classifier. Here, the modified (MVMD) technique is used to reconstruct SH and NSH signals, which have most of power of SH and NSH components inside the bandwidths below and above 10 Hz, respectively, from preprocessed ECG signal [6], [8]. Three above signals are then arranged as the input channels of the CNN.

II. Data, CNN, and SVM

A. Data

We used public databases namely the Creighton University Ventricular Tachyarrhythmia Database (CUDB) and the MIT-BIH Malignant Ventricular Arrhythmia Database (VFDB) similarly in [6]: The annotations are VF, VT, and non-VF in CUDB. Ventricular flutter, VF, and VT are annotated for SH rhythms and others for NSH rhythms in VFDB database. Moreover, artifacts, noise, asystole, transition rhythm, low peak-to-peak amplitude under 200 μV of VF and intermediate rhythms of slow VT for which rate under 150 beats per min are also eliminated from the database. Indeed, the defibrillation archives no benefit for these rhythms, which are neither SH nor NSH rhythms [5]. Removal of these rhythms makes the requirements of algorithm performance appropriate [6] and ensures that the algorithm can identify correct rhythms. TABLE I shows the number of 5s-segments and records used for training and evaluation data.

The databases are preprocessed with five-order moving average filtering to make them smooth. Then, bandpass filter with cutoff frequencies of 1 Hz and 30 Hz is applied to eliminate the base line wander and high frequency.

B. *CNN* (Convolution Neural Network)

The CNN consists of different layers, which are input, convolutional, rectified linear unit, max-pooling layer, fully connected layers, and output layer [10].

Input layer (Inp): The preprocessed ECG segment, SH and NSH signals are arranged as three input channels, which are fed into the CNN.

Convolutional layer (Conv): The Conv layer plays a role of feature detection at different positions in the input signals. The primary element of this layer is the filter, which is a line-shaped object that scans over the input signal to produce an activation map.

TABLE I. Details Of Training And Evaluation Data Of 5S-Segment **Length**

Databases	Training data				Evaluation data				Total			
	SH	NSH	Total	Record	SH	NSH	Total	Record	SH	NSH	Total	Record
VFDB	1059	3756	4815	15	238	2395	2633	7	1297	6151	7448	22
CUDB	489	1711	2200	25	90	538	628	10	579	2249	2828	35
Total	1548	5467	7015	40	328	2933	3261	17	1876	8400	10276	57

Rectified linear unit layer (ReLu): This layer is a nonlinear activation function to speed up the learning convergence by mapping nonlinearity into the data.

Max-pooling layer (MP): To reduce computational complexity, the MP layer is used to subsample the feature maps of the preceding Conv by computing the average or maximum responses in small overlapping neighborhoods.

Fully connected layer (FC): The output of previous layer is connected to every unit in the FC layer, which perform high level inference following by a nonlinear transformation. This layer can be viewed as a special case of Conv with one by one filter size.

Output layer (Out): The output layer contains the Softmax (SM) and classification layers. The SM is used to encode the probability distribution of a specific class defined by the corresponding unit when the CNN is considered as a classifier. The classification layer then labels the output as SH or NSH segments.

C. *SVM* (Support Vector Machine)

The support vector machine (SVM) algorithm with Gaussian kernel is used to construct the optimal model with its corresponding feature set *[11]*. The x_i vector of features is arranged as$\{(t_1, x_1), \dots, (t_n, x_n)\} \in \{0,1\} \times R^m$, where t_i is the label of each vector with $t_i = 1$ for SH segment and $t_i = 0$ for NSH segment. n is for the number of segments and R^m is for the feature space with number of features m. The discriminant function for a segment with feature vector x is given by:

$$f(x) = \sum_{j=1}^{S} \left(\alpha_j t_j e^{-\gamma} \|x - x_j\|^2\right) + b \qquad (1)$$

Where,   are the support vector (SV) j and S is the number of SVs. The   and b are coefficients, which are estimated during training stage. The segments are classified as shock if *f(x) > 0* or non-shock if *f(x) ≤ 0.*

III. METHOD

The procedures of our method are shown in Fig. 1 including channel construction, CNN selection, and CNN validation phases.

In the channel construction phase, the ECG databases are segmented and preprocessed. Three channels used as 3 input of the CNN are then generated by the MVMD technique. In the CNN selection phase, a grid search with nested 5-folds CV is implemented to select a CNNE using support vector machine (SVM) classifier. In the CNN validation phase, the SVM classifier using feature vector extracted by the selected CNNE are validated on evaluation data using 5-folds CV procedure.

