EasiECG: A Novel Inter-Patient Arrhythmia Classification Method using ECG Waves

Chuanqi Han, Ruoran Huang, Fang Yu, Xi Huang, and Li Cui*
{hanchuanqi18b,huangruoran,yufang,huangxi,lcui}@ict.ac.cn.

中科院计算所 Institute of Computing Technology
中国科学院大学 University of Chinese Academy of Sciences

Background

The electrocardiogram, known as ECG, is used to measure and record the electrical activities of heart, and has been widely used as the standard tool in the detection of cardiac abnormalities.In ECG record, a normal heartbeat waveform usually contains a P wave, a QRS complex, and a T wave that correspond to the atrial depolarization, ventricular depolarization and ventricular repolarization, respectively. In the past few decades, automatic classification of arrhythmias relying on Artificial Intelligence have been extensively studied. In this paper, we propose a medical experience-oriented arrhythmia classification method named EasiECG which takes advantage of the above mentioned P,Q,R,S,T waves. The RR intervals, which are the time intervals between two consecutive R waves, are also introduced as input features. Compared with other multi-stage methods, our EasiECG requires simple feature engineering and achieves higher accuracy.

Signal Pre-processing

In this paper, ECG signals are obtained from the MIT-BIH arrhythmia database. To remove noise and baseline wanders, we adopt Empirical Mode Decomposition (EMD), decomposing the ECG signals into several Intrinsic Mode Functions (IMFs). The first two high-frequency IMFs are denoised through the wavelet transform algorithm while the last two IMFs representing low-frequency components are removed. After these steps, signals are recomposed as denoised ECG records that are then split into several segments consisting of 260 samples according to the dominant R peaks. In detail, only 99 samples before the R peak and 160 samples following the R peak are reserved for one heartbeat.

Feature Extraction&Encoding

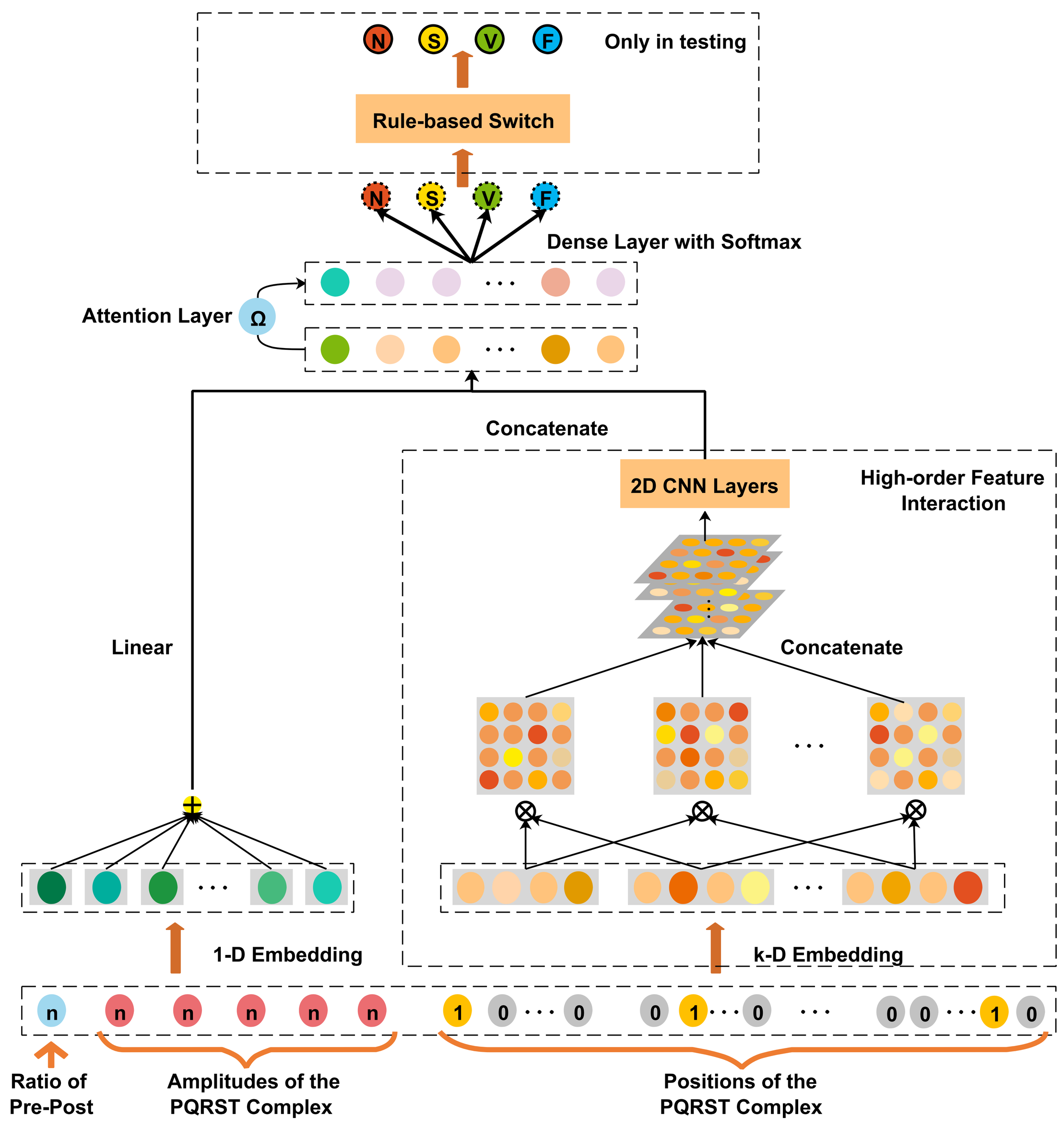
In heartbeat segmentation, two RR intervals are recorded: Pre-RR and Post-RR for each heartbeat. These two intervals, are then transformed to two normalized ones: $ratio_{pre/post} = RR_{pre}/RR_{post}$ and $ratio_{pre/avg} = RR_{pre}/\frac{1}{N}\sum_i^N RR_{pre}^i$. Except for RR intervals, the PQRST waves are located and their amplitudes as well as their locations are extracted as morphology features. Consider the situation that some of these five waves may not be successfully located due to arrhythmia or poor record quality, to facilitate our EasiECG in dealing with samples even with missing waves, these morphology features are embedded to high-dimensional features vectors. More specifically, for detected waves, their positions are numbered with their coordinate values and their amplitudes are assigned with the voltage amplitude values. For missing waves, their positions are numbered with zero, and their amplitudes are assigned with zero as well.These numbered features are then embedded adopting similar strategy to Word2Vec that is widely used in the Natural Language Process. In this way, the EasiECG can deal with all kinds of samples without any modification when faced with data with missing waves.

