Chengyu Liu · Jianqing Li *Editors*

Feature Engineering and Computational Intelligence in ECG Monitoring

Springer

Feature Engineering and Computational Intelligence in ECG Monitoring

Chengyu Liu • Jianqing Li
Editors

Feature Engineering and Computational Intelligence in ECG Monitoring

Springer

Editors
Chengyu Liu
The State Key Laboratory
of Bioelectronics, School of Instrument
Science and Engineering
Southeast University
Nanjing, China

Jianqing Li
School of Biomedical Engineering and
Informatics
Nanjing Medical University
Nanjing, China

ISBN 978-981-15-3823-0 ISBN 978-981-15-3824-7 (eBook)
https://doi.org/10.1007/978-981-15-3824-7

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

Preface

Computational Intelligence Will Benefit Heart Patients

Clinical observation of heart activity from the body surface has become so common that we now rarely stop to think about it. Advances made over the past 100 years have been astonishing. What lies next? It would be all too easy to say that there is little more that can be achieved. That is far from the case, and a well-informed review of modern computational approaches should encourage new developments, with the expectation that many will progress to become essential clinically useful techniques. The book by Chengyu Liu and colleagues (*Feature Engineering and Computational Intelligence in ECG Monitoring*) is just what is needed to encourage such new research.

Electrocardiography started with the development of often cumbersome, complex and costly equipment. As a result of much research and development, devices that create the electrocardiogram (ECG) now appear in every hospital and in most family doctor clinics and ambulances.

The most obvious results from body surface ECGs include the diagnosis of cardiac conditions such as heart attacks and abnormal rhythms, the detection of life-threatening atrial fibrillation, and the continuous monitoring of patients to detect ventricular fibrillation with the ability to initiate life-saving defibrillation.

All such techniques are based on well-founded and obvious features of the ECG. What is the role of more subtle features that undoubtedly relate to cardiac muscle and conduction abnormalities and hidden cardiac conditions?

Readers and researchers can now be introduced to modern computational techniques that use what we would generally consider "intelligence" and examine features that even the trained eye would not see. Then in order to assist future researchers, information is provided on where to find numerous relevant ECG examples that can be used in initial research. No ECG is ever perfect, and in real situations many ECGs are covered in noise that needs to be removed or its influence avoided by further computational methods.

There are many practical areas where new computational techniques would be invaluable, including the ability to record from multiple body sites while subjects continue to pursue their normal activities. This is an important area for future research, as many subjects have abnormal heart conditions that appear only during specific patient situations or occur very infrequently.

Then finally, the review of applications for these new techniques is essential. It is for the benefit of patients that there are many applications, and this list can only but grow with the research techniques described.

Newcastle upon Tyne, UK Alan Murray

Contents

About the Editors

Chengyu Liu received his B.S. and Ph.D. degrees in Biomedical Engineering from Shandong University, China, in 2005 and 2010, respectively. He completed his postdoctoral trainings at Shandong University in China, Newcastle University in UK and Emory University in USA from 2010 to 2017. He is now a Professor of the School of Instrument Science and Engineering and the State Key Laboratory of Bioelectronics in Southeast University and the founding Director of the Wearable Heart-Sleep-Emotion Intelligent Monitoring Lab in Southeast University. He also serves as the founding Chair of the China Physiological Signal Challenge from 2018. He has been a member of the Federation Journal Committee of the International Federation for Medical and Biological Engineering (IFMBE) since 2016. His research topics include wearable ECG and vital signs monitoring, machine learning for medical big data, early detection and device development for cardiovascular diseases.

Jianqing Li received his B.S. and M.S. degrees in Automatic Technology and his Ph.D. degree in Measurement Technology and Instruments from the School of Instrument Science and Engineering, Southeast University, China, in 1986, 1990 and 2000, respectively. He is now a Professor at the School of Biomedical Engineering and Informatics, Nanjing Medical University. He is also a part-time Professor with the School of Instrument Science and Engineering in Southeast University. He serves as the founding Director of the Key Laboratory of Clinical Medical Engineering in Nanjing Medical University and the deputy Director of the Jiangsu Key Lab of Remote Measurement and Control in Southeast University, leading the research works on medical-industry cross-innovation cooperation, medical device development and clinical applications. His research topics include wearable medical sensors, wireless sensor network, rehabilitation robot technology and robot telepresence technology.

Part I
Introduction

Feature Engineering and Computational Intelligence in ECG Monitoring: An Introduction

Chengyu Liu and Jianqing Li

1 Background and Motivation

Recent advances in wearable technology and on devices based on Internet of Things (IoT) have led to an explosion of routinely collected individual electrocardiogram (ECG) data. The use of feature engineering and computational intelligence methods to turn these ever-growing ECG monitoring data into clinical benefits seems as if it should be an obvious path to take. However, this field is still in its infancy, and many nominal "computational intelligence" techniques do not produce substantial help with clinical diagnostic. Many essential concepts and current solutions need to be reviewed and clarified in depth.

ECG Holter is a conventional technology for long-term, dynamic ECG monitoring. Although its application in any hospital is relatively mature, it cannot be easily applied for monitoring people outside the hospital. In the out-of-hospital monitoring environment, subjects have more autonomous activities, inducing a large amount of unexpected noise and hence yielding extreme difficulty in signal analysis. In addition, out-of-hospital monitoring also needs more comfortable monitoring situations. Current wearable ECG technology can be more comfortable and convenient than ECG Holter, but it still needs to maintain a rigorous attitude to its use, with a clear understanding for its practical application, especially for large-scale crowd

C. Liu (✉)
The State Key Laboratory of Bioelectronics, School of Instrument Science and Engineering, Southeast University, Nanjing, China
e-mail: chengyu@seu.edu.cn

J. Li
The State Key Laboratory of Bioelectronics, School of Instrument Science and Engineering, Southeast University, Nanjing, China

School of Biomedical Engineering and Informatics, Nanjing Medical University, Nanjing, China

C. Liu, J. Li (eds.), *Feature Engineering and Computational Intelligence in ECG Monitoring*, https://doi.org/10.1007/978-981-15-3824-7_1

monitoring applications. Real-time, long-term, and dynamic ECG monitoring is helpful for detecting sporadic or hidden arrhythmias. However, we should be alert to the difficulties in medical services caused by a large number of applications, because at present, medical services such as high-quality doctor diagnosis and interpretation for long-term dynamic ECG data have not yet matured into a standardized system. Whether it is patients themselves or equipment suppliers, when they bring a large amount of data to the doctor, it is difficult to obtain full data interpretation and analysis. In addition, the accuracy of the signal processing algorithm itself also needs to be improved. Whether it is a signal quality problem or an algorithm problem, it will cause errors or unreasonable reporting results, and these results will lead to unnecessary medical intervention. In addition, it is important to build accurate dynamic ECG databases, with clear labeling. Manual labeling for massive ECG data is obviously unrealistic. Thus, developing automatic labeling algorithms for dynamic ECGs seems to be a promising direction. It can automatically obtain more reliable feature labels by fusing multiple single algorithm results and effectively improving the reliability of labeling results.

The purpose of this book is to summarize the recent progress in feature engineering and computational intelligence solutions for ECG monitoring, with an emphasis on how these methods can be efficiently used for the emerging needs and challenges—dynamic, continuous, and long-term individual ECG monitoring and real-time feedback and application, aiming to provide a "snapshot" of the state of current research at the interface between ECG signal analysis and machine learning. It could help clarify some dilemmas and encourage further investigations in this field, to explore rational applications of feature engineering and computational intelligence in clinical practices for dynamic ECG monitoring.

2 Open-Source Advances

This book also aims to emphasize the spirit of open-source research, which is characterized by participation, sharing, dedication, and ultimately converging into a community, due to a group behavior, not an individual behavior. One of the authors has participated in the organization of the PhysioNet Challenge with Professor Gari Clifford while working in the United States. People who perform ECG research know the PhysioNet databases. We need to thank Professor Roger G. Mark and Dr. George B. Moody from MIT for their outstanding and leading work in open-access databases. It was their pioneering contributions that allowed later researchers to easily access a large number of clinical physiological data resources and data analysis toolkits and then start their own research work. Since 2000, under the initiative from Roger and George, PhysioNet and the Computing in Cardiology (CinC) conference have jointly launched the annual Physionet/CinC Challenge, which has become an important annual event in the field of physiological signal analysis. After 2015, the organization baton of the Physionet/CinC Challenge was delivered to Professor Gari D. Clifford, who has significantly enhanced the influence

of the challenge worldwide. The author Chengyu Liu participated in the organization of the two competitions in 2016 and 2017. We actively sought valuable competition data from universities, research institutes, and enterprises and then did a lot of work for data labeling and provided some basic signal processing algorithms as benchmarks. The competition provided open-source, free, high-quality databases for researchers worldwide and, more importantly, provided high-value data labeling, benchmark algorithms, and communication platforms, which greatly promoted the progress of open sources. After Chengyu returned to work in China in 2017, he has initiated and organized the annual Chinese Physiological Signal Challenge (CPSC), which promoted the progress of open sources in China.

From a broader perspective, we need to thank Mr. Linus B. Torvalds, who is the father of Linux and an important initiator of the open-source concept; Mr. Dennis M. Ritchie, who is the father of the C language and UNIX as the first modern programming language; and also Mr. Guido van Rossum, who founded the Python language, which greatly accelerated the progress of machine learning.

3 Progress So Far

Open source requires sharing and dedication, but before sharing and dedication, we need to understand the use of existing resources and to participate in the process of co-constructing the resources. In this sense, one of the purposes of this book is to provide readers with the progress of current ECG research and then to share, participate, and contribute under the spirit of open source. Specially, we prepared Chap. 2 to introduce several open-source ECG databases from our group, two of which became the challenge data for the CPSC2018 and CPSC2019. Along with the competition, there are corresponding open-source codes for reference. We should note that, although data and source code are important, they are not the whole and essential spirit of open source. The value of data and source code can be significantly improved with interactive discussions. The true spirit of open source is to form an interactive community. Everyone can take the existing resources as a ladder and eventually achieves further and better goals through participation, sharing, contribution, and interaction.

4 Survey of Contents

"Representative Databases for Feature Engineering and Computational Intelligence in ECG Processing" chapter: With the development of data science and artificial intelligence (AI) technique, data have become the source of novel research. The value of big data has been widely acknowledged by researchers from the machine learning community. Data size/diversity has been regarded as more important than the choice of machine learning algorithms. Efforts to create new databases or even to

increase the size of existing ones, as well as creating standard evaluation protocols, have been made in several research areas, especially to avoid the potential unfair comparisons between methods. The widely used MIT-BIH ECG database does not perfectly fit the algorithm development for wearable ECG analysis. The team from Southeast University has undertaken great work to address this issue and has constructed several databases of wearable ECGs from both cardiovascular disease and healthy subjects. "Representative Databases for Feature Engineering and Computational Intelligence in ECG Processing" chapter reviews five open-access databases and describes the specific composition of each database, by supplying both recording and annotation information since adequate information seems to be the most important element in feature engineering and computational intelligence in ECG monitoring. These databases can be used for training ECG algorithms for feature detection, arrhythmia classification, and signal quality assessment.

"An Overview of Signal Quality Indices on Dynamic ECG Signal Quality Assessment" chapter: Signal quality assessment (SQA) is usually the first step for dynamic ECG analysis, to choose the signal episode with good signal quality and avoid false diagnosis due to artifact or noise. Feifei Liu et al. introduce methods for developing signal quality indexes (SQIs) and for SQA of dynamic ECGs. SQIs were developed from the morphology, time domain, frequency domain, and nonlinear features of ECG signals. The special aspect for SQIs is the newly developed generalized bSQI method, named as GbSQI. Typical machine learning methods were used for developing the models for signal quality classification. This chapter provides a simple review for the commonly used SQIs for quantifying the noise level not only for dynamic ECGs but also for other physiological signals collected from the dynamic environment.

"Signal Quality Features in Dynamic ECGs" chapter: Unlike the former chapter focusing on the SQIs developed by multiple domain features based on the prior knowledge of ECG, this chapter focuses on only self-correlation features. Herein, Yixuan Li et al. present a special method combining both long and short template matching mechanisms and propose a real-time SQA algorithm for wearable ECGs based on multi-template matching and correlation coefficient matrix. The algorithm was verified using the data from three types of signal quality classification.

"Motion Artefact Suppression Method for Wearable ECGs" chapter: Noise due to the artifact and electrode motion is hard to totally eliminate because it overlaps in both time and frequency domains with the real ECG components. Jianlong Zhao and Huanqian Zhang present a review of motion artifact suppression methods in wearable ECGs. By analyzing the source of motion artifact from the view of the skin–electrode interface and electrical model in wearable environments, motion artifact suppression methods from both hardware design and adaptive filter algorithm were developed. This chapter introduces the basic concepts of analog circuit model of skin tissue, electrical circuit architecture, skin electrode impedance measurement of singlechannels and multichannels, as well as the design for adaptive filters based on least means square (LMS) to suppress the motion artifact.

"Data Augmentation for Deep Learning-Based ECG Analysis" chapter: Deep learning has become "hot" technology in recent years owing to its admirable

performance in data processing. Though it is also widely used in ECG signal processing and achieves good results in some typical examples, lack and imbalance of the annotated data seriously hampers the performance of these deep learning-based analysis methods. The challenge also exists in the balance between huge amount of raw data and the limited diversity of patients. Qing Pan et al. summarize the typical data augmentation methods applicable for ECG analysis and examine their performance on a task for detecting atrial fibrillation (AF), verifying the efficiency of data augmentation methods for the unbalanced ECG data analysis.

"Study on Automatic Classification of Arrhythmias" chapter: With the former chapter concerning with the preprocessing step (data augmentation) of machine learning, this chapter discusses automatic classification methods for ECG arrhythmias. There are various types of arrhythmias, and each type is associated with a special pattern, and as such, it is possible to be identified and classified. However, accurately identifying and classifying arrhythmias in dynamic situations is still challenging. Runnan He et al. review the existing studies regarding arrhythmia classification methods and propose a machine learning model composed of several steps, including preprocessing, feature extraction, feature selection, and automatic classification. This chapter provides a helpful and systematic tutorial on the practical operation of machine learning in ECG arrhythmia classification.

"ECG Interpretation with Deep Learning" chapter: Continuing the topic of handling with the arrhythmia classification, Wenjie Cai and Danqin Hu concentrate on the application of convolutional neural networks (CNN) and recurrent neural networks (RNN) for exploring the rational machine learning model. The great success of 2D CNN in image recognition has attracted much attention on the convolution analysis of 1D time series. The addition of RNN can help to enhance the insensitivity of CNN on timing information while increasing the computational cost. This chapter gives a systematical review on CNN-based, RNN-based, as well as CNN&RNN-based intelligent analysis models for automated ECG interpretation.

"Visualizing ECG Contribution into Convolutional Neural Network Classification" chapter: CNN has achieved great success in classification tasks but lacks the interpretability and thus is usually regarded as a "black box," which means at least the following two aspects: (1) the neural network feature or decision logic is difficult to understand at the semantic level; (2) lack of mathematical tools to diagnose and evaluate the network's feature expression capabilities, such as generalization ability and convergence speed. Yaowei Li et al. provide a routine for classification and visualizing the ECG contribution with CNN and deconvolved neural network (DeconvNet) and interpretation and presentation of the CNN model as feature importance degree heatmap (FIDH). This chapter shows that the main ECG features (such as QRS complexes) can be highlighted by a CNN deconvolution process based on the absolute amplitude of ECG signal.

"Atrial Fibrillation Detection in Dynamic Signals" chapter: Rapid development of wearable technology has greatly promoted the progress of dynamic ECG monitoring, and AF detection has become a most important issue since its popularity and universality. The PhysioNet/Computing in Cardiology (CinC) Challenge 2017 aimed to develop a classification algorithm for classifying normal or AF or other

arrhythmia or noise from an ECG segment. Caiyun Ma et al. describe the detailed physiological concepts used to develop AF detection algorithms and redefine several AF features in dynamic ECG analysis. SVM was used to train the AF and non-AF classifier, and the model was verified on the wearable ECG database.

"Applications of Heart Rate Variability in Sleep Apnea" chapter: According to the World Health Organization survey, about one-third of the world's population suffers from sleep disorders, and the proportion has been increasing in recent years. With the development of wearable ECG, sleep apnea-hypopnea syndrome (SAHS) can be monitored and detected from single-channel ECGs. Heart rate variability (HRV) is an important marker of autonomic nervous system (ANS) activity and can be used for sleep apnea analysis. Xiaotong Dong et al. studied multisource HRV features (time and frequency domain and nonlinear analysis) from SAHS patients and found that the features have significant differences between normal sleep and sleep apnea signals, indicating its usefulness as a preliminary screening tool for detecting sleep apnea.

"False Alarm Rejection for ICU ECG Monitoring" chapter: High false alarm rates have been reported in intensive care unit (ICU) monitors, which decrease quality of care by slowing staff response times while increasing patient burden and stress. A National Patient Safety Goals report in 2016 recognized reducing the hazards of clinical alarm fatigue as an important consensus to improve patient safety. Dai Jian et al. provide an overview on the variable reasons of alarm generation in ICU and how to reduce false alarms for ECG monitoring. Clinical alarm management system improvement and intelligent alarms are proposed in this chapter in order to reduce and suppress the meaningless ECG alarm.

"Respiratory Signal Extraction from ECG Signal" chapter: Respiration is an important factor to monitor diseases, and abnormal respiratory rate is recognized as an indicator of catastrophic deterioration of diseases. In special situations, respiratory rate is even more reliable for monitoring cardiac arrest than heart rate. In practice, respiratory signal is usually derived from the ECG signal. Kejun Dong et al. focus on the detection of ECG-derived respiration (EDR). This chapter reviews eight ECG-derived respiration methods and verifies them using the synchronously collected respiratory signals.

"Noninvasive Recording of Cardiac Autonomic Nervous Activity: What Is Behind ECG?" chapter: The autonomic nervous system (ANS) innervates internal organs unconsciously and controls the homeostasis of the human body. Studies show that autonomic nerves are closely related to cardiovascular disease, and its regulation of cardiovascular disease is multifaceted. Yike Zhang et al. overview the relationship between the ANS and cardiovascular diseases, including the functions of ANS, ANS in the cardiovascular system, and ANS dysfunction and related diseases. Then the important features for quantifying the ANS function are introduced, including HRV features, heart rate turbulence features, baroreflex sensitivity features, as well as skin sympathetic nerve activity (SKNA) features. This last SKNA technique is highly emphasized.

"A Questionnaire Study on Artificial Intelligence and Its Effects on Individual Health and Wearable Device" chapter: Development of AI-related techniques has

brought great convenience to our lives. At the same time, it has also promoted the rise of applications of mobile Internet and wearable devices. Meanwhile, wearable techniques have become an important part for personal analysis, by measuring physical conditions, recording physiological parameters, or notifying medication schedules. This evolving technology not only is expected to help people pursue a healthier lifestyle but also provide continuous medical data to proactively track metabolic status, diagnosis, and treatment. To better understand the frequent responses to these current hot topics, Tiange Bu and Fangyuan Li designed and presented a questionnaire survey focusing on the abovementioned hot topics, AI and its impact on personal health and wearables. "A Questionnaire Study on Artificial Intelligence and Its Effects on Individual Health and Wearable Device" chapter introduces this survey and summarizes its results in detail.

Acknowledgements We gratefully acknowledge the many organizations that have facilitated this project, including the National Natural Science Foundation of China (61201049, 61571113, and 81871444), the Distinguished Young Scholars of Jiangsu Province (BK20190014), the Key Research and Development Programs of Jiangsu Province (BE2017735), and the support from the Southeast-Lenovo Wearable Heart-Sleep-Emotion Intelligent Monitoring Lab. We would also like to acknowledge the Springer Nature Publisher for all their editorial help. Finally, we would like to thank all the authors who have contributed to this book.

Part II
Databases Available for Research

Representative Databases for Feature Engineering and Computational Intelligence in ECG Processing

Hongxiang Gao, Chengyu Liu, Qin Shen, and Jianqing Li

Abstract Standard electrocardiogram (ECG) database is a key point for validating the algorithms of feature detection and disease diagnosis. Researchers usually use the ECG databases posted in the PhysioNet platform, which were basically collected from clinical environment with high signal quality. Performance of the developed algorithms from these databases suffers the poor robustness and weak generalization when implemented on the dynamic ECGs typically collected by wearable devices. Standard and accredited dynamic databases are absent. Six open-accessed ECG databases, including signal quality database, China physiological signal challenge (CPSC) 2018 and 2019 databases, arrhythmia database, atrial fibrillation (AF) database, and long-term ECG database, were therefore tidied up and published freely. All the valuable ECG feature information were carefully annotated by cardiologists. We hope these databases may benefit the ECG study on dynamic signal processing.