A. *Channel Construction*

There are three input channels of the CNN, which are preprocessed ECG segment, SH and NSH signals. According to [8], the bandwidths of SH and NSH rhythm complexes are below 10 Hz and over 12 Hz, respectively. Moreover, the NSH components are concentrated on the bandwidth from 13 Hz to 17 Hz [10]. Taking into account the characteristics of power spectrum, the MVMD technique suggested in [6] is used by preselecting 4 values of center frequencies within the bandwidth below 10 Hz to reconstruct 4 SH modes. It is noteworthy that we use the MVMD to decompose the preprocessed ECG segment into 10 modes, and the f rst mode is discarded because its zero center frequency is chosen when starting the MVMD algorithm [6].

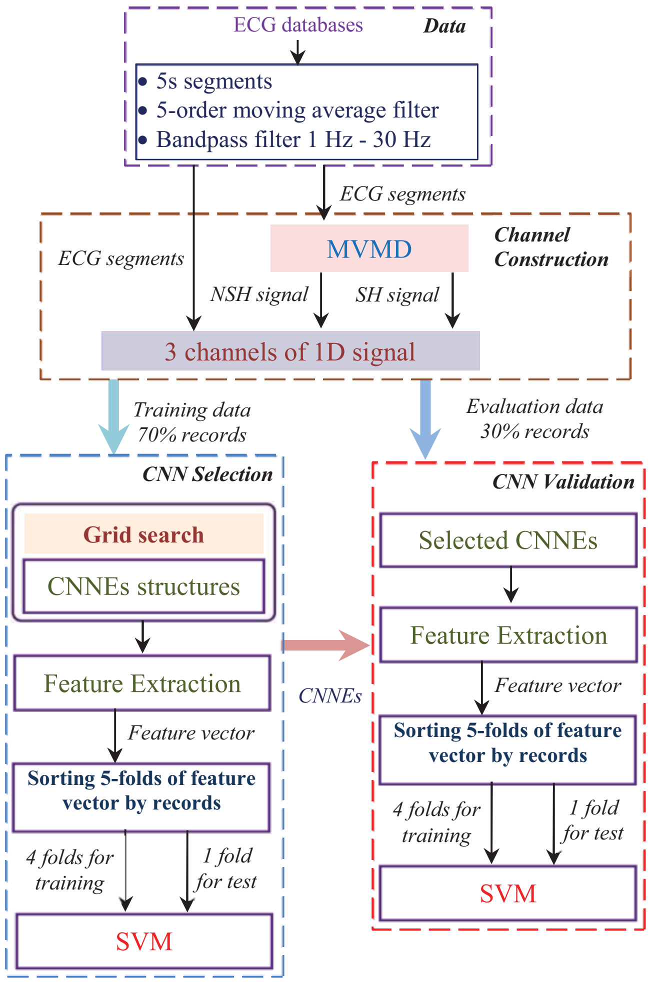

Fig. 1. The Overview Of Method Procedres

B. CNN Selection

A grid search with nested 5-folds CV is deployed to select the best CNN structure in terms of its learning and structure parameters for the CNNEs using the SVM classifiers on the training data. It is well-known that there is no one fitted CNN structure for all biomedical problems. Each CNN structure requires a set of parameters for specific problem and database. The CNNE structure and its parameters are selected for the validation phase based on the minimum BER of corresponding SVM classifier on the training data.

Learning parameters [7]

- Regularization (Re): This parameter prevents the network from becoming overly dependent on any one neuron called overfitting, which is common problem in neural networks.
- Momentum (Mo): A learning property, which causes the weight change to keep going continuously in current direction. This parameter controls the learning speed of the network.
- Learning rate (Lr): To determines how far each iteration of training will take the weight values to help in convergence of the data.

Structure parameters [12]

- Network depth (Nd): Two consecutive layers of a Conv and a ReLu are defined as a block. Nd is number of consecutive blocks, which represent the depth of the network.
- Network section (Ns): Each NS consists of numerous blocks, which are followed by a MP layer. Indeed, the use of multiple Conv layers in a section of network improves the data representations before losing the information due to decrease in data dimensions on MP layer.