Proposed Method: ACFM

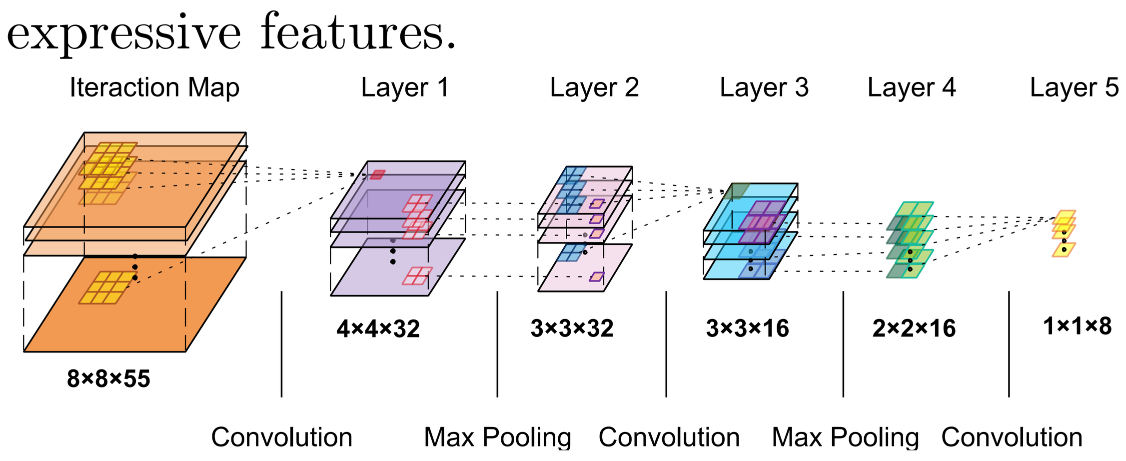
In this paper, we propose a novel adapted model named Attention-based Convolutional Factorization Machine (ACFM) for arrhythmia classification, which is formulated as:

$$\hat{y_m} = Softmax(W\langle a, [y_{linear}, y_{high-order}]\rangle + b) \quad (1)$$

where m is the number of heartbeat type, b is the global bias, y_{linear} and $y_{high-order}$ are the linear interaction and high-order interactions of features, respectively. a is the weight factor discriminating the importance of features' interactions. The following figure illustrates the whole structure of ACFM that mainly consists of an Embedding Layer, a Feature Interaction Part, an Attention Layer and a hierarchical Rule-based Switch.



The ACFM captures linear interactions through linear regression. For high order interaction, we first correlates each pair of embedded features as a feature map through outer product, then we concatenate all feature maps as a 3D tensor T like an "image". In this way, a Convolutional Neural Network is introduced to acquire more expressive features.