Keywords Electrocardiogram (ECG) · Database · Arrhythmia · Wearable

H. Gao · C. Liu (✉)
The State Key Laboratory of Bioelectronics, School of Instrument Science and Engineering, Southeast University, Nanjing, China
e-mail: chengyu@seu.edu.cn

Q. Shen
The First Affiliated Hospital of Nanjing Medical University, Nanjing, China

J. Li
The State Key Laboratory of Bioelectronics, School of Instrument Science and Engineering, Southeast University, Nanjing, China

School of Biomedical Engineering and Informatics, Nanjing Medical University, Nanjing, China

C. Liu, J. Li (eds.), *Feature Engineering and Computational Intelligence in ECG Monitoring*, https://doi.org/10.1007/978-981-15-3824-7_2

1 Introduction

In the field of health management, it is vital to timely discover the abnormal health of the body and the early warning of major disease risks. Traditionally, we will implement this requirement through annual medical examinations, even if the time span of medical examinations is large, and the geographical coverage is not enough, too. In terms of electrocardiogram (ECG) diagnosis, dozens of seconds recording with 12-lead in routine inspection and 24-hour recording for Holter monitoring were widely used in clinical environment. The downside of these methods is, of course, their inefficiency to cover the real-time and long-term diagnostic requirements. Recently, people have turned their attention to wearable devices, which can achieve real-time and long-term monitoring for physical abnormalities of large populations across regions.

The number of available devices that can record and analyze ECG is on the rise [1]. Biswajit et al. [2] searched the Internet and PubMed and found out that most of the available devices are developed in the engineering domain with an isolation from the medical domain. Wearable ECG devices are emerging with their portability and real-time capabilities all appear to be less electrodes and mobile device visualization [3]. They can play a significant role in improving the health and wellness of subjects by increasing the availability and quality of healthcare, with the application for early detection of cardiovascular disease (CVD), as well as for other situations, such as sports and fitness, rehabilitation, elderly care support, emotion, and sleep monitoring [4–8]. Different types of ECG systems [9–12] have been introduced so far to improve the signal quality in the clinical settings. Nevertheless, simple equipment brings up more kinds of data situations and more complicated diagnostic requirements for heart disease. Most of the available ECG wearable devices only monitor the heart rate (HR) without detecting any heart diseases in real time. Therefore, a mature intelligent analysis system is essential to deal with common abnormalities in CVDs and output the necessary medical information.

Although ECG analysis methods have been severely tracked throughout the last several decades and many sophisticated algorithms have been proposed, they are not perfectly suitable for noisy ECG episode or abnormal rhythm waveforms, especially when the ECG recordings are from dynamic wearable situations [3, 4, 13–15]. In contrast to standard ECG leads, signals in wearable devices are acquired by simulated limb leads, which lead to variant ECG morphology. Thus, novel automatic diagnostic algorithms should be trained not only on the existing ECG databases but also on the expert-annotated wearable database, aiming to enhance its efficiency in wearable applications.

Along with the development of data science and artificial intelligence (AI), data has become the important source for novel research. The value of big data has been widely accepted by all parties, and the collection of big data has become an important part. The construction of general well-organized database would be a great challenge because, besides the financial costs involved, they would have to be incorporated into standards such as AAMI standards to reach the desired audience.

The PhysioNet platform [16], established in 1999, is managed by members of the MIT Laboratory for Computational Physiology, offering free access to large collections of physiological and clinical data and related open-source software. This platform provides the most commonly used ECG databases for developing algorithms. AHA database was released since 1982 and completed in 1985, consisting of 80 two-channel excerpts of analog ambulatory ECG recordings, digitized at 250 Hz per channel with 12-bit resolution over a 10 mV range. MIT-BIH Arrhythmia Database [17, 18] contains 48 half-hour excerpts of two-channel ambulatory ECG recordings, obtained from 47 subjects studied by the BIH Arrhythmia Laboratory between 1975 and 1979. In cooperation with the annual Computing in Cardiology (CinC) conference, PhysioNet also hosts an annual series of challenges, which published a lot of clinical data for the unsolved problems in clinic and basic science and attracted lots of researchers and supporters. In the nearly 20 years, the PhysioNet database has enormously boosted the development for research and education.

Unfortunately, dynamic and wearable ECGs, as well as the standardization, such as lead position, lead wire, electrode material, data formats, lack. In the past several years, we have collected dynamic ECGs (usually 24 h) from more than 300 individuals with specific CVD diseases using wearable devices. To address the concern about the shortage of dynamic/wearable ECG database, in the current study, we tidied up five well-annotated wearable databases and one traditional ECG database from the clinical collection (see more profiles in Table 1) and posted them to researchers interested in ECG-related study.

2 Sources and Types of Databases

Six open-accessed ECG databases were prepared to accelerate the development of automatic algorithms used for dynamic/wearable environment. The China Physiological Signal Challenge (CPSC) was held annually since 2018, aiming at providing a platform for the open-source data and algorithms. We included the databases from CPSC 2018 and 2019 here [19, 20].

2.1 Signal Quality Database

In recent years, analysis and evaluation of ECG signal quality has been a hot topic [21–23]. The PhysioNet/CinC Challenge 2011 [24] aimed to develop an efficient algorithm able to run in real time within a mobile phone [25] that can provide useful feedback to a layperson in the process of acquiring a diagnostically useful ECG recording. Due to the poor signal quality caused by the dry electrodes [26, 27], signal quality assessment is considered as a main target. Herein, a specialized database of 300 recordings was designed for signal quality study (see Table 2), with three categories of signal quality: good (A), medium (B), and poor (C) signal quality.

Table 1 Data profiles for six open-accessed ECG databases

Database	Type	#recordings	Profile
Signal Quality Database [44]	Good signal quality (A)	100	• 2-channel • 400 Hz • 10-s length
	Medium signal quality (B)	100	
	Poor signal quality (C)	100	
Database of CPSC 2019 [20]	Pathological arrhythmias	1220	• 1-channel • 500 Hz • 10-s length
	Abnormal sinus rhythm	46	
	Noise/artifacts	734	
Arrhythmias Database [44]	Sinus bradycardia	50	• 1-channel • 400 Hz • 30-s length
	Sinus tachycardia	43	
	Sinus arrest	4	
	Single PAC	100	
	Coupled PAC	50	
	SAT/NSAT	50	
	AF	50	
	Single PVC	100	
	Coupled PVC	20	
	SVT/NSVT	1	
AF Database [44]	AF	20	• 1-channel • 400 Hz • 30-min length
Long-term Arrhythmias [44]	Arrhythmias	2	• 1-channel • 400 Hz • 24-h length
Database of CPSC 2018 [19]	Normal	918	• 12-channel • 500 Hz • Varying lengths (6–60 s)
	AF	1098	
	I-AVB	704	
	LBBB	207	
	RBBB	1695	
	PAC	574	
	PVC	653	
	STD	826	
	STE	202	

PAC premature atrial contract, *PVC* premature ventricular contract, *NSAT/SAT* non-sustained/sustained atrial tachycardia, *AF* atrial fibrillation, *NSVT/SVT* non-sustained/sustained ventricular tachycardia, *I-AVB* first-degree atrioventricular block, *LBBB* left bundle branch block, *RBBB* right bundle branch block, *STD* ST-segment depression, *STE* ST-segment elevation

ECGs for signal quality type A have clear and distinct P-QRS-T morphologies accompanied by slight noise or artifacts occasionally. ECGs for signal quality type B show obvious rhythmical characteristics, but with obvious signal noises and cannot be used for morphology diagnosis. ECGs for signal quality type C are totally unacceptable recordings, due to the large proportion of noise. Typical examples of ECG waveforms are shown in Fig. 1.

Table 2 Specification for signal quality division

Category	Symbol	Definition
Good	A	Signal with apparent P-QRS-T morphologies Signals with slightly baseline drift or transient artifacts
Medium	B	A good recording contaminated severely in a narrow window A good recording with one or a few missing signals A poor recording that may be interpretable with difficulty
Poor	C	Signal usefulness to clinical application (may be caused by misplaced electrodes, poor skin–electrode contact)

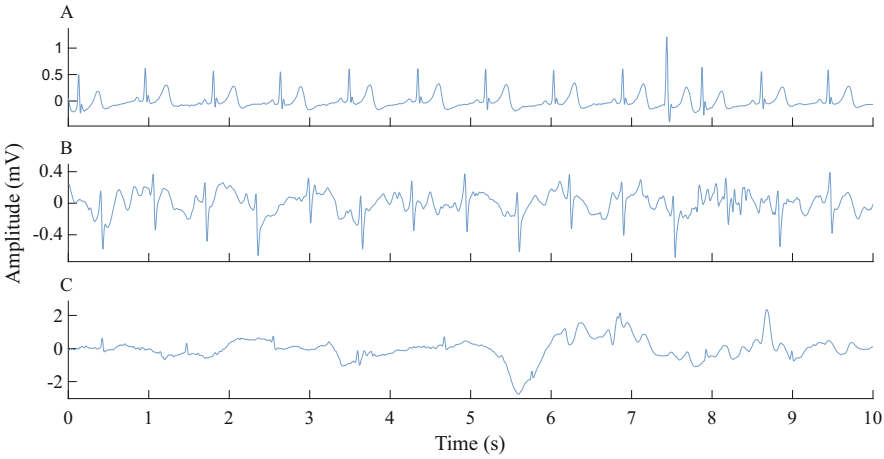

Fig. 1 Typical examples of the three signal quality categories: (**a**) good signal quality, (**b**) medium signal quality, (**c**) poor signal quality

2.2 CPSC 2019 Database

Training set of the CPSC 2019 database [20] consists of 2000 single-lead ECG recordings collected from patients with CVDs. Each recording lasts for 10 s. Test set contains the similar ECG recordings with the same time lengths, but it is unavailable to the public. ECG recordings were sampled as 500 Hz. All recordings were provided in MATLAB format (each including two ".mat" file: one is ECG data and another one is the corresponding QRS annotation file). These 10-s ECGs are challenging for QRS detection, as well as for HR estimation.

2.2.1 Data Structure

1. **Type A: pathological arrhythmias**
   The abnormal heartbeats, generated by the irregularity in the origin/conduction of the cardiac electrical activity, mainly include the following: LBBB, RBBB, and

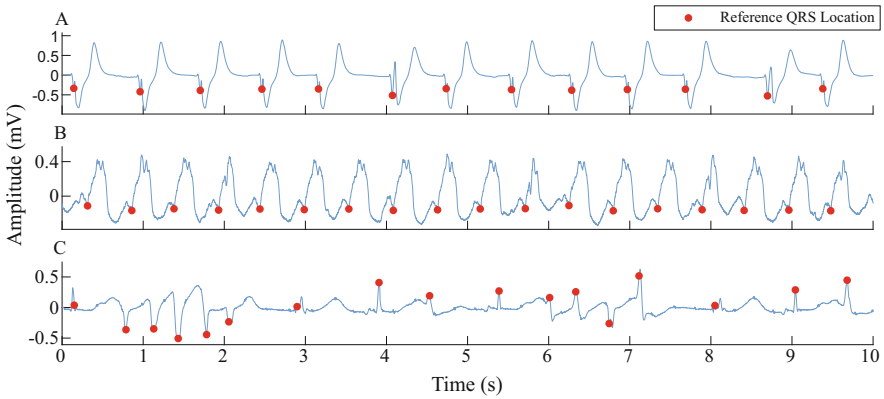

Fig. 2 From top to bottom: LBBB (**a**: data00872), RBBB (**b**: data00295), and PVC (**c**: data00480). Red circles denote the reference QRS locations

Table 3 Diagnostic criteria of RBBB and LBBB

RBBB	LBBB
QRS duration greater than 120 ms	QRS duration greater than 120 ms
rsR' "bunny ear" pattern in the anterior precordial leads (leads V1–V3)	Monomorphic R wave in leads I, V5, and V6 with no Q waves
Slurred S waves in leads I, aVL, and frequently V5 and V6	ST and T wave opposite to the major deflection of the QRS complex

PVC (see Fig. 2). These episodes can be from any of the 12 ECG leads, rather than from a special ECG channel. Thus, the morphology of the ECG episodes varies. Traditional threshold algorithms (usually amplitude threshold) show poor performance when dealt with the small amplitude of QRS complexes caused by abnormal heartbeats.

When bundle branch block occurs, one branch of His bundle delays the conduction of the electrical impulse, and ventricle is activated by the myocardial propagation of electrical activity from other ventricles. Thus, the affected ventricle is depolarized erratically and slowly through an alternative pathway. This delay is shown in ECG with a widening of QRS complex (duration >120 ms) and a change of its pattern, which varies depending on the affected branch, acted as RBBB or LBBB. Specific diagnostic criteria of RBBB and LBBB given by the ACC/ESC consensus document are summarized in Table 3 [28, 29].

2. **Type B: sinus tachycardia and sinus bradycardia**
Sinus tachycardia and sinus bradycardia are sinus rhythms with a rate higher than 100 beats per minute (bpm) or less than 60 bpm. In sinus tachycardia the sinus node fires between 100 and 180 impulses per minute. Maximal HR decreases with age from around 200 to 140 bpm. In sinus bradycardia the sinus node fires at a slow (<60 bpm) rate. More severely, sinoatrial exit block or sinus arrest may occur during sinus bradycardia and cause a long break. All these sinus

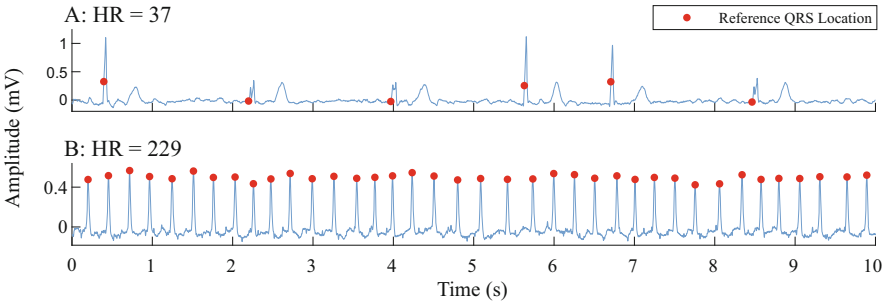

Fig. 3 Example of bradycardia (**a**: data00134) and tachycardia (**b**: data00470). Red circles denote the reference QRS locations

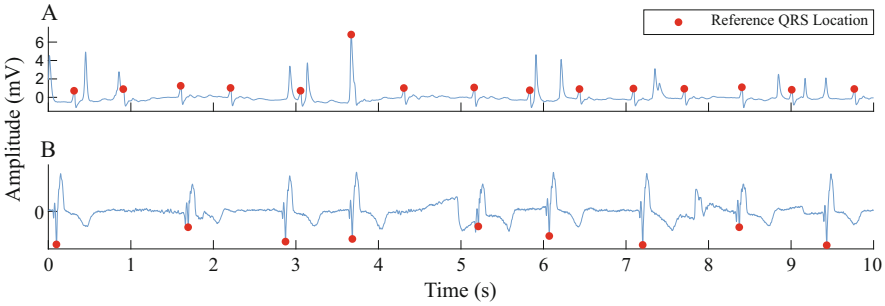

Fig. 4 Examples of poor signal quality ECG episodes due to artifacts (**a**: data00079) and noise (**b**: data00573). Red circles denote the reference QRS locations

tachycardia and sinus bradycardia put a challenge to the fixed threshold algorithms. Figure 3 shows two examples of sinus tachycardia and sinus bradycardia.

3. **Type C: poor signal quality due to artifact and noise**
   Dynamic/wearable ECGs are easily contaminated by artifacts and noises. Denoising approaches are limited here in frequency domain since the overlap between the noise frequency content and signal interest or in time domain since the similarity between QRS complex and some special artifacts (as shown in Fig. 4). The typical artifacts and noises [30] are from (1) electrode contact noise: loss of contact between the electrode and skin manifesting as sharp changes with saturation on the ECGs (usually due to an electrode being nearly or completely pulled off); (2) electrode movement artifacts, electrode movement away from the contact area on the skin, leading to variations in the impedance between the electrode and skin, which will cause potential variations in the ECG and usually manifest themselves as rapid (but continuous) baseline jumps or complete saturation; and (3) device noise, noises generated by the hardware of the device.

   Unfortunately, ECG is often contaminated by noise in similar morphologies caused the interest signal nearly invisible by the human eyes. To remove all

noises completely is impossible, so it is important to quantify the nature of noises in a particular dataset and choose an appropriate algorithm.

PVCs are conducted by the specialized conduction system and therefore are broad. The QRS width is at least 120 ms but often very broad around 160–200 ms. PVCs have many types and can be: monomorphic (QRS complexes with similar morphologies), multiformic (QRS complexes with different morphologies), bigeminy (every sinus beat followed by a PVC), or trigemini (every second sinus beat followed by a PVC).

2.2.2 Annotation

All QRS locations were beat-by-beat annotated by the Pan and Tompkin (P&T) detector and then manually hand-corrected by visual inspection. For more information about the CPSC 2019, please visit the URL (http://www.icbeb.org/Challenge.html).

2.3 Arrhythmia Database

2.3.1 Data Acquisition

ECGs in the hospital are huge but usually locally available. In addition, raw signals are usually invisible to public, and the data collected from different hospitals or different devices in the same hospital do not have unified format, making the sharing difficult. A database contains only wearable ECGs from 200 arrhythmia patients which was constructed. The patients aged between 18 and 82 years. All subjects were trained to wear a wearable ECG monitor device for at least 24 h even to 72 h to cover all possible onsets of arrhythmia. ECGs were collected with a sampling rate of 400 Hz, using a 12-bit sampling accuracy and a frequency response bandwidth of 0.05–45 Hz. The acquired signals are two-channel of the simulated limbs (lead I, II), which contain the information of various activities (sport, rest, and sleep), fully reflecting the real-world situation of patients' dynamic ECG signals.

2.3.2 Annotation Workflow

Performance of an algorithm or system must be evaluated against reference or "gold standard" annotations. An annotation platform was developed by co-operation of automatic classification algorithms and three cardiologists. First, ECG episodes were uploaded to the annotation platform in the standard ECG waveform presentation format. Then, an automatic step was applied to coarse annotate with commonly used algorithms. After that, two clinical cardiologists independently corrected the

automatic labels results. The third expert finally checked the results and identified the labels with different opinions and made a determination.

2.3.3 Data Type

The ECGs included in this database are intended to be used for the evaluation of algorithms for arrhythmias analysis on rhythm changes. However, subject to the small sample of heart disease, several subtypes only contain a few recordings (Table 4). The ANSI/AAMI EC57:1998/(R) 2008 standard [31] specifies that records of patients using pacemakers should not be considered. In addition, segments of data containing ventricular flutter or fibrillation (VF) were also excluded from the analysis.

1. **Sinus arrhythmia**
   For sinus arrhythmias, ECG appears periodically with a stable PR interval shorter than 0.2 s. Normal P wave in lead II usually shows upright wave in morphology and consistent lasting time in duration. Bradycardia is the situation where HR is less than 60 bpm, and tachycardia is the situation where HR is greater than 100 bpm. A long interval twice as long as normal PP interval between the second and the third P waves (see picture arrest in Fig. 5) indicates a sinus arrest. In this situation, the sinus node ceases to generate the electrical impulses for a variable period of time.
2. **Atrial arrhythmia**
   Premature beat initiates outside the sinoatrial node. Atrial premature beats, also called PAC, are ectopic beats that originate in the atria. Typically, atrial impulse propagates normally through the atrioventricular node into the cardiac ventricles, resulting in a normal, narrow QRS complex. Atrial premature beat is associated with an incomplete compensatory pause, meaning that the interval between the preceding and following sinus beats is less than twice the complete cycle. Single PAC clearly manifests a regular underlying rhythm; however, there is a premature beat which can be identified by irregular P wave with different sizes and shapes.

   Bigeminy is an abnormal pulse characterized by two beats in rapid succession followed by a pause. Atrial bigeminy occurs with pairing of atrial beats, as when an atrial extrasystole is coupled to each sinus beat. PACs may occur frequently or sporadically. Two PACs occurring consecutively are referred to an atrial couplet (coupled PAC). Paroxysmal atrial tachycardia has a high regular rate of about 140–250 bpm. In this situation, P waves are generally invisible and PR intervals are not measurable. AF has an atrial rate of more than 400 bpm and is distinguishable due to its irregular ventricular rate.

   Figure 6 shows examples of different atrial arrhythmias except AF. Typical features were marked to present the rule of clinical diagnosis. Figure 7 shows examples of AF pattern.