For the simplicity of problem formulation, we only investigate above 5 CNN parameters within specific ranges: Ns = {1, 2, 3, 4}; Nd = {1, 2, 3}; Lr = {0.001, 0.005, 0.01}; Mo = {0.7, 0.8, 0.9, 0.95}; and Re = {0.1, 0.15, 0.2}. Other parameters are selected as conventional values of these parameters, which are commonly used in previous researches [7].

It is noteworthy that we use two consecutive FC layers for which the output of last FC represents the binary classes. Moreover, The FC layer determines the nonlinearity of the high level features extracted by previous layer. Therefore, the first FC is used to extract the feature vector, which is then fed into various ML classifiers.

C. CNN Validation

The 5-folds CV procedure is implemented by dividing evaluation data into 5 parts by record for which 4 parts are used as training data and remaining part is for testing data. The completed process contains 5 runs to ensure that every single part is used as the testing data. This procedure is used to validate the performance of SVM classifier using a feature vector, which is extracted by the CNNE from the evaluation data. A repetition of 100 times is applied for the 5-folds CV procedure to compute the mean and standard deviation of the performances of the SVM classifier.

IV. RESULTS AND DISCUSSIONS

We use 4 main measures, which are accuracy (Ac), sensitivity (Se), specificity (Sp), and balanced error rate (BER) to estimate the performance of the ML classifiers using the feature vector extracted by the CNNE as shown in [6].

A. CNN Selection

There are 108 specific CNN structures for each Network Section with different values of parameters investigated on the training data using nested 5-folds CV. TABLE II shows the CNNEs corresponding to different Network sections and best values of parameters, which are selected on the 5s-segment training data.

TABLE II. CNN EXTRACTORS CORRESPHONDING TO NETWORK SECTION AND SELECTED PARAMETERS

CNN extractors	**Ns**	**Nd**	**Lr**	**Mo**	**Re**
CNNE1	1	2	0.01	0.9	0.1
CNNE2	2	1	0.005	0.7	0.1
CNNE3	3	1	0.005	0.8	0.15
CNNE4	4	1	0.005	0.7	0.1

TABLE III. PERFORMANCE OF SVM CLASSIFIER USING THE FEATURE VECTORS EXTRACTED BY CNNEs FROM 5S-SEGMENT EVALUATION DATA

CNN extractors	**Ac (%)**	**Se (%)**	**Sp (%)**	**BER (%)**
CNNE1	98.73±0.54	93.84±1.45	99.10±0.71	3.53±0.81
CNNE2	98.80±0.38	94.78±1.17	99.19±0.41	3.01±0.62
CNNE3	**99.02±0.32**	**95.21±1.28**	**99.31±0.38**	**2.74±0.65**
CNNE4	98.87±0.34	95.11±1.01	99.15±0.43	2.87±0.54

B. CNN Validation

TABLE III presents the validation performance results of the SVM classifier using feature vectors extracted by the CNNEs on evaluation data. Generally, the SVM classifiers show relative high validation performance for all CNNEs and meet the AHA recommendations for the AED. The highest performance is produced by the SVM classifier using the feature vector extracted by the CNNE3.

C. Discussions

The first research applying CNN for detection of SH/NSH rhythms is proposed in [7]. However, the performance does not meet the AHA recommendations, which are Ac of 93.18%, Se of 95.32%, and Sp of 91.04%. This may be due to only a full CNN structure of 11 layers is investigated with stand-alone ECG segment as the input channel. Moreover, the authors of [7] do not pay attention to the CNN as the feature extractor, which can generate effective feature vector used as the input of the machine learning classifiers.

In our research, we investigate different CNN structures and parameters by employing an exhaustive grid search with nested 5-folds CV to select the best values of the learning and structure parameters among conventional values. The highest validation performance is archived for the SVM classifier with Ac of 99.02%, Se of 95.21%, and Sp of 99.31%

on evaluation data, which is higher than that in [7]. TABLE IV shows the details of the selected CNNE.

A larger number of methods are proposed for classification of SH/NSH rhythms using ML technique, conventional feature extraction, and feature selection algorithms, which are time consuming and require expertise human knowledge. Inconstrast, the CNN does not require those algorithms or any knowledge of heart disease but improve certainly the final classification in terms of SH/NSH rhythm detection.