In order to classify heartbeats more accurately, the most relevant feature interactions should be assigned of greater importance. Therefore, an attention mechanism is implemented in the Attention Layer. Finally, since individual difference introduced by the inter-patient paradigm may decrease the model's output accuracy, we add a rule-based switch according to ratio of Pre-Avg RR intervals in the end when testing.

Acknowledgements

The paper is supported by the National Natural Science Foundation of China (NSFC) under Grant No. 61672498 and the National Key Research and Development Program of China under Grant No. 2016YFC0302300.

Experiment

In our experiment, we aim to classify heartbeats into AAMI heartbeat classes of Non-ectopic beats(N), Supra-ventricular ectopic beats(S), Ventricular ectopic beats(V) and Fusion beats(F). We adopt the inter-patient paradigm which means there is no individual overlap between training and testing sets. In detail, 44 records in MIT-BIH arrhythmia database are split into two subsets, DS1 and DS2. Each subset consists of the following records. DS1 = {101, 106, 108, 109, 112, 114, 115, 116, 118, 119, 122, 124, 201, 203, 205, 207, 208, 209, 215, 220, 223, 230} DS2 = {100, 103, 105, 111, 113, 117, 121, 123, 200, 202, 210, 212, 213, 214, 219, 221, 222, 228, 231, 232, 233, 234} During training, a focal loss which is formulated as: $focal\ loss = -\alpha_t(1-\hat{y}_{true})^{\gamma}log(\hat{y}_{true})$ is introduced where we assign α as 0.5, 0.9, 0.6, 1.0 for class N,S,V,F,respectively. γ is valued 2.0.

Results

To evaluate the performance of the proposed method, three measures are adopted: Sensitivity(Se), Positive predictive value($+P$) and Specificity($Spec$).The following table presents the performance comparison between our proposed method and several typical state-of-the-arts, where our method ACFM has achieved the best overall performance that the sensitivities in class N and class S are top 1 while that of class V ranks No.2.

Method	N			S			V		
	Se	$+P$	$Spec$	Se	$+P$	$Spec$	Se	$+P$	$Spec$
Llamedo et al.[13]	95	98	-	77	39	-	81	87	-
Lin and Yang [14]	91.0	99.0	-	81.0	31.0	-	86.0	73.0	-
Garcia et al.[32]	94.0	98.0	82.6	62.0	53.0	97.9	87.3	59.4	95.9
Zheng et al.[7]*	90.5	99.6	96.2	90.3	45.2	96.0	91.5	51.3	94.5
Anwar et al.[18]*	94.8	99.7	97.4	85.9	45.5	96.2	**95.1**	78.5	98.3
Proposed Method	**95.4**	99.5	95.2	**91.6**	52.1	96.9	92.9	78.8	98.4

To comprehensively understand the functions of components and mechanisms proposed in our model, we further carry out several ablation experiments. The results are demonstrated in the table below.

Methods	N	S	V	F
Proposed Method	38497	1488	2587	82
Ablation Expt 1[1]	36352 (-5.1%)[5]	401 (-66.0%)	2594 (+0.3%)	90 (+2.1%)
Ablation Expt 2[2]	37693 (-1.9%)	1495 (+0.4%)	2342 (-8.8%)	72 (-2.6%)
Ablation Expt 3[3]	38444 (-0.1%)	1411 (-4.7%)	2573 (-0.5%)	0 (-21.4%)
Ablation Expt 4[4]	38815 (+0.8%)	1434 (-3.3%)	2422 (-6.0%)	0 (-21.4%)

[1] Proposed model without Rule-based Switch
[2] Proposed model without attention mechanism
[3] Proposed model using cross-entropy as loss fuction when training
[4] Model based on classical FMs
[5] Numbers in brackets indicate the percentage of increasement(+) / decreasement(-) in sensitivities

To more explicitly validate our model's robust performance on samples with missing waves, we designed the following experiment. We randomly select 10%, 20%, 60%, 50% samples of the N, S, V, F classes, respectively in the training and testing sets and remove their P waves. Then we train the ACFM using the same strategy. The corresponding results are summarized below,where an overall accuracy of 87.6% is achieved.

Class	Correct	False	Total
N	3896 (89.4%[1])	460	4356
S	307 (91.1%[1])	30	337
V	1488 (89.2%[1])	181	1669
F	52 (27.1%[1])	140	192
Total	5743 (87.6%[2])	811	6554

[1] Sensitivity of each class
[2] Overall accuracy := $\sum_{i=N,S,V,F} TP_i / \sum_{i=N,S,V,F} Num_i$, TP_i and Num_i denote True Positive and total samples of class i, respectively.

We therefore get the conclusion that our method EasiECG is still applicable even to datasets containing samples with missing waves.