Table 4 The structure of the arrhythmia database

Class	Subtype	Number	ECG characteristics
N	Sinus bradycardia	N1_001 ~ N1_050	Upright, consistent, and normal in morphology and duration of P waves in analog leads I and II Regular PP intervals, and PR interval is between 0.12 and 0.25 s HR less than or equal to 60 bpm
	Sinus tachycardia	N2_001 ~ N2_43	Upright, consistent, and normal in morphology and duration of P waves in analog leads I and II Regular PP intervals, and PR interval is between 0.12 and 0.20 s HR greater than or equal to 100 bpm
	Sinus arrest	N3_001 ~ N3_004	Upright, consistent, and normal in morphology and duration of P waves in analog leads I and II PP interval prolonged significantly No common multiple relationship between the longest PP interval and basic sinus PP interval
A	Single PAC	A11_001 ~ A11_050 A12_001 ~ A12_050	P waves, as opposed to sinus P waves, appear before the normal narrow QRS with normal axis PR interval is equal or greater than PR interval which is between 0.12 and 2 s Atrial premature beats are associated with an incomplete compensatory pause, meaning that the interval between the preceding and following sinus beats is less than length of two complete cycles
	Bigeminy PAC		If an atrial premature beat follows every sinus beat, atrial bigeminy is said to exist
	Coupled PAC	A2_001 ~ A2_050	Two consecutive P waves appear before the normal narrow QRS, followed by a long compensatory interval
	NSAT/ SAT	A3_001 ~ A3_050	No less than three P waves appear continuously before the normal narrow QRS followed by a compensatory interval NSAT lasts no longer than 30 s, otherwise known as SAT
	AF	A4_001 ~ A4_050	Absence of distinct P waves and irregular RR intervals Abrupt onset of an irregularly narrow-complex tachycardia Small and irregular baseline wave with variable shape and amplitude, called F waves Frequency of F waves is about 350–600 bpm
V	Single PVC	V11_001 ~ V11_050 V12_001 ~ V12_050	Occurs early on or shortly after the T wave of the preceding sinus beat and is associated with a wide QRS complex Accompanied by secondary ST-T change (ST segment down and T wave inversion) Followed by a long interval. The PP interval around the premature ventricular complex is twice as the underlying sinus rate (PP interval)
	Bigeminy PVC		If a ventricular premature beat follows every sinus beat, bigeminy PVC is said to exist
	Coupled PVC	V2_001 ~V2_020	Two consecutive wide and deformed QRS complexes, followed by a long compensatory interval
	NSVT/ SVT	V3_001	No less than three wide and deformed QRS complexes appeared continuous and followed by a long compensatory interval NSVT lasts no longer than 30 s, otherwise known as SVT

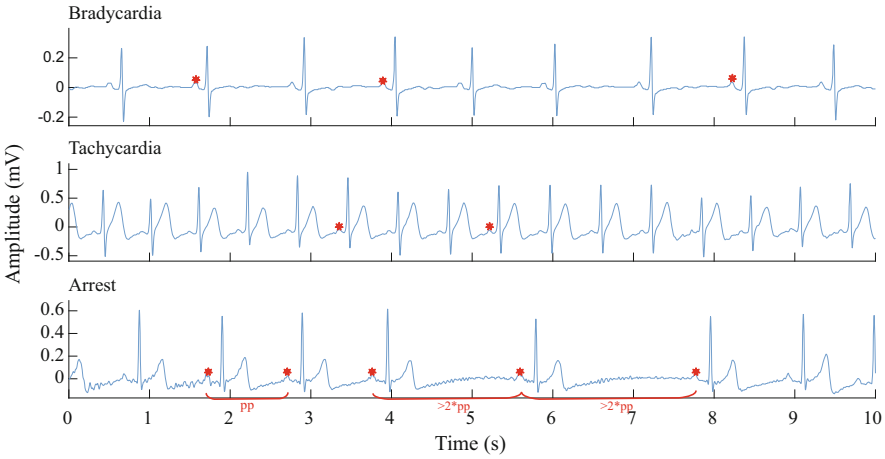

Fig. 5 From top to bottom, they are sinus bradycardia, sinus tachycardia, and sinus arrest; red asterisks represent the sinus P waves

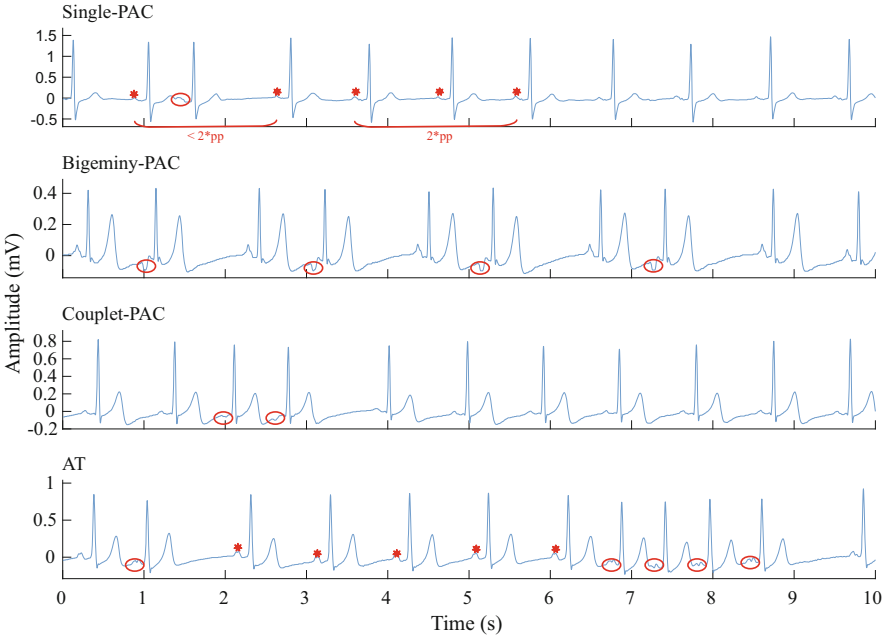

Fig. 6 From top to bottom, they are single PAC, bigeminy PAC, and couplet PAC and AT; red asterisks represent the sinus P waves, and abnormal P waves are circled by red oval

3. Ventricular arrhythmia

PVCs are premature ectopic beats originating in the ventricles of the heart and sometimes produces accompanying palpitations. A ventricular beat is initiated by

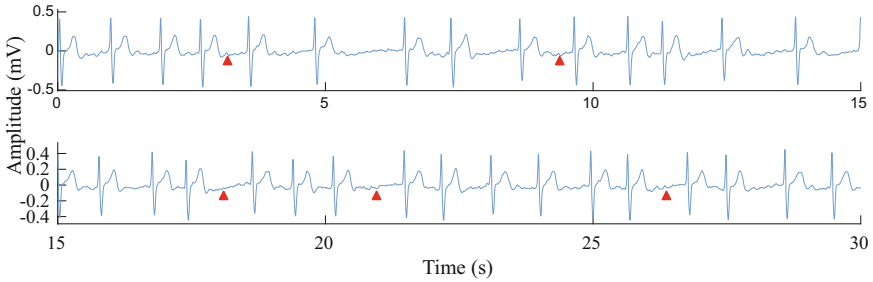

Fig. 7 The typical waveform of AF lasts for 30 s; red upper triangles represent irregular U waves

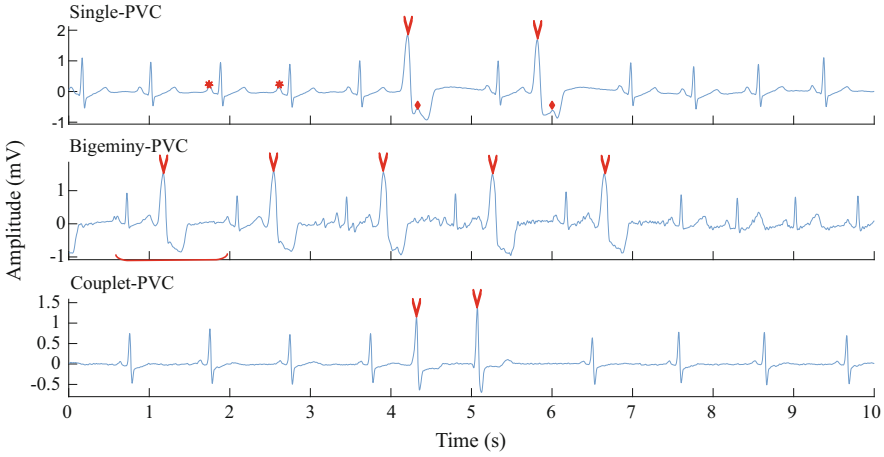

Fig. 8 From top to bottom, they are single PVC, bigeminy PVC, and couplet PVC; red asterisks represent the sinus P waves; abnormal P waves are assigned with red rhombic and big "V" points to the widened, bizarre QRS complex

an ectopic focus, which occurs before the usual sinoatrial beat, characterized by premature, widened, and bizarre QRS complexes, not preceded by a P wave. PVCs can occur during sinus rhythm or any other prevailing cardiac rhythm or cause a compensatory pause, if no retrograde atrial activation is present or if retrograde atrial encounters entrance block at the sinus node, thereby not producing disturbance in the sinus firing rate. PVC may be expressed as different forms: (1) single unifocal complexes, (2) multiple morphologies and fusion complexes, (3) multiple coupling intervals to the preceding sinus beats, (4) couplets or triplets, and (5) non-sustained or sustained ventricular tachycardia (Fig. 8).

2.4 AF Database

AF is the most common sustained cardiac arrhythmia [32], occurring in 1–2% of the general population, and is associated with significant mortality and morbidity through association of risk of death, stroke, heart failure and coronary artery disease, etc. [33, 34].

AF can be classified into paroxysmal AF, persistent AF, and permanent AF according to the European Society of Cardiovascular Diseases and North American Association of Pacing and Electrophysiology Arrhythmia [35]. Clinical 12-lead ECG detection is effective to diagnose patients suffering from persistent AF but may miss many cases of paroxysmal AF [36, 37]. For AF detection, atrial activity-based [38] and ventricular response-based methods are two commonly used methods. The success of the PhysioNet/CinC Challenge 2017 of "AF Classification from a short single lead ECG recording" significantly promotes the process of AF detection research [39]. Herein, the AF database contains 20 single-lead normal and AF recordings sampled at 400 Hz and lasts for 5 min.

2.5 Long-Term Arrhythmia ECGs

Abnormality of the cardiac conduction system can induce arrhythmia, abnormal heart rhythm, which can frequently lead to other cardiac diseases and complications, sometimes life-threatening [40]. In clinic, the number of abnormal heartbeats for a long-time monitoring window (such as 24 h) is a vital indicator for diagnostic decision. Thus, we collected and annotated this long-term arrhythmia ECG database, which contains two long-term ECG recordings, lasting for 24.51 and 23.11 h, respectively. The heartbeats are manually classified into the most common types: normal sinus (N), PAC (A), and PVC (V), represented numerically as 0, 1, and 2, respectively.

2.6 CPSC 2018 Database

A number of studies have investigated the performances of different detection/classification methods for the abnormal ECG types. However, many studies are generally limited in applicability because (1) the classification of normal and only one single abnormality was performed; (2) the data were not sufficient without the use of a separate out of sample test dataset, and only a small number of subjects were used, almost certainly resulting in over-fitting of the model and inflated statistics; and (3) failure to post the data (and any code to process the data) publicly so others may compare their results directly. Therefore, CPSC 2018 [19] contributes to a more comprehensive database to address these issues, including one normal type and eight

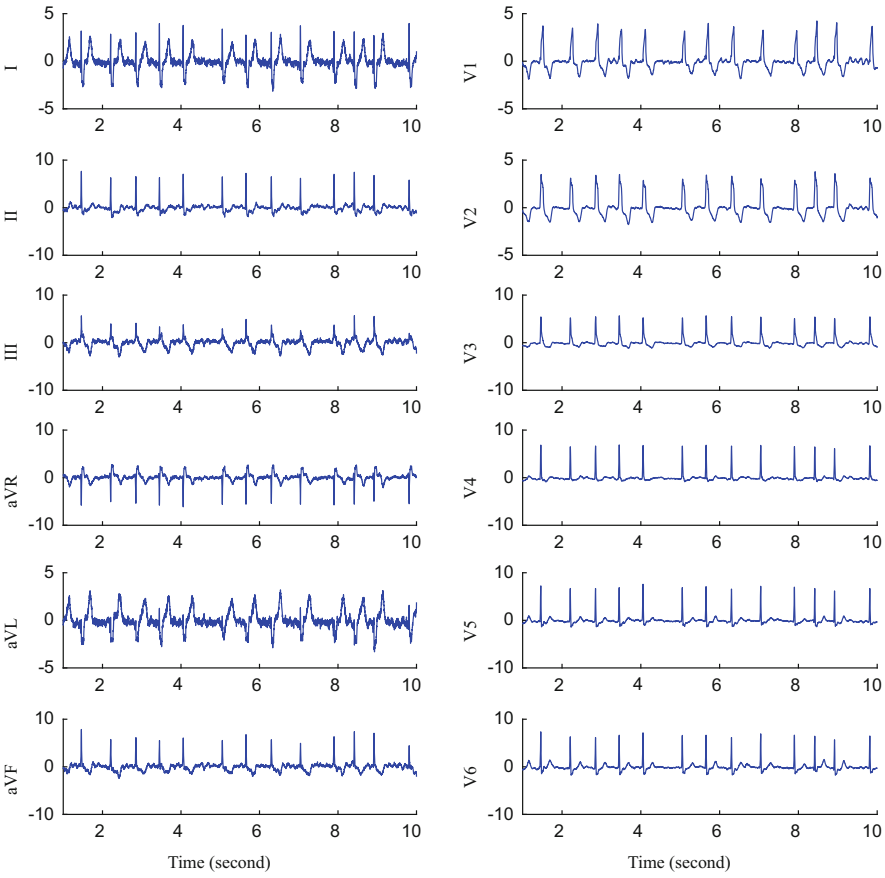

Fig. 9 Example of a 12-lead ECG waveform (AF and RBBB)

abnormal types, which are detailed as (1) AF, (2) I-AVB, (3) LBBB, (4) RBBB, (5) PAC, (6) PVC, (7) STD, and (8) STE.

ECG recordings were collected from 11 hospitals. The training set contains 6877 (female, 3178; male, 3699) 12-lead ECG recordings lasting from 6 s to just 60 s, sampled as 500 Hz. Figure 9 illustrates an example of the 12-lead ECG waveforms with AF and RBBB. Table 1 shows the details of the training data. All data are provided in MATLAB format (each recording is a .mat file containing the ECG data, as well as the patient sex and age information). For the most data, each recording has only one label from the nine types, while some recordings have two or three labels because the patient who provided the signals suffers from multiple diseases simultaneously. There are 477 recordings of this multi-label type in the training set. For more information on the Challenge and to download the data, please visit the website (http://www.icbeb.org/Challenge.html).

3 Discussion

Researchers have made great progress related to the automatic ECG diagnosis. Results presented in literatures are usually based on the analysis of the MIT-BIH database. However, unbalance issue hinders the model development of in-patient scheme [41]. The size and diversity of databases play a more and more important role in machine learning-based techniques [42]. One of the obstacles to achieve advances in this research domain is the databases, which have strict annotations and can be freely accessed online. Thus, elaboration of new databases is essentially important but challenging.

Noted that, many of approaches developed for ECG classification and QRS detection presented very high accuracy even nearly 100% in the commonly used MIT-BIH database but are hardly worked effectively in real environment. Considering the significance of individual diversity, databases existing for now are insufficient, especially for the dynamic and long-term ECG databases. Efforts to create new databases or even to increase the size of existing ones, as well as creating the corresponding annotations, have been made in several research areas involving image recognition, speech recognition, etc., to avoid unfair comparisons between methods [43].

This study describes the design and implementation of several annotated dynamic ECG database for signal quality assessment, QRS detection, and arrhythmia classification. These databases emphasize the challenge of signal processing and abnormal detection for dynamic ECGs, which are usually very noisy due to the unlimited physical activities. Currently, portable battery-operated systems such as mobile phones with wireless ECG sensors have the potential to be used in continuous and real-time cardiac function assessment that can be easily integrated into daily life. With the development in equipment and algorithms, new databases are needed. We hope this work can benefit the study of dynamic ECG analysis. We release the data listed in this chapter in our Lab website. If you want these data for research, please visit http://www.shelab.cn/Data for downloading the data.

References

1. Bansal, A., Joshi, R.: Portable out-of-hospital electrocardiography: a review of current technologies. J Arrhythm. **34**, 129 (2018)
2. Mishra, B., Arora, N., Vora, Y.: Wearable ECG for real time complex P-QRS-T detection and classification of various arrhythmias. In: 2019 11th International Conference on Communication Systems & Networks (COMSNETS), vol. 870. IEEE (2019)
3. Elgendi, M., Eskofier, B., Dokos, S., Abbott, D.: Revisiting QRS detection methodologies for portable, wearable, battery-operated, and wireless ECG systems. PLoS One. **9**, e84011 (2014)
4. Liu, C., Zhang, X., Zhao, L., Liu, F., Chen, X., Yao, Y., Li, J.: Signal quality assessment and lightweight QRS detection for wearable ECG smartvest system. IEEE Int. Things J. **6**, 1363 (2019)

5. Catarinucci, L., De Donno, D., Mainetti, L., Palano, L., Patrono, L., Stefanizzi, M.L., Tarricone, L.: An IoT-aware architecture for smart healthcare systems. IEEE Int. Things J. **2**, 515 (2015)
6. Yang, Z., Zhou, Q., Lei, L., Zheng, K., Xiang, W.: An IoT-cloud based wearable ECG monitoring system for smart healthcare. J. Med. Syst. **40**, 286 (2016)
7. Satija, U., Ramkumar, B., Manikandan, M.S.: Real-time signal quality-aware ECG telemetry system for IoT-based health care monitoring. IEEE Int. Things J. **4**, 815 (2017)
8. Zheng, Y., Ding, X., Poon, C.C.Y., Lo, B.P.L., Zhang, H., Zhou, X., Yang, G., Zhao, N., Zhang, Y.: Unobtrusive sensing and wearable devices for health informatics. IEEE Trans. Biomed. Eng. **61**, 1538 (2014)
9. Nemati, E., Deen, M.J., Mondal, T.: A wireless wearable ECG sensor for long-term applications. IEEE Commun. Mag. **50**, 36 (2012)
10. Lin, B., Chou, W., Wang, H., Huang, Y., Pan, J.: Development of novel non-contact electrodes for mobile electrocardiogram monitoring system. IEEE J. Transl. Eng. Health Med. **1**, 1 (2013)
11. Gargiulo, G., Bifulco, P., Cesarelli, M., Ruffo, M., Romano, M., Calvo, R.A., Jin, C., van Schaik, A.: An ultra-high input impedance ECG amplifier for long-term monitoring of athletes. Med. Devices (Auckland, NZ). **3**(1), (2010)
12. Arcelus, A., Sardar, M., Mihailidis, A.: Design of a capacitive ECG sensor for unobtrusive heart rate measurements. In: 2013 IEEE International Instrumentation and Measurement Technology Conference (I2MTC), vol. 407. IEEE (2013)
13. Friesen, G.M., Jannett, T.C., Jadallah, M.A., Yates, S.L., Quint, S.R., Nagle, H.T.: A comparison of the noise sensitivity of nine QRS detection algorithms. IEEE Trans. Biomed. Eng. **37**, 85 (1990)
14. Gribok, A.V., Chen, X., Reifman, J.: A robust method to estimate instantaneous heart rate from noisy electrocardiogram waveforms. Ann. Biomed. Eng. **39**, 824 (2011)
15. Khamis, H., Weiss, R., Xie, Y., Chang, C., Lovell, N.H., Redmond, S.J.: QRS detection algorithm for telehealth electrocardiogram recordings. IEEE Trans. Biomed. Eng. **63**, 1377 (2016)
16. Goldberger, A.L., Amaral, L.A., Glass, L., Hausdorff, J.M., Ivanov, P.C., Mark, R.G., Mietus, J.E., Moody, G.B., Peng, C., Stanley, H.E.: PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. Circulation. **101**, e215 (2000)
17. Moody, G.B., Mark, R.G.: The impact of the MIT-BIH arrhythmia database. IEEE Eng. Med. Biol. Mag. **20**, 45 (2001)
18. Mark, R., Moody, G.: MIT-BIH arrhythmia database directory. Massachusetts Institute of Technology, Cambridge (1988)
19. Liu, F., Liu, C., Zhao, L., Zhang, X., Wu, X., Xu, X., Liu, Y., Ma, C., Wei, S., He, Z., Li, J., Kwee, E.N.Y.: An open access database for evaluating the algorithms of electrocardiogram rhythm and morphology abnormality detection. J. Med. Imaging Health Inf. **8**, 1368 (2018)
20. Gao, H., Liu, C., Wang, X., Zhao, L., Shen, Q., Ng, E., Li, J., An Open-Access, E.C.G.: Database for algorithm evaluation of QRS detection and heart rate estimation. J. Med. Imaging Health Inf. **9**, 1853 (2019)
21. Redmond, S.J., Lovell, N.H., Basilakis, J., Celler, B.G.: ECG quality measures in telecare monitoring. In: 2869. IEEE (2008)
22. Behar, J., Oster, J., Li, Q., Clifford, G.D.: ECG signal quality during arrhythmia and its application to false alarm reduction. IEEE Trans. Biomed. Eng. **60**, 1660 (2013)
23. Li, Q., Mark, R.G., Clifford, G.D.: Robust heart rate estimation from multiple asynchronous noisy sources using signal quality indices and a Kalman filter. Physiol. Meas. **29**, 15 (2007)
24. Silva, I., Moody, G.B., Celi, L.: Improving the quality of ECGs collected using mobile phones: the Physionet/Computing in Cardiology Challenge 2011. In: 2011 Computing in Cardiology, vol. 273. IEEE (2011)
25. Clifford, G.D., Behar, J., Li, Q., Rezek, I.: Signal quality indices and data fusion for determining clinical acceptability of electrocardiograms. Physiol. Meas. **33**, 1419 (2012)