TABLE IV.DETAILS OF CNN3 STRUCTURE

Order	Layer	Filter size	Filter number	Stride	Padding
1	Input	1250x3 channels			
2,3	Convolution+ReLU	101x1	10	1	50
4	Max Pooling	11x1		2	
5,6	Convolution+ReLU	101x1	20	1	50
7	Max Pooling	11x1		2	
8,9	Convolution+ReLU	101x1	40	1	50
10	Max Pooling	11x1		2	
11	Fully Connected 1	100 ouputs			
12	Fully Connected 2	2 outputs			
13	Soft max				
14	Classification	2 classes			

V. CONCLUSION

In this work, we proposed a set of feature and a SAA applied for AED applications using the combination of machine learning and deep learning techniques. The CNN is used as the feature extractor to activate the feature vector, which is then fed into a SVM classifier to identify a 5s-segment as shock or non-shock. Here the MVMD is applied to reconstructed the SH and NSH signals from preprocessed ECG segment used as three input channels of the CNNE, which enable an improvement of learned feature quality. Moreover, the structure and parameters of the final CNNE are selected carefully on the training data adopting an exhaustive grid search with nested 5-folds CV procedure. The validation performance results of this proposed SAA is relative high with Ac of 99.02%, Se of 95.21%, and Sp of 99.31%.

ACKNOWLEDGMENT

This research was supported by the Brain Research Program through the National Research Foundation (NRF) of Korea funded by the Ministry of Science, ICT & Future Planning (2016M3C7A1905477).

REFERENCES

[1] M. T. Nguyen, A. Shahzad, B. V. Nguyen, and K. Kim, “Diagnosis of shockable rhythms for automated external defibrillators using a reliable support vector machine classifier,” Journal of Biomedical Signal Processing and Control, vol. 44, pp. 258–269, Jul. 2018.

[2] P. Cheng and X. Dong, “Life-Threatening Ventricular Arrhythmia Detection with Personalized Features,” IEEE Access, vol. 5, pp. 14195-14203, Jul. 2017.

[3] S-H. Lee, “Development of Ventricular Fibrillation Diagnosis Method Based on Neuro-fuzzy Systems for Automated External Defibrillators,” Int. J. Fuzzy Syst., vol. 19, no. 2, pp. 440-451, Apr. 2017.

[4] Y. Xu, D. Wang, W. Zhang, P. Ping, and L. Feng, “Detection of ventricular tachycardia and fibrillation using adaptive variational mode decomposition and boosted-CART classifier,” Biomedical Signal Processing and Control, vol. 39, pp. 219-229, 2018.

[5] I. Guyon, S. Gunn, M. Nikravesh, and L. A. Zadeh, Feature extraction Foundations and application, Book, NY, USA: Springer Berlin Heidelberg, 2006.

[6] M. T. Nguyen, B. V. Nguyen, and K. Kim, “Shockable rhythm diagnosis for Automated External Defibrillator Using a Modified Variational Mode Decomposition Technique,” IEEE Trans. Industrial Informatics, vol. 13, no. 6, pp. 3037-3046, Dec. 2017.

[7] U. R. Acharya et al, “Automated identification of shockable and non-shockable life-threatening ventricular arrhythmias using convolutional neural networks,” Future Generation Computer Systems, vol. 79, no. 3, pp. 952-959, Aug. 2017.

[8] K. Minami, H. Nakajima, and T. Toyoshima, “Real-time discrimination of ventricular tachyarrhythmia with Fourier-transform neural network,” IEEE Trans. Biomed. Eng., vol. 46, no. 2, pp. 179-185, 1999.

[9] R. E. Kerber et al, “Automatic external defibrillators for public access defibrillation: recommendations for specifying and reporting arrhythmia analysis algorithm performance, incorporating new waveforms, and enhancing safety,” Circulation, vol. 95, no. 6, pp. 1677-1682, 1997.

[10] J. Heaton, Artificial Intelligence for Humans Volume 3: Deep Learning and Neural Networks, Book, Chesterfield, England: Heaton Research Inc, 2015.

[11] C.-C. Chang and C.-J. Lin, “LIBSVM: a Library for support vector machines,” ACM Trans. Intell. Syst. Technol., vol. 2, no. 27, pp. 1-27, 2011.

[12] L. Liu, C. Shen, and A. V. D Hengel, “Cross-Convolutional-Layer Pooling for Image Recognition,” IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 39, no. 11, pp. 2305-2313, Nov. 2017.