26. Searle, A., Kirkup, L.: A direct comparison of wet, dry and insulating bioelectric recording electrodes. Physiol. Meas. **21**, 271 (2000)
27. Baek, J., An, J., Choi, J., Park, K., Lee, S.: Flexible polymeric dry electrodes for the long-term monitoring of ECG. Sens. Actuators A Phys. **143**, 423 (2008)
28. Tang, A.S., Wells, G.A., Talajic, M., Arnold, M.O., Sheldon, R., Connolly, S., Hohnloser, S.H., Nichol, G., Birnie, D.H., Sapp, J.L.: Cardiac-resynchronization therapy for mild-to-moderate heart failure. N. Engl. J. Med. **363**, 2385 (2010)
29. Huang, H., Liu, J., Zhu, Q., Wang, R., Hu, G.: Detection of inter-patient left and right bundle branch block heartbeats in ECG using ensemble classifiers. Biomed. Eng. Online. **13**, 72 (2014)
30. Clifford, G.D., Azuaje, F., Mcsharry, P.: ECG statistics, noise, artifacts, and missing data. Adv. Methods Tools ECG Data Anal. **6**, 18 (2006)
31. AAMI A., EC A.ANSI/AAMI. EC57:1998/(R)2008–Testing and reporting performance results of cardiac rhythm and ST segment measurement algorithms. American National Standards Institute, Arlington (2008)
32. Lip, G.Y.H., Fauchier, L., Freedman, S.B., Van Gelder, I., Natale, A., Gianni, C., Nattel, S., Potpara, T., Rienstra, M., Tse, H., Lane, D.A.: Atrial fibrillation. Nat. Rev. Dis. Primers. **2**, 16016 (2016)
33. Clifford, G.D., Liu, C., Moody, B., Li-Wei, H.L., Silva, I., Li, Q., Johnson, A.E., Mark, R.G.: AF classification from a short single lead ECG recording: the PhysioNet/Computing in Cardiology Challenge 2017. In: 2017 Computing in Cardiology (CinC), vol. 1. IEEE (2017)
34. European Heart Rhythm Association, European Association for Cardio-Thoracic Surgery, Camm, A.J., Kirchhof, P., Lip, G.Y., Schotten, U., Savelieva, I., Ernst, S., Van Gelder, I.C.: Guidelines for the management of atrial fibrillation: the Task Force for the Management of Atrial Fibrillation of the European Society of Cardiology (ESC). Eur. Heart J. **31**, 2369 (2010)
35. Levy, S., Camm, A.J., Saksena, S., Aliot, E., Breithardt, G., Crijns, H.J., Davies, D.W., Kay, G. N., Prystowsky, E.N., Sutton, R.: International consensus on nomenclature and classification of atrial fibrillation: a collaborative project of the Working Group on Arrhythmias and the Working Group of Cardiac Pacing of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. J. Cardiovasc. Electrophysiol. **14**, 443 (2003)
36. Zhao, L., Liu, C., Wei, S., Shen, Q., Zhou, F., Li, J.: A new entropy-based atrial fibrillation detection method for scanning wearable ECG recordings. Entropy. **20**, 904 (2018)
37. Camm, A.J., Corbucci, G., Padeletti, L.: Usefulness of continuous electrocardiographic monitoring for atrial fibrillation. Am. J. Cardiol. **110**, 270 (2012)
38. Ladavich, S., Ghoraani, B.: Rate-independent detection of atrial fibrillation by statistical modeling of atrial activity. Biomed. Signal Process. Control. **18**, 274 (2015)
39. Moody, G.: A new method for detecting atrial fibrillation using RR intervals. Comput. Cardiol. **10**, 227 (1983)
40. Oh, S.L., Ng, E.Y., San, T.R., Acharya, U.R.: Automated diagnosis of arrhythmia using combination of CNN and LSTM techniques with variable length heart beats. Comput. Biol. Med. **102**, 278 (2018)
41. Luz, E.J.D.S., Schwartz, W.R., Cámara-Chávez, G., Menotti, D.: ECG-based heartbeat classification for arrhythmia detection: a survey. Comput. Methods Prog Biomed. **127**, 144 (2016)
42. Banko, M., Brill, E.: Scaling to very very large corpora for natural language disambiguation. In: Proceedings of the 39th Annual Meeting on Association for Computational Linguistics, vol. 26. Association for Computational Linguistics (2001)
43. Torralba, A., Efros, A.A.: Unbiased look at dataset bias. In: CVPR, vol. 1, pp. 7. Citeseer (2011)
44. http://www.shelab.cn/Data

Part III
Relevance of Signal Quality

An Overview of Signal Quality Indices on Dynamic ECG Signal Quality Assessment

Feifei Liu, Shoushui Wei, Fei Lin, Xinge Jiang, and Chengyu Liu

Abstract With the rapid development of wearable ECG medical devices, it is an imperious demand to evaluate the quality of dynamic ECG signals. Thus, a lot of signal quality indices (SQIs) have been proposed in the past few years. In this chapter, we review the analysis performances of SQIs from time-domain, frequency-domain, joint time-frequency, self-correlation, cross-correlation, entropy methods. We then illustrate the SQI performances using real clinical data, allowing a comparison of the SQIs. The performance of 26 SQIs was analyzed and discussed systematically.

Keywords Electrocardiogram (ECG) · Signal quality assessment (SQA) · Signal quality index (SQI)

F. Liu (✉)
School of Science, Shandong Jianzhu University, Jinan, China

School of Instrument Science and Engineering, Southeast University, Nanjing, China

S. Wei
School of Control Science and Engineering, Shandong University, Jinan, China
e-mail: sswei@sdu.edu.cn

F. Lin
Shandong ZhongYang Health Technology Group Co. Ltd, Jinan, China

X. Jiang
School of Information Science and Electrical Engineering, Shandong Jiaotong University, Jinan, China

C. Liu (✉)
The State Key Laboratory of Bioelectronics, School of Instrument Science and Engineering, Southeast University, Nanjing, China
e-mail: chengyu@seu.edu.cn

C. Liu, J. Li (eds.), *Feature Engineering and Computational Intelligence in ECG Monitoring*, https://doi.org/10.1007/978-981-15-3824-7_3

1 Introduction

Cardiovascular disease (CVD) is the leading cause of death globally: more people die annually from CVDs than from any other causes. According to the World Health Organization [1], CVDs take 17.9 million lives every year, accounting for 31% of all global deaths. Early detection and prevention are very important for the treatment of CVDs. Electrocardiogram (ECG) is an important medical tool for diagnosing heart disease and evaluating cardiac function. Recent years have witnessed rapid developments of wearable ECG systems for continuous ECG recording [2]. It is a needful tool for the early detection of cardiovascular diseases (CVDs) [3]. However, because of the bad electrode contacting or wrong electrode positioning on account of wrong operation without supervision in a remote condition, dynamic ECG signals were subjected to poor signal quality [4]. These poor signal qualities bring great challenges to the wearable ECG system, which hinders the reliable manual or automated measurement [5], increases the risk of false alerts and misdiagnosis [6], and increases the workload of doctors [7]. All of these problems will limit the telemedicine application in remote areas [5]. So automated signal quality assessment (SQA) is an important work need to be done. First, an alarm of degraded signal quality can remind the user to check the location of the electrode [8]. Then, the unavailable ECGs with the poor signal quality can be deleted directly, so it avoids the network congestion in the process of data transmission [9]. Therefore, the reliable signals can be provided for CVD detecting [10].

The issue of ECG quality evaluation has recently become a hot topic. A large number of signal quality indices (SQIs) have emerged [11–17], including time-domain features, frequency-domain feature, the features of the QRS waves, nonlinear characteristic, etc. In 2008, Redmond et al. developed a signal artifact masking method, which was the first study about the ECG signal quality in a remote medical system [12]. In the same year, Li et al. [14] developed four SQIs (bSQI, iSQI, kSQI, and sSQI). Based on these four SQIs, Clifford et al. got very good grades in the 2011 PhysioNet Computing in Cardiology Challenge: Improving the Quality of ECGs Collected using Mobile Phones [15]. Since then, a lot of SQIs have been developed [3], including six wave features from time domain proposed by Marco [16] and a signal quality matrix employed by Xia [10]. Typical SQIs were extensively studied in previous works [8, 18, 19]. However, according to the existing literature, the performance of these SQIs has not been analyzed and discussed systematically. The aim of this chapter is to do this work. First, the issue of lead-fall detection, a very important part for ultra-long-term (14-day) dynamic ECGs, needs to be considered. Next, the SQIs based on different signal characteristics such as time- and frequency-domain features or QRS wave information are discussed.

2 Typical Signal Quality Indexes for Dynamic ECG Signal Quality Assessment

The research in the literature [44] found that for high signal quality ECG recordings, most QRS detection algorithms had very high detection accuracy (F1 >99%), whereas the F1 results decrease significantly for the poor signal quality ECG signals (all <80%). For the resting ECG recordings, the signal qualities are very well. And the preprocessing program could solve the problem from the slight noise. However, dynamic ECGs suffer from the problem of lead fall due to the bad electrode contact or incorrect electrode positioning due to unsupervised operators in a remote condition [4]. A variety of SQIs have sprung up for dynamic ECGs, including time-domain, frequency-domain, joint time-frequency, self-correlation, cross-correlation, entropy methods, etc.

2.1 Lead-Fall Detection

Liu et al. found that lead-fall detection used as an initial classification could decrease data volume and computation load for further analysis [19]. ECG is detected as lead-fall signal not only when the most portion of samples keep constant, but also the random fluctuation without any rules because of the outside noises, as shown in Fig. 1. This figure shows several typical cases of lead fall from the 2011 PhysioNet/CinC Challenge. Many research works have been done for lead-fall detection, including the following several methods:

1. Constant voltage ECG excerpts of at least 1 s length are searched in all leads [20].
2. Range (min–max) below a global disconnection threshold or signal amplitude above low limit for at least certain percent of the total lead duration [21].
3. Low-frequency time marginal energy defined as the peak value of the energy distribution in the frequency band [0, 0.5] Hz. High values of this parameter indicated the presence of prominent baseline drift or constant amplitude level ("flat line") in the original ECG signal [16].
4. Leads with constant voltage were labeled as "missing." The voltage is considered constant if voltage variations in a given lead are less than or equal to four times the maximum voltage resolution (4×5 μV) [22].
5. Zero-line detection for 80% of the lead or >40% portion of samples above ±2 mV [23].

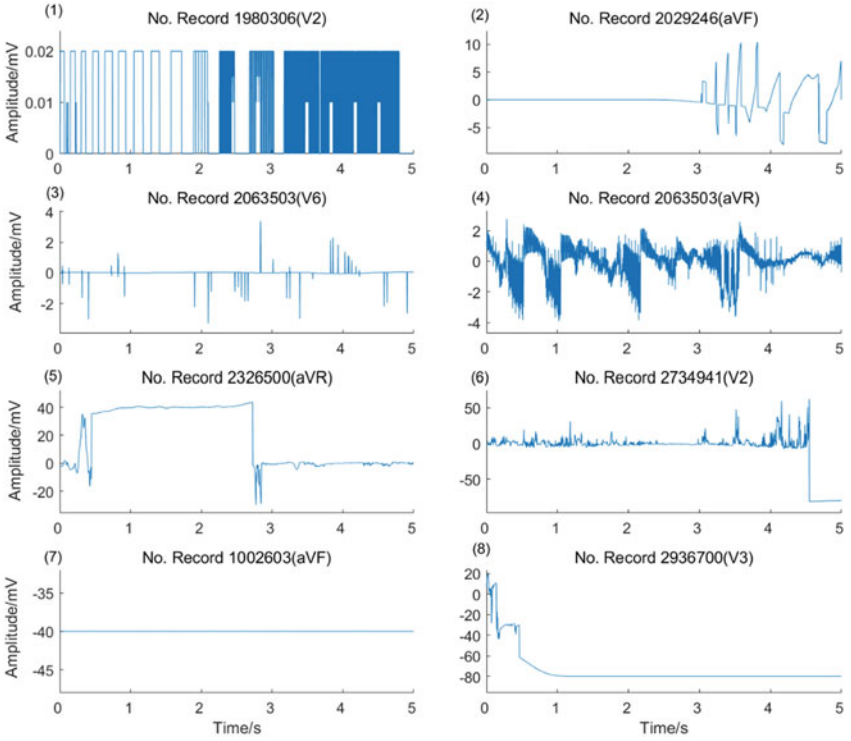

Fig. 1 Typical cases of lead fall from the 2011 PhysioNet/CinC Challenge

2.2 *Time-Domain Features*

1. **sSQI:** sSQI is the third moment (skewness) of the ECG signal distribution [8, 14] and is defined as:

$$\mathrm{sSQI} = \left| \frac{1}{M} \sum_{i=1}^{M} \left[\frac{x_i - \mu}{\sigma} \right]^3 \right| \tag{1}$$

where x_i is the ECG signal with N sample points, μ is the signal mean, σ is the standard deviation (SD), and $|\cdot|$ means the absolute value.

2. **kSQI:** kSQI is the fourth moment (kurtosis) of the ECG signal distribution [8, 14] and is defined as:

$$\mathrm{kSQI} = \frac{1}{M} \sum_{i=1}^{M} \left[\frac{x_i - \mu}{\sigma} \right]^4 \tag{2}$$

where all parameters have the same meanings with sSQI.

3. **SDN-SQI:** Estrella et al. [24] developed the noise maps for quantitative and clinical severity toward long-term ECG monitoring. Traditionally, dealing with poor-quality ECG has been faced as a denoising problem for which the target consists of improving some quality metrics, such as the root mean square (RMS) or the signal-to-noise ratio (SNR), which are often measured on artificially contaminated ECGs. Recent quantitative analysis has been performed within this framework using different signal processing techniques. SDN is a novel measure, which was designed to take into account events that make the signal unreadable, such as signal loss or gain saturation due to electrode disconnection. The proposed SDN-SQI was extracted by the following steps: (a) the standard deviation of the signal was computed in blocks of 0.5 s; (b) for every 10 blocks, the mean and the standard deviation were calculated; and (c) finally, the mean plus twice the standard deviation was used as a measure of the noise for each block.
4. **PLI-SQI:** The powerline interference (PLI) was also analyzed in the literature [24], because it is the most usual types of noise in cardiac records, and they are also easy to extract from the ECG. The quantification of PLI-SQI was made with a notch filter with the center frequency at 50 Hz.

2.3 Frequency-Domain Features

1. **purSQI:** Signal purity of ECG [25]:

$$\text{purSQI} = \frac{(\varpi_2(k))^2}{(\varpi_0(k)\varpi_4(k))} \tag{3}$$

where $\varpi_n = \int_{-\pi}^{\pi} \omega^n P(e^{j\omega}) \mathrm{d}\omega$, $P(e^{j\omega})$ is the power spectrum of the ECG in the analysis window, and $\omega = 2\pi f$.

2. **basSQI:** The relative power in the baseline:

$$\text{basSQI} = 1 - \int_{0\ \text{Hz}}^{1\text{Hz}} P(f)\mathrm{d}f \Big/ \int_{0\ \text{Hz}}^{40\text{Hz}} P(f)\mathrm{d}f \tag{4}$$

3. **pSQI:** pSQI assesses the power spectrum distribution feature. ECG waveform usually has a frequency range of 0.05–125 Hz for clinical diagnosis and a frequency range of 0.05–45 Hz for clinical monitoring. High signal quality ECGs usually have a distinguishable QRS complex, which has a frequency range from several to a dozen of Hz [8, 14]. So, the ratio of power spectral density in the QRS energy band to that in the overall energy band provides a useful measure, and thus, pSQI is defined as:

$$\mathrm{RpSQI} = \int_{5\ \mathrm{Hz}}^{15\mathrm{Hz}} P(f)\mathrm{d}f \Big/ \int_{5\ \mathrm{Hz}}^{40\mathrm{Hz}} P(f)\mathrm{d}f \tag{5}$$

where $P(f)$ is the autoregressive (AR) model spectrum and the Burg algorithm is adopted for parameter estimation.

4. **LpSQI:** The low-frequency power spectrum distribution feature: The ratio of the power spectral density in low-frequency band compared to the power spectral density in the overall signal is:

$$\mathrm{LpSQI} = \int_{0\ \mathrm{Hz}}^{3\mathrm{Hz}} P(f)\mathrm{d}f \Big/ \int_{0\ \mathrm{Hz}}^{100\mathrm{Hz}} P(f)\mathrm{d}f \tag{6}$$

5. **MpSQI:** The main frequency power spectrum distribution of ECG waveforms: The ratio of the sum of the power of the low-frequency ECG between frequencies, f, of 5 and 35 Hz to the power between 0 and 100 Hz is:

$$\mathrm{MpSQI} = \int_{5\ \mathrm{Hz}}^{35\mathrm{Hz}} P(f)\mathrm{d}f \Big/ \int_{0\ \mathrm{Hz}}^{100\mathrm{Hz}} P(f)\mathrm{d}f \tag{7}$$

6. **HpSQI:** The high-frequency power spectrum distribution feature: The ratio of the sum of the power of the high-frequency ECG between frequencies, f, of 40 and 100 Hz to the power between 0 and 100 Hz is:

$$\mathrm{HpSQI} = \int_{40\ \mathrm{Hz}}^{100\mathrm{Hz}} P(f)\mathrm{d}f \Big/ \int_{0\ \mathrm{Hz}}^{100\mathrm{Hz}} P(f)\mathrm{d}f \tag{8}$$

2.4 SQI Based on the QRS Waves

1. **GbSQI:** The original bSQI proposed by Li et al. was based on the comparison of two beat detectors on a single lead [14]. In the original bSQI, "ep_limited" [26] and "wqrs" [27] QRS detecting methods were employed. Liu et al. found that the performance of bSQI heavily depends on QRS detecting methods [28]. Therefore, GbSQI was proposed in the Liu's study based on ten QRS detectors, which is defined as:

$$\mathrm{GbSQI}(k,w) = \frac{(n-1) \times \mathrm{N\_matched}(k,w)}{\mathrm{N\_Method}_1(k,w) + \mathrm{N\_method}_2(k,w) + \cdots + \mathrm{N\_method}_n(k,w) - \mathrm{N\_matched}(k,w)} \tag{9}$$

   where N_matched(k,w) is the beat number agreed upon (within $\gamma = 150$ ms) and N_Methodn(k,w) is the beat number detected by the nth QRS detectors. Therefore, GbSQI ranges between 0 and 1 inclusively. Seven patterns of GbSQI (bSQI-2, bSQI-3, bSQI-4, bSQI-5, bSQI-6, bSQI-7, and bSQI-8) were tested. The GbSQI definition was sufficiently flexible to allow the use of an arbitrary number of R wave detectors. U3 [29], UNSW [30], DOM [31], and OKB [32] detectors were recommended for calculating GbSQI.
2. **iSQI:** Liu et al. developed the iSQI index [2]. It assesses the interval abnormal index for RR interval time series (with QRS locations) with a fixed time window. RR intervals are sorted in ascending order, and then, the 15% percentile value RR15 and 85% percentile value RR85 are selected, and iSQI is defined as:

$$\mathrm{iSQI} = \mathrm{RR}_{15}/\mathrm{RR}_{85} \tag{10}$$

   Similar to GbSQI, detection performance iSQI was heavily dependent on the performance of QRS detector selection.
3. **rsdSQI:** The relative standard deviation (STD) of QRS complex [33]:

$$\mathrm{rsdSQI} = \frac{1}{N}\sum_{i=1}^{N}\frac{\sigma r_i}{\sigma a_i * 2} \tag{11}$$

   where σr_i is the STD of each QRS (from R −0.07 to R +0.08 s) and σa_i is the STD around each QRS (from R −0.2 to R +0.2 s).
4. **eSQI:** The relative energy in the QRS complex [33]:

$$\mathrm{eSQI} = \frac{\sum_i Er_i}{Ea} \tag{12}$$

   where $Er_i = \sum x^2$ is the energy of each QRS segment (from R −0.07 to R +0.08 s), $i = 1, 2, 3, \ldots$ is the detected QRS complex in the analysis window, and Ea is the energy of the analysis window.

5. **bsSQI:** Baseline wander check in time domain [33]:

$$\mathrm{bsSQI} = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{Ra_i}{Ba_i}\right) \tag{13}$$

where Ra_i is the peak-to-peak amplitude of the ECG waveform around each QRS complex (from R −0.07 to R +0.08 s), R is the fiducial marker of QRS complex of QRS detector, $i = 1, 2, 3, \ldots$, and N is the detected QRS complex in the analysis window. Ba_i is the peak-to-peak amplitude of baseline (filtered by a 1-Hz low-pass filter, $H(z) = 0.0503/(1 - 0.9497z^{-1})$) around each QRS complex (from R −1 to R +1 s).
6. **hfSQI:** The relative amplitude of high-frequency noise [33]:

$$\mathrm{hfSQI} = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{Ra_i}{H_i}\right) \tag{14}$$

where the ECG signal (x) was multiplied by integer coefficient high-pass filter, with the difference equation $y(j) = x(j) - 2x(j-1) + x(j-2)$. Then, the filtered signal (y) was summed for every six points $s(j) = |y(j)| + |y(j-1)| + \cdots + |y(j-5)|$. Ra_i is the peak-to-peak amplitude of each QRS complex. H_i is the mean of $s(j)$ before each QRS complex (from R −0.28 to R −0.05 s).
7. **tSQI:** tSQI assesses the morphology consistency of any two ECG beats (with QRS locations) within a fixed time window. The correlation matrix $C = [c_{ij}]$ is constructed, where C_{ij} is the correlation coefficient between the ith beat and the jth beat. tSQI is defined as:

$$\mathrm{tSQI} = \frac{\sum_{i=1}^{M}\sum_{j=1}^{M} c_{ij}}{M^2} \tag{15}$$

where M is the beat number in a fixed time window.
8. **pcaSQI:** A ratio comprising of the sum of the eigenvalues associated with the five principal components over the sum of all eigenvalues obtained by principal component analysis applied to the time-aligned ECG cycles detected in the window by the eplimited algorithm, segmented at 100 ms either side of the R-peak [18].
9. **picaSQI:** Periodic component analysis (PiCA) periodicity measure of the ECG waveform nonlinear characteristic [33]. We define the covariance of the signal $X(t)$ as:

$$C_X(\tau) = E_t\{X(t+\tau)X(t)\} \tag{16}$$

where $E_t\{\bullet\}$ indicates averaging over $t\tau$, τ is a constant time lag. In order to apply this to the ECG signal, we replace τ with a variable τ_t that is calculated from beat-

to-beat ECG. Therefore, in each ECG cycle, the sample at time instant t is compared with the sample $t + \tau_t$, which is the sample with the same phase value in the succeeding ECG beat. Then,

$$\text{picaSQI} = \left|\frac{C_X(\tau_t)}{C_X(0)}\right| = \left|\frac{E_t\{X(t+\tau_t)X(t)\}}{E_t\left\{X(t)^2\right\}}\right| \tag{17}$$

2.5 Entropy Measures

Entropy as a measure of the complexity of an unstable time series has been widely used for signal processing and analyzing. Many algorithms about entropy have emerged. The approximate entropy (ApEn), which was proposed by Pincus in 1991, represents a simple index for the overall complexity and predictability of time series. It has been widely applied to clinical physiological signal studies [34]. It is derived from the computation of the correlation integral [35]. However, the ApEn produces biased estimation for the complexity of physiological signals with self-matching. To relieve this bias, Richman et al. proposed another statistic, the sample entropy (SampEn) [36]. SampEn is derived from approaches developed by Grassberger and coworkers. SampEn(m, r, N) is precisely the negative natural logarithm of the conditional probability that two sequences similar for m points remain similar at the next point, where self-matches are not included in calculating the probability. Thus, a lower value of SampEn also indicates more self-similarity in the time series.

However, for the ApEn and SampEn, the poor statistical stability has not been solved. The inherent reason for their poor statistical stability is that the two entropy measures are based on the Heaviside function of the classical sets, which is basically a two-state classifier that judges two vectors as either "similar" or "dissimilar," with no intermediate states [37, 38]. To overcome this problem, Chen et al. [12, 16] proposed a statistic named fuzzy entropy (FuzzyEn), in which the Heaviside function was replaced by the Zadeh fuzzy sets. However, these entropy measures are heavily dependent on the predetermined parameters and confined to data length. Therefore, Li [39] developed a novel measure—distribution entropy (DistEn). The DistEn took full advantage of the inherent information underlying the vector-to-vector distances in the state space by probability density estimation.

In this chapter, we analyze the performances of these four entropy measures on the ECG signal quality assessment. The detailed calculation processes of ApEn, SampEn, FuzzyEn, and DistEn are given in the Appendix.

2.6 Lempel–Ziv (LZ) Complexity

Lempel–Ziv (LZ) complexity is a measure of signal complexity and has been applied to a variety of biomedical signals. It is a useful approach for evaluating the irregularity of physiological time series. In most cases, the classic LZ complexity (CLZ) algorithm is executed by transforming an original signal into a binary sequence by comparing it with a preset median or mean value as the threshold [40]. That is, whenever the signal is larger than the threshold, one maps the signal to 1, otherwise, to 0. So the ECG signal should be coarse grained and transformed into a symbol sequence before the LZ calculation.

However, some studies have found most exiting LZ algorithms (CLZ and MLZ) with signal irregularity rather than complexity. Furthermore, the LZ values from random signals overlap with those from chaotic signals, corroborating to the inaccuracies found in LZ algorithms. Therefore, Zhang et al. developed a novel encoding LZ (ELZ) algorithm to directly and accurately quantify the irregularity, rather than the complexity, of a physiological time series.

In this chapter, we analyze the performances of LZ and ELZ algorithm on the ECG signal quality assessment. The detailed calculation processes of ApEn, SampEn, FuzzyEn, and DistEn are given in the Appendix.

3 Performance Analysis of SQIs

3.1 Database

In this study, 1000 recordings of standard 12-lead ECGs were employed to detect the performance of each SQIs. These data came from the 2011 PhysioNet Computing in Cardiology Challenge [9, 41, 42]. Each ECG recording was 10 s long, sampled 500 Hz. Of all these recordings, 773 were marked as "acceptable," 225 were marked as "unacceptable," and 2 were "unascertainable." But these tags were determined for 12-lead signals, not for the single lead. Therefore, the database could not be used to evaluate the performance of these indicators on single-lead ECG signal. For this reason, we relabel the ECG recording on a single ECG channel. For more information about relabeling, please refer to the reference [28]. At last, we obtained 12,000 10-s marked ECG records, a total of 9941 acceptable and a total of 2059 unacceptable 10-s ECG segments. In this chapter, lead-fall detection was used firstly, and 1071 10-s ECG segments from unacceptable group were detected as lead fall. Therefore, only 988 unacceptable segments were used for performance analysis of SQIs. Table **1** shows the detail description of the database.

Table 1 The description of the database

Description	No. of recordings	Sample frequency (Hz)	Record length (s)	Source
Lead fall	1071	500	10	2011 PhysioNet/CinC Challenge training set https://physionet.org/challenge/2011/
Unacceptable	9941	500	10	
Acceptable	988	500	10	
Total	12,000	–	–	

3.2 Evaluation Methods

In this chapter, sensitivity (Se), specificity (Sp), and modified accuracy (mAcc) were used for measuring the performance, which are defined in (18–20).

$$\mathrm{Se} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \times 100\% \tag{18}$$

$$\mathrm{Sp} = \frac{\mathrm{TN}}{\mathrm{TN} + \mathrm{FP}} \times 100\% \tag{19}$$

$$\mathrm{mAcc} = \frac{\mathrm{Se} + \mathrm{Sp}}{2} \times 100\% \tag{20}$$

where TP was the number of marked unacceptable segments correctly classified as "unacceptable." FP was the number of marked acceptable signals falsely classified as "unacceptable." TN was the number of marked acceptable signals correctly classified as "acceptable." FN was the number of marked acceptable signals falsely classified as "acceptable" [23].

In this chapter, 26 SQI indexes were selected to analyze the performances on ECG signal SQA. We tested the performances from both single SQI feature-based classifier and multiple SQI feature-based classifier. For single SQI feature-based classifier, each SQI feature was input to an SVM classier for training a classification model. For multiple SQI feature-based classifiers, the selected (sorted based on the performance of single SQI feature) features were input to an SVM classier for training a classification model. The Gaussian kernel was used in SVM. C and γ were optimized using a grid search method with the search range over C (from 0.5 to 724) and γ (from 4 to 32). For each test, a tenfold cross-validation was used. Figure 2 demonstrates the evaluation process. The threshold was also used for classifying.

3.3 Results

Table 2 shows the total classification performances (mAcc, Se, Sp) for single GbSQI feature-based classifiers. In order to analyze and compare the performances of each SQI better, we sorted the results of mAcc from high to low and presented the results

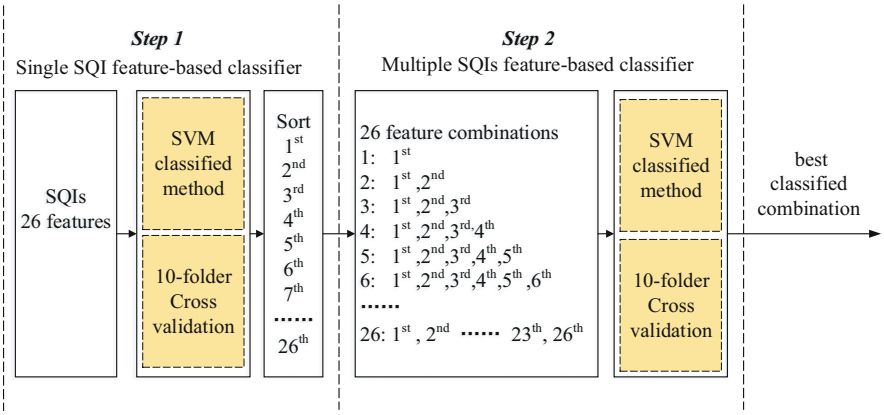

Fig. 2 Demonstration of the evaluation process

in Fig. 3. Table 3 and Fig. 4 show the total classification performances (mAcc, Se, Sp) for multiple GbSQI feature-based classifiers. In Table 2, the threshold detecting results are also presented.

As shown in Table 2 and Fig. 3, the SQI based on the QRS waves presented best performances, especially bSQI-2, bSQI-4, tSQI, and picaSQI (mAcc all >90%). These SQIs were highly dependent on the QRS detection algorithm. Second, the SQIs based on the nonlinear characteristic presented better results, especially SampEn-SQI, Fuzzy-SQI, and ELZ-SQI (mAcc all >85%). However, SQIs based on the time-domain features and frequency-domain features did not show good performances. In Table 2, the results from threshold detecting were shown. Because some SQI results are difficult to judge the range of the threshold with the naked eye, threshold detection was not carried out, such as basSQI, pSQI, HpSQI, and pcaSQI. For other SQIs, the results of threshold detecting and SVM classer were equal.

Based on the results of single SQI, we sorted the results mAcc from high to low and selected (sorted based on the performance of single SQI feature) features to training a multifeature classification model. The performance of the classifier increased initially and then held steady as the features increased. The classification performance was not proportional to feature number. As shown in Table 3 and Fig. 4, when the first 14 SQIs were selected, the classification model presented the best result (mAcc = 95.2%). The continuous increase of the SQIs reduced the performance of classification. Especially, if the four bottom SQIs (PLI-SQI, rsdSQI, purSQI, and hfSQI) were selected, the performance of classification was very terrible. Figure 4 did not show the results of the last four models in order to display the previous results better.

Table 2 The classified results of single feature for 26 SQIs

Category	No.	Signal quality index	SVM			Threshold detection
			mAcc	Se	Sp	
Time domain features	1	sSQI	80.06±2.14	87.84±4.55	72.95±1.47	81.02
	2	kSQI	80.42±3.04	84.18±6.17	77.37±4.34	82.51
	3	SDN-SQI	83.41±2.34	83.00±4.59	83.86±2.58	85.83
	4	PLI-SQI	68.08±1.87	65.25±5.72	70.41±5.52	74.55
Frequency domain features	5	purSQI	52.78±0.22	99.47±0.53	0.79±0.76	79.10
	6	basSQI	75.49±2.70	63.49±4.30	87.92±2.29	--
	7	bsSQI	82.22±1.81	69.73±4.32	94.85±1.60	81.31
	8	pSQI	77.37±1.44	61.96±3.21	92.78±1.59	--
	9	HpSQI	83.66±2.19	73.13±4.06	93.93±0.92	--
	10	LpSQI	78.11±2.29	81.40±3.91	74.92±2.65	79.75
	11	MpSQI	82.99±1.60	92.57±3.32	73.28±1.56	85.15
SQI based on the QRS waves	12	tSQI	91.98±1.00	96.10±1.68	88.12±1.70	92.45
	13	iSQI	87.31±2.88	87.09±4.73	87.87±3.51	89.29
	14	rsdSQI	67.18±12.03	97.70±7.27	9.91±31.34	87.50
	15	hfSQI	50.00±0.00	100.00±0	0.00±0.00	79.21
	16	eSQI	76.22±2.09	85.66±4.57	66.57±3.81	84.33
	17	picaSQI	90.83±1.35	94.84±1.65	86.93±2.39	91.66
	18	pcaSQI	74.24±2.45	90.37±3.87	57.73±2.59	--
	19	bSQI-4	93.70±1.49	93.43±2.98	93.79±1.25	94.00
	20	bSQI-2	93.84±1.47	93.43±2.58	94.18±1.45	94.13
Nonlinear characteristic	21	ApEn-SQI	83.28±2.33	84.46±3.96	81.70±1.91	75.69
	22	SampEn-SQI	85.57±2.52	79.64±4.92	91.42±2.55	85.23
	23	Fuzzy-SQI	86.33±2.60	80.47±4.98	91.89±1.91	85.41
	24	DistEn-SQI	80.55±2.42	62.09±4.62	98.98±0.57	76.01
	25	LZ-SQI	83.34±1.52	77.42±4.56	89.23±2.97	83.89
	26	ELZ-SQI	86.61±2.25	80.99±4.96	91.72±1.25	86.55

4 Discussion and Conclusion

Results demonstrated that the SQIs based on the QRS waves presented best performances. These SQIs were highly dependent on the QRS detection algorithm. Therefore, QRS detector selection was quite important. Literature [28] pointed out that two aspects should be considered for the QRS selection: one is the own character of the QRS detection algorithm, and the other is the relationship between the selected detectors. Different QRS detection algorithms were sensitive to different types of noise [43]. First, the selected detectors should have high sensitivity and specificity in QRS detection. Second, the selected detectors should be complementary. In this chapter, the constructions of bSQI-2 and bSQI-4 are in full compliance with the literature [28] recommendations. For bSQI-2, the QRS detectors U3 [29] and UNSW [30] were selected, and for bSQI-4, the QRS detectors OKB [32], UNSW, U3, and

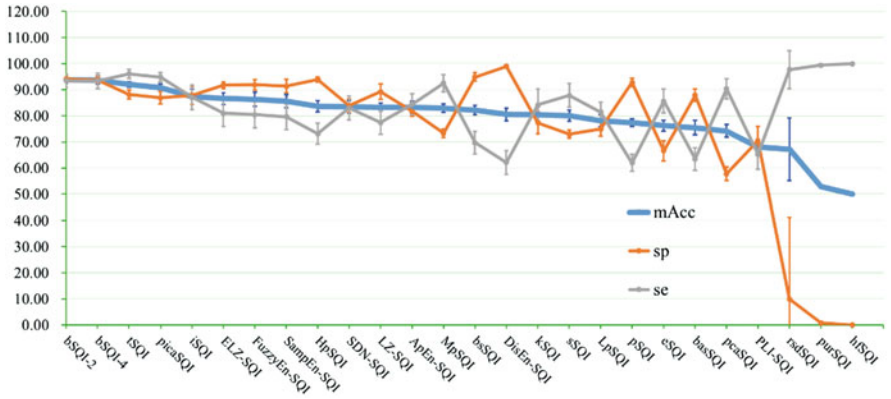

Fig. 3 The line chart of the classified results (mAcc, Se, Sp) of single feature

Table 3 The classified results of multifeature combination for 26 SQIs

The number of	Signal quality	mAcc	Se	Sp
1	bSQI-2	93.82±1.47	94.21±1.42	93.43±2.58
2	bSQI-4	94.36±1.44	94.38±1.45	94.35±2.91
3	tSQI	94.11±1.05	92.35±1.60	95.86±2.22
4	picaSQI	93.85±0.96	91.94±1.91	95.76±2.12
5	iSQI	93.92±1.06	92.03±1.96	95.81±2.26
6	ELZ-SQI	94.05±0.85	92.33±1.82	95.76±1.96
7	FuzzyEn-SQI	94.33±0.94	92.52±1.88	96.15±1.66
8	SampEn-SQI	94.24±1.02	92.42±1.98	96.06±1.64
9	HpSQI	94.58±0.98	93.05±1.93	96.10±1.73
10	SDN-SQI	94.63±1.01	93.39±1.77	95.86±1.75
11	LZ-SQI	94.64±0.92	93.42±1.67	95.86±1.72
12	ApEn-SQI	95.03±0.85	93.96±1.82	96.10±2.96
13	MpSQI	95.00±1.00	93.90±1.95	96.10±2.31
14	bsSQI	95.20±0.83	94.25±1.71	96.16±1.85
15	DisEn-SQI	94.47±0.86	92.69±1.81	96.25±1.43
16	kSQI	94.46±0.84	92.68±1.80	96.25±1.43
17	sSQI	94.42±0.73	92.54±1.67	96.30±1.65
18	LpSQI	94.48±0.76	92.57±1.70	96.39±1.60
19	pSQI	94.35±0.78	92.40±1.83	96.30±1.37
20	eSQI	94.35±0.83	92.41±1.82	96.30±1.45
21	basSQI	94.35±0.87	92.41±1.83	96.30±1.45
22	pcaSQI	94.34±0.84	92.44±1.86	96.25±1.37
23	PLI-SQI	50.00	0.00	100.00
24	rsdSQI	50.00	0.00	100.00
25	purSQI	50.00	0.00	100.00
26	hfSQI	50.00	0.00	100.00

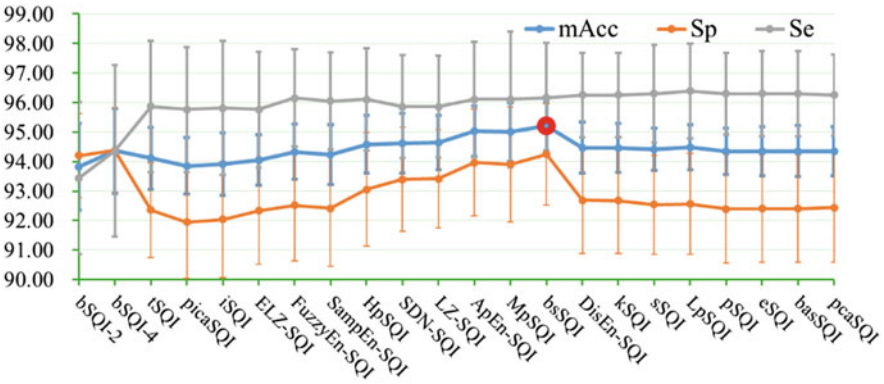

Fig. 4 The line chart of the classified results (mAcc, Se, Sp) of multifeature combination

DOM [31] were used. For other indexes based on QRS detecting, OKB detector was employed. This detector presented the best R-wave detecting performance on the six ECG databases, especially dynamic ECG signals [44]. For the nonlinear characteristic, ELZ-SQI, Fuzzy-SQI, and SampEn-SQI had good performances (mAcc all >85%).

From Table 2, the reason for the difference between SVM and threshold detecting is that the tenfold cross-validation was employed for the SVM and the best threshold was chosen from all the samples. As the results shown, the SQI selection is very important. The best result from bSQI-2 was 93.84%, whereas the worst result of purSQI was only 52.78%. The results are very different. In this chapter, the SQIs (bSQI-2, bSQI-4, tSQI, picaSQI, SampEn-SQI, Fuzzy-SQI, and ELZ-SQI) were recommended.

There are some limitations in this study. First, in this study, only 26 SQIs were analyzed although there are lots of SQIs at present, which were not considered because of space constrain. Second, because some algorithms were published in a theoretical way without online code and some literatures only include a few guidelines for real implementation and do not fully explain the necessary preprocessing operations, some SQIs were coded by ourselves. Therefore, the detection results in this study may be different from those in the other literatures, but these differences are slight. Third, a unified signal normalization processing was performed before SQI calculation for the fair comparisons among different SQI methods.

In this chapter, a systematical evaluation work was performed on 26 widely used SQI algorithms. Seven SQIs (bSQI-2, bSQI-4, tSQI, picaSQI, SampEn-SQI, Fuzzy-SQI, and ELZ-SQI) were recommended. The first 14 SQI combinations showed the best performances.

Appendix

A.1 ApEn_SQI

To solve the problems of short and noisy recordings in physiological signals, Pincus [45] presented approximate entropy (ApEn) as a measure of complexity that is applicable to noisy, medium-sized datasets. For an N sample time series $\{\boldsymbol{u}(\boldsymbol{i}) : \mathbf{1} \leq \boldsymbol{i} \leq \boldsymbol{N}\}$, given m, form vector sequences $\mathbf{X}_{\mathbf{1}}^{\boldsymbol{m}}$ through $\mathbf{X}_{\boldsymbol{N}-\boldsymbol{m}+\mathbf{1}}^{\boldsymbol{m}}$ as

$$X_i^m = \{u(i), u(i+1), \ldots, u(i+m-1)\}, \qquad i = 1, \ldots, N-m+1 \tag{21}$$

where m is the length of compared window. For each $\boldsymbol{i} \leq \boldsymbol{N} - \boldsymbol{m} + \mathbf{1}$, let $\boldsymbol{C}_{\boldsymbol{i}}^{\boldsymbol{m}}(\boldsymbol{r})$ be $(N - m + 1)^{-1}$ times the number of vectors $\mathbf{X}_{\boldsymbol{j}}^{\boldsymbol{m}}$ within r of $\mathbf{X}_{\boldsymbol{i}}^{\boldsymbol{m}}$. By defining

$$\phi^m(r) = (N-m+1)^{-1} \sum_{i=1}^{N-m+1} \ln C_i^m(r) \tag{22}$$

where ln is the natural logarithm, Pincus defined the parameter:

$$\mathrm{ApEn}(m, r) = \lim_{x \to \infty} \left[\phi^m(r) - \phi^{m+1}(r)\right] \tag{23}$$

In this chapter, ApEn_SQI was defined as:

$$\mathrm{ApEn\_SQI} = \mathrm{ApEn}(m, r) \tag{24}$$

where $m = 2$, r is equal to the 0.15 times of standard deviation of the signal.

SampEn_SQI

Sample entropy (SampEn) is a modification of approximate entropy (ApEn), used for assessing the complexity of physiological time-series signals, diagnosing diseased states [46]. Now assume we have a time-series data set of length $N = \{x_1, x_2, x_3, \ldots, x_N\}$ with a constant time interval τ. We define a template vector of length m, such that $X_m(i) = \{x_i, x_{i+1}, x_{i+2}, \ldots, x_{i+m-1}\}$ and the distance function $d[X_m(i), X_m(j)](i \neq j)$ is to be the Chebyshev distance. We define the sample entropy to be

$$\mathrm{SampEn}(m, r, N) = -\ln\left[\frac{A^m(r)}{B^m(r)}\right] \tag{25}$$

where B_i^m is $(N-m-1)^{-1}$ times the number of vectors X_j^m within r of X_i^m, where j ranges from 1 to $N-m$, and $j \neq i$ to exclude self-matches, and then define

$$B^m(r) = (N-m)^{-1} \sum_{i=1}^{N-m} B_i^m(r) \tag{26}$$

where A_i^m is $(N-m-1)^{-1}$ times the number of vectors X_j^{m+1} within r of X_i^{m+1}, where j ranges from 1 to $N-m$, and $j \neq i$ to exclude self-matches, and then define

$$A^m(r) = (N-m)^{-1} \sum_{i=1}^{N-m} A_i^m(r) \tag{27}$$

In this chapter, ApEn_SQI was defined as:

$$\text{SampEn\_SQI} = \text{SampEn}(m, r, N) \tag{28}$$

where N is equal to the length of the signal, $m = 2$, r is equal to the 0.15 times of standard deviation of the signal.

FuzzyEn_SQI

FuzzyEn [46] excludes self-matches and considers only the first $N-m$ vectors of length m to ensure that X_i^m and X_i^{m+1} are defined for all $1 \leq i \leq N-m$. For times series $\{u(i) : 1 \leq i \leq N\}$, form vectors:

$$\left\{X_i^m = \{u(i), u(i+1), \ldots, u(i+m-1)\} - u0(i), i = 1, \ldots, N-m+1\right\} \tag{29}$$

where $u0(i) = m^{-1}\sum_{j=0}^{m-1} u(i+j)$.

For finite datasets, FuzzyEn can be estimated by the statistic:

$$\text{FuzzyEn}(m, r, N) = \ln \varphi^m(r) - \ln \varphi^{m+1}(r) \tag{30}$$

where $\varphi^m(r) = (N-m)^{-1}\sum_{i=1}^{N-m} \phi_i^m(r)$ and $\varphi^{m+1}(r) = (N-m)^{-1}\sum_{i=1}^{N-m} \phi_i^{m+1}(r)$. And $\phi_i^m(r) = (N-m-1)^{-1}\sum_{j=1, j\neq i}^{N-m} D_{ij}^m$.

Given vector X_i^m, calculate the similarity degree D_{ij}^m of its neighboring vector X_j^m to it through the similarity degree defined by a fuzzy function:

$$D_{ij}^m = \mu\left(d_{ij}^m, r\right) \tag{31}$$

where d_{ij}^m is the maximum absolute difference of the corresponding scalar components of X_i^m and X_j^m. For each vector $X_i^m(i-1, \ldots, N-m+1)$, averaging all the similarity degree of its neighboring vectors $X_j^m(\ j-1, \ldots, N-m+1, \text{and } j \neq i)$, we get $\phi_i^m(r)$.

For a ECG signal,

$$\text{FuzzyEn\_SQI} = \text{FuzzyEn}(m, r, N) \tag{32}$$

where N is equal to the length of the signal, $m = 2$, r is equal to the 0.15 times of standard deviation of the signal.

DistEn_SQI

Distribution entropy was established by Li et al. [39]. For times series $\{u(i) : 1 \leq i \leq N\}$, form vectors:

$$\{X(i) = \{u(i), u(i+1), \ldots, u(i+m-1)\} i = 1, \ldots, N-m\} \tag{33}$$

Here, m indicates the embedding dimension. Define the distance matrix $\mathbf{D} = \{d_{i,j}\}$ among vectors $\mathbf{X}(i)$ and $\mathbf{X}(j)$ for all $1 \leq i, j \leq N-m$, wherein $d_{i,j} = \max\{|u(i+k) - u(j+k)|, 0 \leq k \leq m-1\}$ is the Chebyshev distance between $\mathbf{X}(i)$ and $\mathbf{X}(j)$. The distribution characteristics of all $d_{i,j}$ for $1 \leq i, j \leq N-m$ should be complete quantification of the information underlying the distance matrix $\mathbf{D}$. We here apply the histogram approach to estimate the empirical probability density function of $\mathbf{D}$. If the histogram has M bins, we use p_t, $t = 1, 2, \ldots, M$ to denote the probability of each bin. To reduce bias, elements with $i = j$ are excluded when estimating the empirical probability density function.

Define the DistEn of $u(i)$ by the classical formula of Shannon entropy, that is

$$\text{DistEn}(m) = -\frac{1}{\log_2(M)} \sum_{t=1}^{M} p_t \log_2(p_t) \tag{34}$$

LZ Complexity

The calculation process of LZ complexity is summarized as follows [17, 40]. For CLZ complexity, the coarse-graining process is performed by comparing signal X with a threshold to transform X into a binary sequence R. That is, whenever the signal is larger than the threshold, one maps the signal to 1, otherwise, to 0. The

mean or median of the signal is usually selected as the threshold. The MLZ converts signal $X = x_1, x_2, \ldots, x_n$ to a 0, 1, 2, …, γ-sequence S, where γ is an integer number higher than 3. After the coarse-graining process, the LZ complexity counter $c(n)$ of the new symbol sequence can be calculated according to the rules. Let S and Q, respectively, denote two strings, and SQ is the concatenation of S and Q, whereas string $SQ\pi$ is derived from SQ after its last character is deleted (π means the operation to delete the last character in the string). Let $v(SQ\pi)$ denote the vocabulary of all different substrings of $SQ\pi$. Initially, $c(n) = 1$, $S = s_1$, and $Q = s_2$, and thus $SQ\pi = s_1$. In summary, $S = s_1, s_2, \ldots, s_r$, and $Q = s_{r+1}$, and thus $SQ\pi = s_1 s_2, \ldots, s_r$. If Q belongs to $v(SQ\pi)$, then s_{r+1}, that is, Q is a substring of $SQ\pi$, and so S does not change, and renew Q to be $s_{r+1}s_{r+2}$, and then judge if Q belongs to $v(SQ\pi)$ or not. This process is repeated until Q does not belong to $v(SQ\pi)$. Next, $Q = s_{r+1}s_{r+2}, \ldots, s_{r+i}$, which is not a substring of $SQ\pi = s_1 s_2, \ldots, s_r s_{r+1}, \ldots, s_{r+i-1}$; therefore, $c(n)$ is increased by 1. Subsequently, S is renewed to be $S = s_1 s_2, \ldots, s_{r+i}$, and $Q = s_{r+i+1}$. The procedures are repeated until Q is the last character. Concurrently, $c(n)$ is the number of different substrings (new pattern) contained in the new sequence. Finally, $c(n)$ can be normalized as:

$$C(n) = c(n)\frac{\log_a(n)}{n} \tag{35}$$

where n is the length of signal X, α is the number of possible symbols contained in the new sequence, and $C(n)$ is the normalized LZ complexity and denotes the arising rate of new patterns within the sequence. In practice, the normalized complexity $C(n)$, instead of $c(n)$, is considered.

ELZ Complexity

ELZ transforms each x_i contained within the original signal [17] $X = x_1, x_2, \ldots, x_n$ into a three-bit binary symbol $b_1(i)b_2(i)b_3(i)$, and the process is described as.

The first binary digit $b_1(i)$ is determined by comparing x_i with a threshold T_{mean} which is the mean of signal X, and it is defined as:

$$b_1 = \begin{cases} 0 & \text{if } x_i < T_{\text{mean}} \\ 1 & \text{if } x_i \geq T_{\text{mean}} \end{cases}, \quad i = 1, 2, \ldots n \tag{36}$$

The second binary digit $b_2(i)$ is determined by the difference between x_i and x_{i-1}, and it is defined as:

$$b_2 = \begin{cases} 0 & \text{if } x_i - x_{i-1} < 0 \\ 1 & \text{if } x_i - x_{i-1} \geq 0 \end{cases}, \quad i = 1, 2, \ldots n \tag{37}$$

where $b_2(1)$ is set to 0.

For the third binary digit $b_3(i)$, a variable Flag is first denoted as:

$$\text{Flag}(i) = \begin{cases} 0 & \text{if } |x_i - x_{i-1}| < \text{dm} \\ 1 & \text{if } |x_i - x_{i-1}| \geq \text{dm} \end{cases}, \quad i = 2, 3, \ldots n \tag{38}$$

where dm is the mean distance between adjacent points within signal X. If Flag(i) is 0, point x_i is relatively close to point x_{i-1}; otherwise, the two points are relatively far away. Subsequently, $b_3(i)$ is calculated as:

$$b_3(i) = \text{NOT}(b_2(i) \text{XOR Flag}(i)), \quad i = 2, 3, \ldots, n \tag{39}$$

where $b_3(1)$ is 0. Moreover, $b_2(i) = 1$ and Flag$(i) = 1$ mean that x_i is not only higher than x_{i-1} but also relatively farther from x_{i-1} compared with $b_2(i) = 1$ and Flag $(i) = 0$.

References

1. World Health Organization: Cardiovascular diseases (CVDs). https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds) (2019)
2. Liu, C., et al.: Signal quality assessment and lightweight QRS detection for wearable ECG SmartVest system. IEEE Internet Things J. **6**(2), 1363–1374 (2019)
3. Clifford, G.D., Moody, G.B.: Signal quality in cardiorespiratory monitoring. Physiol. Meas. **33** (9), 1–6 (2012)
4. Maan, A.C., Van Zwet, E.W., Man, S., Oliveira-Martens, S.M.M.: Assessment of signal quality and electrode placement in ECGs using a reconstruction matrix. In: Computing in Cardiology, Hangzhou, China, pp. 289–292 (2011)
5. Jekova, I., Krasteva, V., Christov, I., Abacherli, R.: Threshold-based system for noise detection in multilead ECG recordings. Physiol. Meas. **33**(9), 1463–1477 (2012)
6. Liu, C., Li, P., Zhao, L., Liu, F.: Real-time signal quality assessment for ECGs collected using mobile phones. In: Computing in Cardiology, Hangzhou, China, pp. 357–360 (2011)
7. Allen, J., Murray, A.: Assessing ECG signal quality on a coronary care unit. Physiol. Meas. **17** (4), 249–258 (1996)
8. Clifford, G.D., Behar, J., Li, Q., Rezek, I.: Signal quality indices and data fusion for determining clinical acceptability of electrocardiograms. Physiol. Meas. **33**(9), 1419–1433 (2012)
9. Silva, I., Moody, G.B., Celi, L.: Improving the quality of ECGs collected using mobile phones: the PhysioNet/Computing in Cardiology Challenge 2011. In: Computing in Cardiology, Hangzhou, China, vol. 6801, no. 1, pp. 273–276 (2011)
10. Xia, H., Garcia, G.A., Bains, J., Wortham, D.C., Zhao, X.: Matrix of regularity for improving the quality of ECGs. Physiol. Meas. **33**(9), 1535–1548 (2012)
11. Moody, G.B., Mark, R.G.: QRS morphology representation and noise estimation using the Karhunen-Loeve transform. In: Proceedings of the Computers in Cardiology, Jerusalem, Israel, pp. 269–272. IEEE (1989)

12. Redmond, S.J., Lovell, N.H., Basilakis, J., Celler, B.G.: ECG quality measures in telecare monitoring. In: 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Vancouver, BC, Canada, 20-25 Aug. 2008, pp. 2869–2872
13. Redmond, S.J., Xie, Y., Chang, D., Basilakis, J., Lovell, N.H.: Electrocardiogram signal quality measures for unsupervised telehealth environments. Physiol. Meas. **33**(9), 1517–1533 (2012)
14. Li, Q., Mark, R.G., Clifford, G.D.: Robust heart rate estimation from multiple asynchronous noisy sources using signal quality indices and a Kalman filter. Physiol. Meas. **29**(1), 15–32 (2008)
15. Clifford, G., Lopez, D., Li, D., Rezek, I.: Signal quality indices and data fusion for determining acceptability of electrocardiograms collected in noisy ambulatory environments. In: Computing in Cardiology, Hangzhou, China, pp. 285–288. IEEE (2011)
16. Di Marco, L.Y., et al.: Evaluation of an algorithm based on single-condition decision rules for binary classification of 12-lead ambulatory ECG recording quality. Physiol. Meas. **33**(9), 1435–1448 (2012)
17. Zhang, Y., Wei, S., Liu, H., Zhao, L., Liu, C.: A novel encoding Lempel-Ziv complexity algorithm for quantifying the irregularity of physiological time series. Comput. Methods Programs Biomed. **133**(8), 7–15 (2016)
18. Behar, J., Oster, J., Li, Q., Clifford, G.D.: ECG signal quality during arrhythmia and its application to false alarm reduction. IEEE Trans. Biomed. Eng. **60**(6), 1660–1666 (2013)
19. Zhang, Y.-t., Liu, C.-y., Wei, S.-s., Wei, C.-z., Liu, F.-f.: ECG quality assessment based on a kernel support vector machine and genetic algorithm with a feature matrix. J. Zhejiang Univ. Sci. C. **15**(7), 564–573 (2014)
20. Langley, P., et al.: An algorithm for assessment of quality of ECGs acquired via mobile telephones. In: 2011 Computing in Cardiology, pp. 281–284 (2011)
21. Noponen, K., Karsikas, M., Tiinanen, S., Kortelainen, J., Huikuri, H., Seppänen, T.: Electrocardiogram quality classification based on robust best subsets linear prediction error. In: 2011 Computing in Cardiology, pp. 365–368 (2011)
22. Johannesen, L., Galeotti, L.: Automatic ECG quality scoring methodology: mimicking human annotators. Physiol. Meas. **33**(9), 1479–1489 (2012)
23. Hayn, D., Jammerbund, B., Schreier, G.: QRS detection based ECG quality assessment. Physiol. Meas. **33**(9), 1449–1461 (2012)
24. Everss-Villalba, E., et al.: Noise maps for quantitative and clinical severity towards long-term ECG monitoring. Sensors. **17**(11), 2448 (2017)
25. Nemati, S., Malhotra, A., Clifford, G.D.: Data fusion for improved respiration rate estimation. EURASIP J. Adv. Signal Process. **2010**, 10 (2010)
26. Hamilton, P.S., Tompkins, W.J.: Quantitative investigation of QRS detection rules using the MIT/BIH arrhythmia database. IEEE Trans. Biomed. Eng. **33**(12), 1157–1165 (1986)
27. Zong, W., Moody, G., Jiang, D.: A robust open-source algorithm to detect onset and duration of QRS complexes. In: Computers in Cardiology, Thessaloniki Chalkidiki, Greece, pp. 737–740. IEEE (2003)
28. Liu, F., et al.: Dynamic ECG signal quality evaluation based on the generalized bSQI index. IEEE Access. **6**, 41892–41902 (2018)
29. Paoletti, M., Marchesi, C.: Discovering dangerous patterns in long-term ambulatory ECG recordings using a fast QRS detection algorithm and explorative data analysis. Comput. Methods Programs Biomed. **82**(1), 20–30 (2006)
30. Khamis, H., Weiss, R., Xie, Y., Chen, C.W., Lovell, N., Redmond, S.: QRS detection algorithm for telehealth electrocardiogram recordings. IEEE Trans. Biomed. Eng. **63**(7), 1377–1388 (2016)
31. Yeh, Y.C., Wang, W.J.: QRS complexes detection for ECG signal: the difference operation method (in English). Comput. Methods Prog. Biomed. **91**(3), 245–254 (2008)
32. Elgendi, M.: Fast QRS detection with an optimized knowledge-based method: evaluation on 11 standard ECG databases. Plos One. **8**(9), e735571–e735518 (2013)

33. Li, Q., Rajagopalan, C., Clifford, G.D.: A machine learning approach to multi-level ECG signal quality classification. Comput. Methods Programs Biomed. **117**(3), 435–447 (2014)
34. Liu, C., et al.: Analysis of heart rate variability using fuzzy measure entropy. Comput. Biol. Med. **43**(2), 100–108 (2013)
35. Pincus, S., Singer, B.H.: Randomness and degrees of irregularity. Proc. Natl. Acad. Sci. U. S. A. **93**(5), 2083–2088 (1996)
36. Richman, J.S., Moorman, J.R.: Physiological time-series analysis using approximate entropy and sample entropy. Am. J. Physiol. Heart Circ. Physiol. **278**(6), H2039–H2049 (2000)
37. Castiglioni, P., Rienzo, M.D.: How the threshold "r" influences approximate entropy analysis of heart-rate variability. In: 2008 Computers in Cardiology. Bologna, pp. 561–564 (2008)
38. Liu, C., et al.: Comparison of different threshold values for approximate entropy: application to investigate the heart rate variability between heart failure and healthy control groups. Physiol. Meas. **32**(2), 167–180 (2011)
39. Li, P., et al.: Assessing the complexity of short-term heartbeat interval series by distribution entropy. Med. Biol. Eng. Comput. **53**(1), 77–87 (2015)
40. Aboy, M., Hornero, R., Abásolo, D., Alvarez, D.: Interpretation of the Lempel-Ziv complexity measure in the context of biomedical signal analysis. IEEE Trans. Biomed. Eng. **53**(11), 2282–2288 (2006)
41. Celi, L.A., Sarmenta, L., Rotberg, J., Marcelo, A., Clifford, G.: Mobile care (Moca) for remote diagnosis and screening. J. Health Inform. Dev. Countries. **3**(1), 17–21 (2009)
42. Goldberger, A.L., et al.: Physiobank, physiotoolkit, and physionet components of a new research resource for complex physiologic signals. Circulation. **101**(23), e215–e220 (2000)
43. Friesen, G.M., Jannett, T.C., Jadallah, M.A., Yates, S.L., Quint, S.R., Nagle, H.T.: A comparison of the noise sensitivity of nine QRS detection algorithms. IEEE Trans. Biomed. Eng. **37**(1), 85–98 (1990)
44. Liu, F., et al.: Performance analysis of ten common QRS detectors on different ECG application cases. J. Healthc. Eng. **2018**, 8 (2018). Art. no. 9050812
45. Pincus, S.M., Gladstone, I.M., Ehrenkranz, R.A.: A regularity statistic for medical data analysis. J. Clin. Monit. **7**(4), 335–345 (1991)
46. Chen, W., Zhuang, J., Yu, W., Wang, Z.: Measuring complexity using FuzzyEn, ApEn, and SampEn. Med Eng Phys. **31**(1), 61–68 (2009)

Signal Quality Features in Dynamic ECGs

Yixuan Li, Chengyu Liu, Yuwei Zhang, and Jianqing Li

Abstract The electrocardiogram (ECG) recorded by a wearable device usually shows a different signal quality due to the diversity of noise and individual activities. It is often found that the ECG segment only includes noise without any clinically useful information. Therefore, real-time evaluation and feedback of signal quality are required. We propose a real-time dynamic ECG quality evaluation algorithm based on multi-template matching and correlation coefficient matrix. We then illustrate the algorithm and analyze the results using the data from three classification databases of ECG quality. The total Kappa coefficient of this algorithm is 0.705 and the total accuracy is 81.25%.

Keywords Wearable ECG · ECG quality evaluation · Template matching · Correlation coefficient matrix

1 Introduction

Cardiovascular disease (CVD) is a kind of circulatory system disease related to the heart or blood vessels that seriously endangers human health and safety. In many developing countries, the incidence of and mortality due to CVD are increasing. According to the 2017 report on cardiovascular disease in China published in 2018, it is estimated that more than 40% of residents' deaths due to diseases are caused by CVD, which is far higher than other major diseases such as tumors. In recent years,

Y. Li
School of Instrument Science and Engineering, Southeast University, Nanjing, China

University of British Columbia, Vancouver, BC, Canada

C. Liu (✉)
The State Key Laboratory of Bioelectronics, School of Instrument Science and Engineering, Southeast University, Nanjing, China
e-mail: chengyu@seu.edu.cn

Y. Zhang · J. Li
School of Instrument Science and Engineering, Southeast University, Nanjing, China

C. Liu, J. Li (eds.), *Feature Engineering and Computational Intelligence in ECG Monitoring*, https://doi.org/10.1007/978-981-15-3824-7_4

the mortality rate because of CVD in rural areas has been higher than that in urban areas, and the number of CVDs will rise rapidly in the next decade.

While smartphones and mobile networks are common, primary health care is still poorly equipped. Many rural populations around the world rely on clinics staffed by nonspecialist volunteers, and these clinics identify patients who need secondary care through healthcare specialists in distant urban hospitals. Thence, the availability of some inexpensive medical devices, such as electrocardiographs, is becoming increasingly feasible to rural clinics. These electrocardiographs transmit digital ECG to smartphones for storage and display, thereby extending the reach of diagnostic doctors to remote areas. But technology alone cannot provide consistently available information without quality control. The quality of collected data can be improved to assist clinical diagnosis, and growing interest in mobile medicine is driving the vision of combining smartphones with medical data to provide point-to-point diagnostic services for underserved populations [1].

At the same time, the study shows that long-duration ECG can effectively improve the detection rate of arrhythmia. However, the ECG signal is at the millivolt level, which is very weak and sensitive to environmental noise. The quality of the dynamic ECG will be far worse than that of routine ECG due to individual activities, which seriously affects the accuracy of the results. Therefore, it is important to evaluate the quality of dynamic ECG and filter the noise. Also, due to the limitation of power consumption on the mobile terminal and the demand for real-time results, the classification algorithm using many signal features and the complex network is inappropriate, so it becomes extremely urgent to design an algorithm with simple operation and real-time feedback of ECG classification results.

In 2011, the PhysioNet/CinC challenge called for the development of an algorithm that achieves real-time quality assessment of ECG signals and noise filtering.

Clifford et al. [2] quantified the spectral energy distribution, higher-order moments and inter-channel and inter-algorithm agreement through a series of signal quality metrics. Seven indicators (12 leads, a total of 84 features) of each channel were calculated and submitted to a support vector machine (SVM) or a multilayer perceptual neural network (MLP) for classification.

Di Marco et al. [3] used time-frequency analysis to evaluate ECG quality by identifying ECG contaminants (baseline drift, flat line, QRS artifact, stray spikes, gradually changing amplitude, other noises, etc.) on a single lead. Classification was based on cascaded single-condition decision rules (SCDRs) tested levels of contaminants against classification thresholds.

Hayn et al. [4] evaluated four ECG quality indexes: a no signal detector, a spike detector, a lead crossing point analysis, and the quantization of the robustness of QRS detection. Conditions and thresholds were set respectively to realize classification.

Johannesen and Galeotti designed a two-step algorithm [5]. First, a signal with macroscopic errors (signal missing, large voltage offset, and signal saturation) was marked as unacceptable and filtered. Then, the baseline drift, power line noise, and myoelectric noise were quantified. Finally, the global noise was obtained after weighted addition, and the threshold value was set to realize classification.

The objective of these algorithms is to evaluate the quality of clinical available and unavailable ECG signals. Remond et al. [6] separately labeled 300 single-lead

ECG records. By identifying motion artifacts, QRS signal position and quality, and comparing the original unfiltered signal with the filtered signal, the three-level quality indexes (good, middle, and poor) were obtained. Seven groups of features based on template and signal morphology were introduced into the Parzen window supervised statistical classifier model to realize signal three classification.

In this chapter, an ECG quality evaluation algorithm mainly based on the signal quality matrix is proposed. Many researchers used the matrix to represent features of the signal and quantified the signal quality to score or classify it.

Henian Xia et al. [7] proposed a matrix of regularity (MoRE) to measure the degree of irregularities in the ECG. They used some tests such as missing signals, overlaps between leads, and irregular beats to construct the MoRE. MoRE is a $12*12$ matrix, where each column corresponds to a lead. The elements represent the influences of irregularities due to artifacts in the corresponding lead or other leads on the present lead. Therefore, R will be a zero matrix if ECG quality is perfect since no irregularity exists, and R will be 'away' from the zero matrix if the artifact increases. Finally, they computed properties of the MoRE as scores to grade the quality of ECGs.

Yatao Zhang et al. [8] calculated and quantified the power spectrum, baseline drifts, amplitude difference, and other time-domain features to build the feature matrix. The matrix is assessed using KSVM and GA to determine the ECG quality degree.

Yun Chen and Hui Yang [9] utilized Dower transform to derive the 3-lead VCG from the 12-lead ECG to preserve the useful information and reduce the redundant part, thereby being more efficient in the computer processing and analysis of the ECG signal quality. The transformation of an ECG into a VCG was done using a $3*8$ Dower transformation matrix (DTM) and the back operation was done by an $8*3$ inverse DTM.

2 Real-Time Dynamic ECG Quality Evaluation Algorithm Based on Multitemplate Matching and a Correlation Coefficient Matrix

In this chapter, a new real-time dynamic ECG quality evaluation algorithm based on multitemplate matching and a correlation coefficient matrix is proposed. It consists of six parts: (1) ECG waveform pretreatment; (2) QRS detection; (3) standard deviation of the RR interval; (4) ECG signal multitemplate matching and a correlation coefficient matrix; (5) processing of the correlation coefficient matrix; and (6) a fusion evaluation algorithm. Figure 1 shows the specific process of the algorithm.

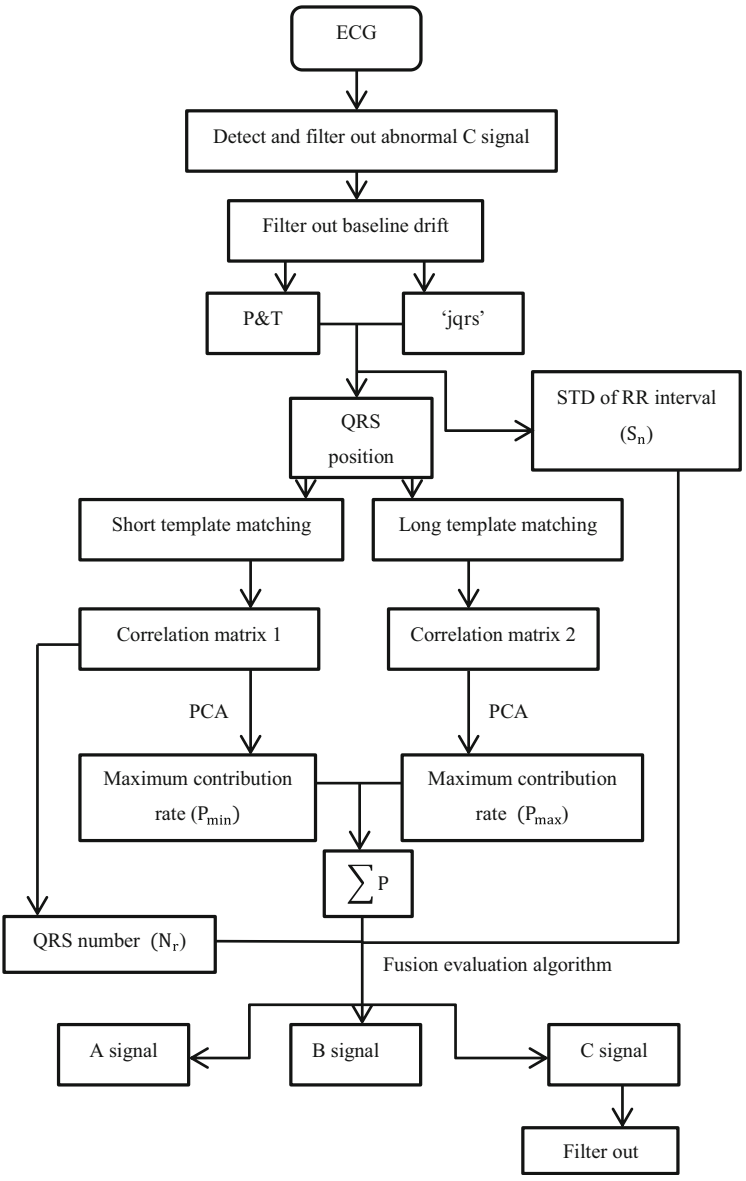

Fig. 1 The specific process of the algorithm

2.1 ECG Waveform Pretreatment

ECG is often contaminated by noise and artifact, which may superimpose on a clean signal or cover the original ECG signal in the frequency band of interest and appear

in morphology like an ECG itself. Therefore, it is necessary to perform noise preprocessing of an ECG signal for subsequent processing.

2.1.1 Abnormal Signal Detection

Abnormal signals, including step signals, linear signals, and saturation signals, are classified as C. Multiple continuous QRS complexes cannot be detected on these signals, which are of low signal quality and clinical uselessness, mostly caused by the following noise or artifact:

Patient-electrode motion artifacts: movement of the electrodes away from the skin contact area, resulting in impedance changes between the skin and the electrodes, leading to potential changes in the ECG. It usually manifests as a rapid but continuous baseline jump or complete saturation for 0.5 s.

Data collecting device noise: artifacts produced by signal processing hardware, such as signal saturation.

According to experimental simulation, if abnormal signals are not detected and filtered first, they will be deformed after baseline drift processing, which may lead to QRS false detection. Therefore, abnormal signals have to be detected and filtered at the beginning.

If there is a straight segment longer than 0.25 s in the 10 s ECG signal, it indicates that the signal has an obvious step, straight line, or saturation segment, which should be classified as a clinically useless C signal.

2.1.2 Baseline Drift Filtering

Baseline drift is one of the main artifacts in ECG recording, which is usually caused by human respiration, body shaking, and temperature drift of the component equipment. It is a low-frequency noise with a frequency of 0.15–0.3 Hz. It presents a sinusoidal curve that changes slowly and overlaps with the S-T frequency band, thus affecting ECG signal detection [10].

The baseline drift can be removed by passing the original signal through a high-pass filter. A Butterworth filter has the characteristics of simple design, comprehensive function, low Q value requirement for components, and easy manufacture. The frequency response curve of the Butterworth filter achieves maximum smoothness and uniformity in the passband and gradually drops to zero in the stopband. The transfer function of the filter is as follows:

$$H(u,v) = \frac{1}{1 + \left(\frac{D_0}{D(u,v)}\right)^{2n}} \tag{1}$$

where n is the order of the filter and $D(u, v)$ is the cut-off frequency. A Butterworth high-pass filter with a cut-off frequency of 0.5 Hz is used to filter baseline drift noise and improve the SNR.

2.2 QRS Detection

The QRS position is detected by the classical algorithm Pan and Tompkins based on difference threshold judgment and the 'jqrs' algorithm based on wavelet transform, and the QRS detected jointly by the two methods is regarded as the final QRS position within the tolerance error range of 60 ms. This way can ensure the accuracy of detection, laying a foundation for template matching of ECG signals and the calculation of other indicators such as the standard deviation of RR interval and QRS number.

2.3 The Standard Deviation of the RR Interval

An RR interval sequence is defined as:

$$\mathrm{RR} = \{\mathrm{RR}_1, \mathrm{RR}_2, \ldots, \mathrm{RR}_n\} \tag{2}$$

The standard deviation (STD) of RR interval is defined as:

$$S_n = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\mathrm{RR}_i - \overline{\mathrm{RR}}\right)^2} \tag{3}$$

where n is the number of the RR interval, i is the serial number of the RR interval, $\overline{\mathrm{RR}}$ is the average of the RR interval. S_n is used to measure the dispersion degree of RR interval sequences. If obvious abnormal values are detected in RR interval sequences, it indicates that normal QRS complexes could not be detected in this segment and there is a large noise.

2.4 ECG Signal Multitemplate Matching

The principle of template matching and correlation coefficient method is to place the template image in the larger image to be searched, and to achieve the research objective by moving the template image and calculating the similarity between the subimage and the template image in each place.

The index measuring the similarity level of two vectors i and j is the correlation coefficient ρ_{ij}. In the broad sense, it can represent the mathematical distance. It has the following properties:

$$| \rho_{ij} | \leq 1 \tag{4}$$

The higher the value of $|\rho_{ij}|$, the higher the correlation between the two vectors, and the lower the value of $|\rho_{ij}|$, the lower the correlation between the two vectors, accordingly. When $|\rho_{ij}| = 0$, the two vectors are not relevant.

The QRS complexes' positions at the first and last ends of the QRS sequence detected in Sect. 2.2 are discarded to ensure that the template length will not exceed the signal boundary. The template center is n QRS complexes after the removal of both ends. By setting two templates with different lengths, the searching of signal images in different scopes is realized. The template lengths of the two groups are the integral value of 0.5 times the average RR interval length (short template) and 1 time the average RR interval length (long template), respectively.

Extract the heartbeats with the above two template lengths, match each heart beat template with all the heart beat templates including itself, and get the correlation coefficient matrix R, which is defined as:

$$R = \begin{bmatrix} 1 & \rho_{12} & \cdots & \rho_{1n} \\ \rho_{21} & 1 & \cdots & \rho_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{n1} & \rho_{n2} & \cdots & 1 \end{bmatrix} \tag{5}$$

The ρ_{ij} $(i,j = 1,2,\ldots n)$ is the correlation coefficient obtained by template matching between T_i, the ith template, and T_j, the jth template.

Correlation coefficient ρ_{ij} is defined as:

$$\rho_{ij} = \frac{\sum_{n=1}^{\text{length}} \left[T_i(n) - \overline{T_i}\right]\left[T_j(n) - \overline{T_j}\right]}{\sqrt{\sum_{n=1}^{\text{length}} \left[T_i(n) - \overline{T_i}\right]^2} \sqrt{\sum_{n=1}^{\text{length}} \left[T_j(n) - \overline{T_j}\right]^2}} \tag{6}$$

where length is the template length, $\overline{T_i}$ and $\overline{T_j}$ are the average of template T_i and template T_j, respectively.

2.5 Correlation Coefficient Matrix Processing

2.5.1 QRS Number

QRS number (N_r) can be derived from the correlation coefficient matrix.

$$N_r = n + 2 \tag{7}$$

If the QRS number (N_r) of ECG signals with a length of 10 s satisfies $N_r < 5$ or $N_r > 40$, the signal is abnormal, which is a noise signal and should be filtered out [5].

2.5.2 PCA Processing

The correlation coefficient matrix contains the correlation degree information between two pairs of each heart beat template centered on QRS complexes. However, in order to evaluate the overall quality of the signal, the correlation coefficient matrix is required to process and reflect the overall correlation of the signal.

Principal Component Analysis (PCA) is a multivariate statistical method proposed by Pearson in 1901 and further developed by Hotelling in 1933. It is a technique of data dimensionality reduction. By referring to linear fitting, high-dimensional data of n-dimension are divided into data projected on k new orthogonal axes, and a large number of related variables are converted into k independent variables (principal components).

The correlation coefficient matrix obtained by template matching is analyzed by PCA, and the eigenvalues and eigenvectors are calculated and arranged in order from large to small, and the maximum contribution rate P is calculated and defined as:

$$P = \frac{\lambda_{\max}}{\sum_{i=1}^{n} \lambda_i} \tag{8}$$

where $\lambda_{\max}$ is the greatest eigenvalue.

The proportion of the variance of a principal component in the total variance, that is, the proportion of the eigenvalue in the total eigenvalue, is defined as the contribution rate. Since the variance of each principal component is decreasing, the information contained gradually decreases, too. Through the observation, the better the signal quality is, the higher the interpretation of the first principal component to the signal data is, and the higher the value of P is. On the other hand, the poorer the signal quality is, the lower the interpretation of the first principal component to the signal data is, the lower the value of P is, and more principal components are needed to explain this ECG information.

Here, two maximum contribution rates, $P_{\min}$ and $P_{\max}$, are used to represent the correlation degree after matching of short and long templates. High value of the P indicates that the ECG signal in this segment is more consistent with the characteristic of approximate periodic fluctuation, and the signal quality is good, which is suitable for clinical use; otherwise, the signal quality is poor and clinically useless.

2.6 Fusion Evaluation Algorithm

By integrating STD of RR interval, QRS number, maximum contribution rate and other indicators, thresholds are set according to the ROC curve and Youden index to classify signal quality levels and filter out type C signals that could not be used clinically.

Vector $X = [\text{T\_line}, S_n, N_r, P_{\min}, P_{\max}]$, where T _ line is the duration of straight segment, S_n is the STD of the RR interval, N_r is the QRS number, $P_{\min}$ is the maximum contribution rate obtained by the short template, and $P_{\max}$ is the maximum contribution rate obtained by the long template.

For the acquired original dynamic ECG signals, set the threshold value of each indicator and divide the signal quality levels. The specific steps are as follows:

(a) If T _ line > 0.25 s, the signal segment is judged to be type C and clinically useless; otherwise, step b is entered.
(b) If the STD of the RR interval satisfies $S_n > 500$ ms, the signal segment is judged to be type C and clinically useless; otherwise, step c is entered.
(c) If the QRS number N_r satisfies $N_r < 5$ or $N_r > 40$, the signal segment is judged to be type C and clinically useless; otherwise, step d is entered.
(d) The maximum contribution rates $P_{\min}$ and $P_{\max}$ are obtained by template matching and correlation coefficient matrix processing. If the sum of $P_{\min}$ and $P_{\max}$ satisfies $\sum P < 55$, the signal segment is judged to be type C and clinically useless; If $\sum P > 125$, the signal segment is judged to be type A and of good quality; If $55 \leq \sum P \leq 125$, the signal segment is judged to be type B, clinical available but the signal quality is relatively poor;
(e) Finally, type C ECG signals are filtered out to realize real-time quality evaluation of dynamic ECG.

3 Demonstration of ECG Three-Classification Quality Evaluation Algorithm

3.1 ECG Waveform Pretreatment

3.1.1 Abnormal Signal Detection

Figure 2 shows an abnormal signal, which includes step signals, linear signals, and saturation signals. If this signal is not filtered out first, QRS complexes will be detected falsely after baseline drift processing.

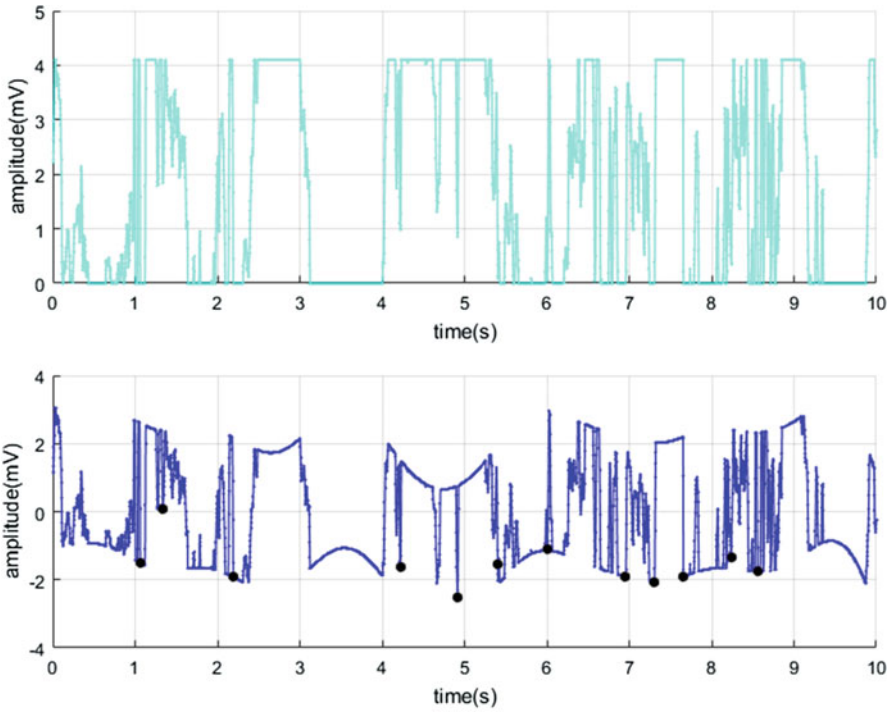

Fig. 2 Abnormal ECG signal

3.1.2 Baseline Drift Filtering

Figure 3 demonstrates the comparison of the effects after baseline drift filtering processing. The baseline signal becomes stable after filtering.

3.2 The Standard Deviation of the RR Interval

After experimental induction and analysis, it is found that if $S_n > 500$ ms, it indicates that the discretization degree of the RR interval sequence is serious, and many abnormal values are included. If the value of S_n is larger than 500 ms, the signal is determined as type C. As shown in Fig. 4, the STD of the signal $S_n = 702$ ms; this is a noise signal.

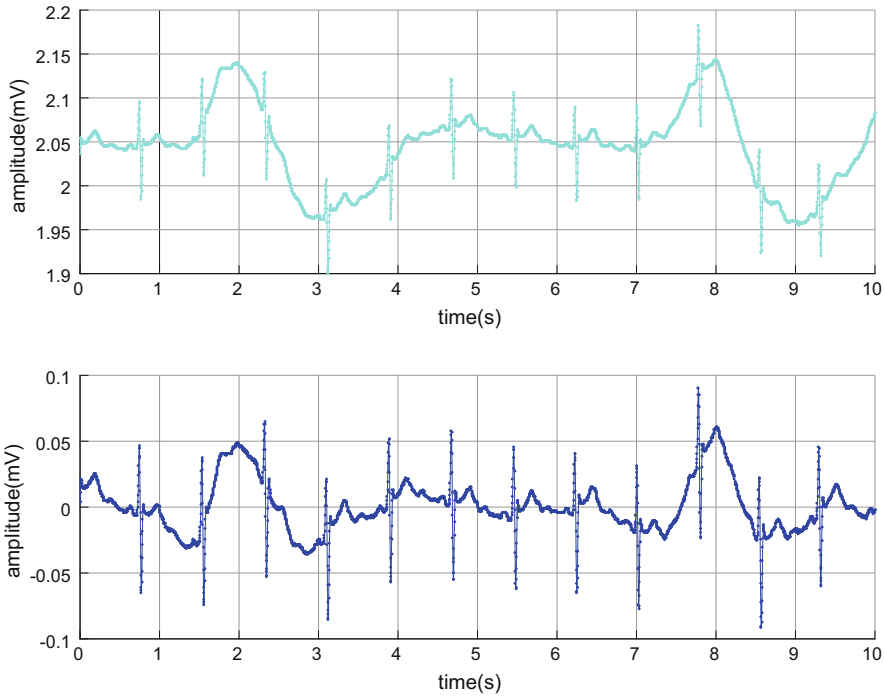

Fig. 3 Baseline drift filtering

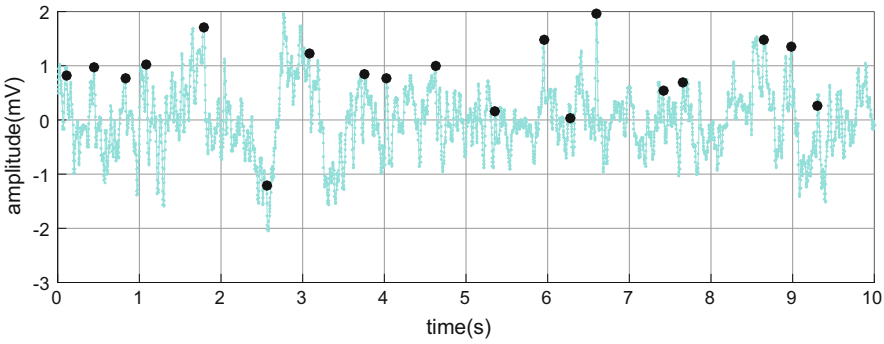

Fig. 4 A type C signal with $S_n = 702\,\text{ms}$

3.3 ECG Signal Multitemplate Matching

The deep blue curves displayed in Fig. 5 are the short template and the long template, respectively.

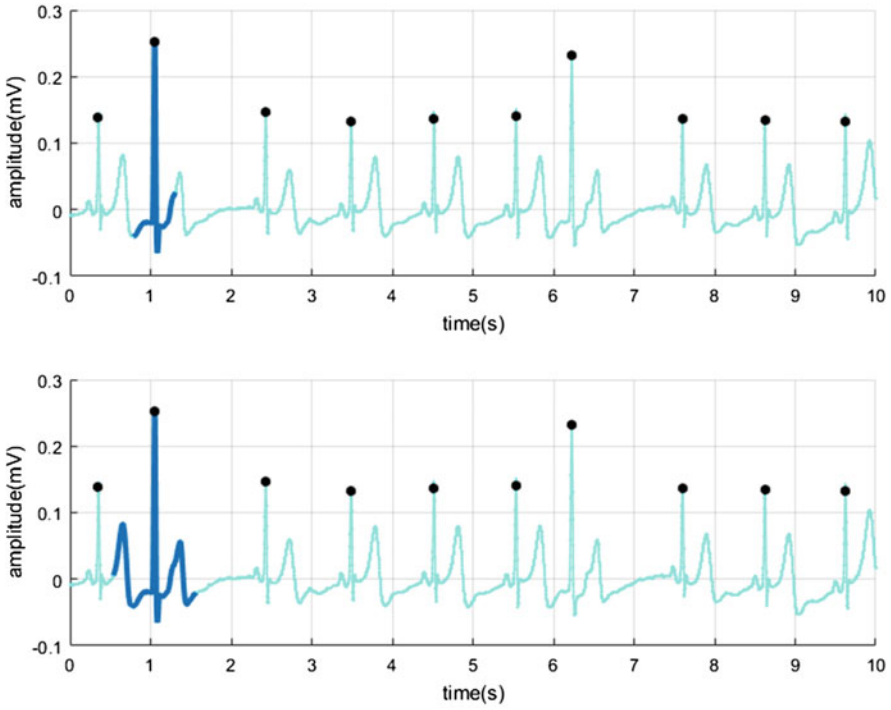

Fig. 5 Short template and long template

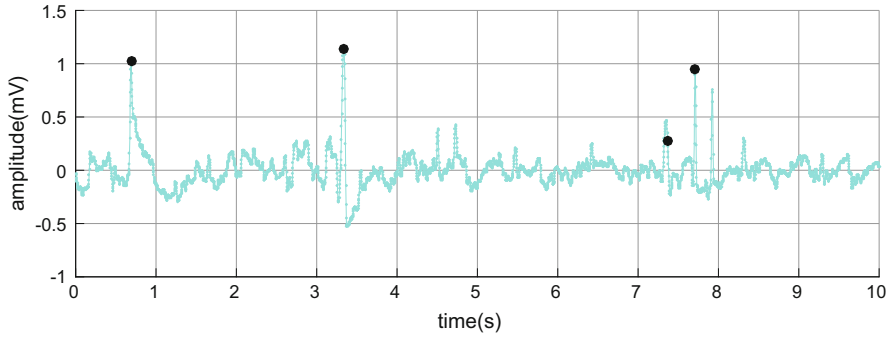

Fig. 6 A type C signal with $N_r = 4$

3.4 QRS Number Obtained from the Matrix

If the number of QRS complexes is less than 5, this signal is classified as type C. Figure 6 illustrates that three QRS complex waves are detected, and the number of waves is not within the range of the QRS waves of the normal ECG signal.

Therefore, the signal should be classified into class C and filtered according to the rules.

3.5 The Signal Examples Mainly Classified by the Matrix

Type A ECG signals satisfy $\sum P > 125$, $\sum P$ is the sum of $P_{\min}$ and $P_{\max}$. In Fig. 7, $\sum P$ equals 131.2, 137.7, and 153.2, respectively.

Type B ECG signals satisfy $55 \leq \sum P \leq 125$, $\sum P$ is the sum of $P_{\min}$ and $P_{\max}$. In Fig. 8, $\sum P$ equals 71.4, 62.3, and 56.9, respectively.

Type C ECG signals satisfy $\sum P < 55$, $\sum P$ is the sum of $P_{\min}$ and $P_{\max}$. In Fig. 9, $\sum P$ equals 41.4, 42.6, and 54.1, respectively.

3.6 Classification Result Analysis

In this work, three evaluation indicators named F-measure, Kappa coefficient, and accuracy (ACC) are employed to evaluate the reliability of the training model. The

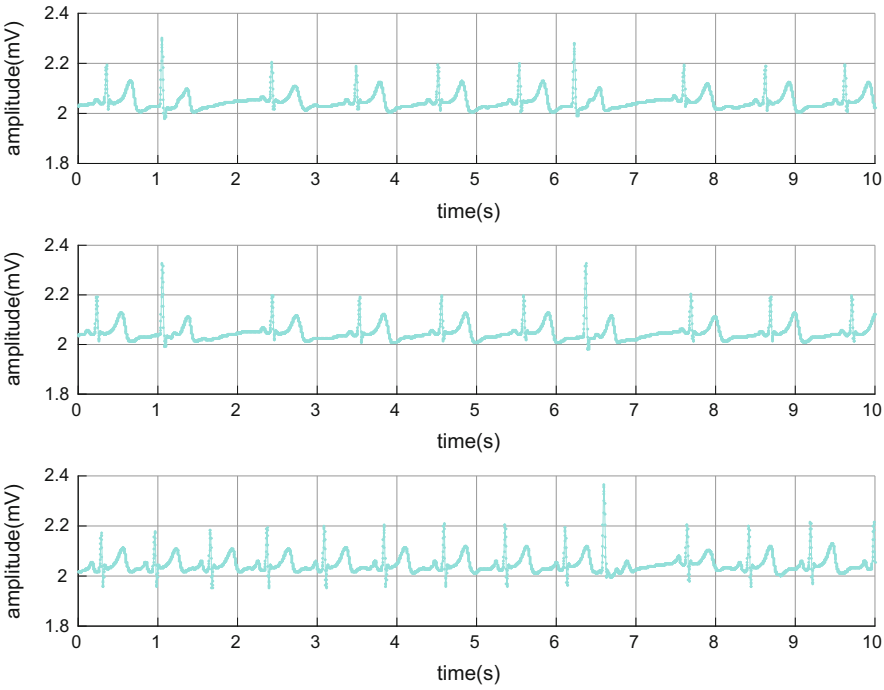

Fig. 7 ECG signals of type A

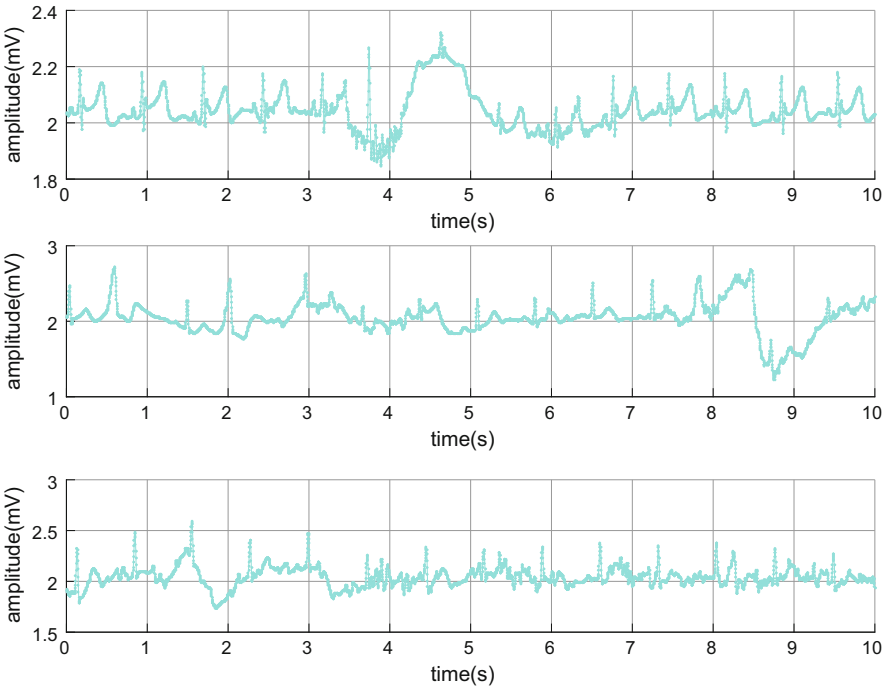

Fig. 8 ECG signals of type B

definitions of recall rate (RE_k), precision rate (PE_k), and accuracy rate (ACC) are below:

$$\mathrm{RE}_k = \frac{C1_k}{C1_k + C2_k} \tag{9}$$

$$\mathrm{PE}_k = \frac{C1_k}{C1_k + C3_k} \tag{10}$$

$$\mathrm{ACC} = \frac{\sum_{k=1}^{3} C1_k}{\sum_{k=1}^{3} C1_k + \sum_{k=1}^{3} C2_k} = \frac{\sum_{k=1}^{3} C1_k}{\sum_{k=1}^{3} C1_k + \sum_{k=1}^{3} C3_k} = \frac{2\sum_{k=1}^{3} C1_k}{2\sum_{k=1}^{3} C1_k + \sum_{k=1}^{3} C2_k + \sum_{k=1}^{3} C3_k} \tag{11}$$

where $C1_k$ represents the number of type k ECG signals that are identified as type k correctly. $C2_k$ represents the number of type k ECG signals falsely identified as other two types; $C3_k$ represents the number of other two types of ECG signals falsely identified as type k signals.

Conflict exists in the index of precision and recall sometimes, so they need to be considered comprehensively. F-measure is a weighted and average of precision rate

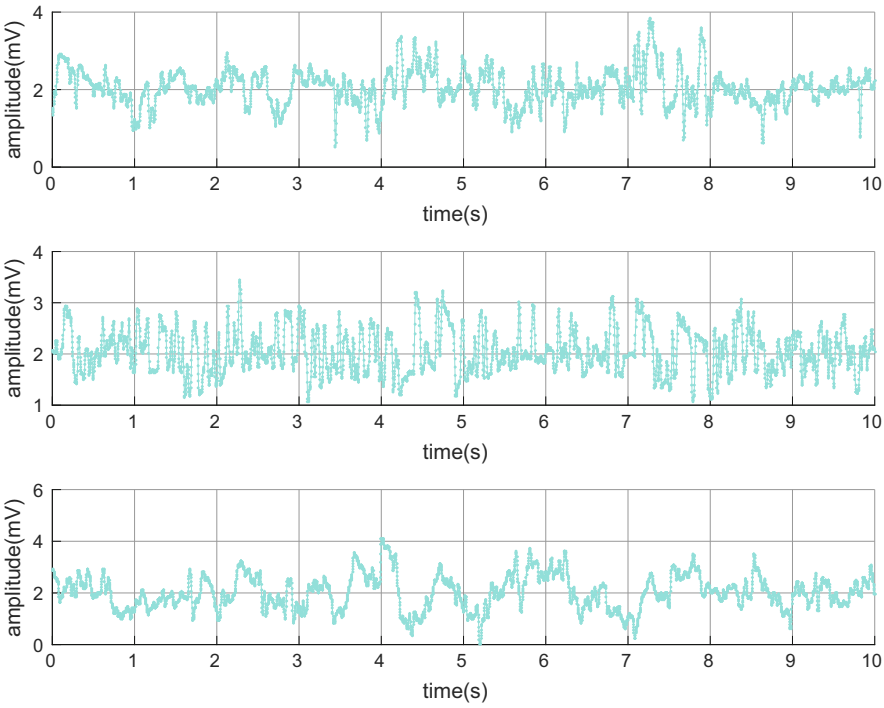

Fig. 9 ECG signals of type C

and recall rate, which can comprehensively reflect the classification of the model. The index of F-measure is defined as:

$$F_\beta = \frac{(\beta^2 + 1)\mathrm{PE}_k \times \mathrm{RE}_k}{\beta^2 \mathrm{PE}_k + \mathrm{RE}_k} \tag{12}$$

The value of β is assigned 1 under normal circumstances. The definition of F_1 is as follows:

$$F_1 = \frac{2 \times \mathrm{PE}_k \times \mathrm{RE}_k}{\mathrm{PE}_k + \mathrm{RE}_k} \tag{13}$$

For a classification algorithm, the higher the $F1$ measure, the more effective the classification method is.

Kappa coefficient is mostly used for testing consistency and measure classification accuracy, which can be obtained through the confusion matrix. The value of Kappa coefficient ranges from 0 to 1. When Kappa value is in the interval from 0.61 to 0.80, it indicates a high degree of consistency between the classification results and the actual results. When Kappa value is in the interval from 0.81 to 1, it means

Fig. 10 Confusion matrix

Proposed / Referenced	A	B	C
A	40	20	0
B	10	103	7
C	8	15	117

Table 1 Measurement performance of this chapter

*F*1-measure			Kappa	ACC
A	B	C		
0.678	0.798	0.886	0.705	0.8125

that the classification results are almost identical with the actual results. Kappa coefficient is defined below as:

$$\text{Kappa} = \frac{p_0 - p_e}{1 - p_e} \tag{14}$$

where the p_0 value equals the accuracy rate. Suppose that the number of actual samples of each type is a_1, a_2, a_3, and the number of predicted samples of each type is b_1, b_2, b_3, and the total number of samples is n. p_e is defined as:

$$p_e = \frac{a_1 \times b_1 + a_2 \times b_2 + a_3 \times b_3}{n^2} \tag{15}$$

The confusion matrix is derived from 320 pieces of data (type A: 60, type B: 120, type C: 140) in the three-classification database, as shown in Fig. 10 below.

Table 1 lists the measurement performance of the dynamic ECG quality evaluation algorithm proposed in detail.

Then, using the test set (type A: 10, type B: 10, type C: 10) for evaluation, total accuracy is 76.7%.

The result analysis above indicates that the algorithm can effectively realize the quality evaluation of dynamic ECG signals.

4 Summary

This chapter mainly introduces the principle and specific steps of a real-time ECG quality evaluation algorithm based on multitemplate matching and correlation coefficient matrix. Computation of three classification indexes of *F*-measure, Kappa coefficient, and accuracy of the three-classification database proves that the algorithm is not only simple to calculate, but also has high accuracy and a good classification effect. It can be effectively applied to the dynamic ECG quality assessment in real time.

References

1. Clifford, G.D., Moody, G.B.: Signal quality in cardiorespiratory monitoring. Physiol. Meas. **33** (9), 1414–1418 (2012)
2. Clifford, G.D., Behar, J., Li, Q., et al.: Signal quality indices and data fusion for determining clinical acceptability of electrocardiograms. Physiol. Meas. **33**(9), 1419–1433 (2012)
3. Di Marco, L.Y., Duan, W., Bojarnejad, M., et al.: Evaluation of an algorithm based on single-condition decision rules for binary classification of 12-lead ambulatory ECG recording quality. Physiol. Meas. **33**(9), 1435–1448 (2012)
4. Hayn, D., Jammerbund, B., Schreier, G.: QRS detection based ECG quality assessment. Physiol. Meas. **33**(9), 1449–1461 (2012)
5. Johannesen, L., Galeotti, L.: Automatic ECG quality scoring methodology: mimicking human annotators. Physiol. Meas. **33**(9), 1479–1489 (2012)
6. Redmond, S.J., Xie, Y., Chang, D., et al.: Electrocardiogram signal quality measures for unsupervised telehealth environments. Physiol. Meas. **33**(9), 1517–1533 (2012)
7. Xia, H., Garcia, G.A., Bains, J., et al.: Matrix of regularity for improving the quality of ECGs. Physiol. Meas. **33**(9), 1535–1548 (2012)
8. Zhang, Y., Liu, C., Wei, S.: ECG quality assessment based on a kernel support vector machine and genetic algorithm with a feature matrix. J. Zhejiang Univ. Sci. C. **15**(7), 564–573 (2014)
9. Chen, Y., Yang, H.: Self-organized neural network for the quality control of 12-lead ECG signals. Physiol. Meas. **33**(9), 1399–1418 (2012)
10. Cortez, D., Sharma, N., Cavanaugh, J., et al.: The spatial QRS-T angle outperforms the Italian and Seattle ECG-based criteria for detection of hypertrophic cardiomyopathy in pediatric patients. J. Electrocardiol. **48**(5), 826–833 (2015)

Motion Artefact Suppression Method for Wearable ECGs

Huanqian Zhang and Jianlong Zhao

Abstract Due to the development of the Internet of Things, there has been increasing interest in the use of wearable electrocardiograms (WECGs) in the outdoor environment instead of in a resting state at a hospital. During daily activities, the WECG signals will suffer additional motion artefacts (MAs) originating from the interface between the electrode and the conductive adhesive and the stretching of the skin. However, MAs in WECG signals are highly difficult to suppress because MAs and WECG signals have similar frequency spectra.

In this review, we briefly discuss motion artefact suppression methods, from the origin of the motion artefacts to detecting the motion artefacts and then suppressing the motion artefacts.

The metabolic difference between the live skin cells of the inner layer and the dead skin cells of the stratum corneum create an 'injury current'. When a force is applied to the skin, the membrane of the dead skin cell breaks, and then, sodium will flow into the cells through the crack and ultimately form the 'injury current'. When the current flows through the resistor of the stratum corneum, there will be a potential change, which is the MA.

Adaptive filters (AFs) have been extensively applied in biomedical engineering because of their simplicity, real-time processing ability and robustness. These filters can remove MAs from WECG signals by using a reference signal that is correlated with MAs and uncorrelated with the WECG. We also describe two concepts of reference signal detection.

Because of the nonstationary properties of motion artefacts, low filter output distortion and high QRS beat detection accuracy cannot be simultaneously

H. Zhang (✉)
Chinese Academy of Science, Shanghai Institute of Microsystem and Information Technology, Shanghai, China
e-mail: henry.zhang@nus.edu.sg

J. Zhao (✉)
Chinese Academy of Science, Shanghai Institute of Microsystem and Information Technology, Shanghai, China
e-mail: jlzhao@mail.sim.ac.cn

C. Liu, J. Li (eds.), *Feature Engineering and Computational Intelligence in ECG Monitoring*, https://doi.org/10.1007/978-981-15-3824-7_5

generated. Hence, we propose a new feed forward combined adaptive filter algorithm to overcome this limitation.

Finally, we provide an overview of recent findings for the adaptive filter algorithm-based motion artefact suppression method.

Keywords Wearable ECG · Motion artefacts · Electrode tissue impedance · Noise suppression · Adaptive filter

1 Introduction

With the increase in human life expectancy, the trend of ageing is evident. In 2050, the population of elderly people (greater than 60 years old) will exceed that of young people (less than 15 years old). There will be a large expense in medical care because of ageing, and the health of elderly people will determine the overall cost of medical insurance and the frequency of using advanced medical equipment. Therefore, people must remain healthy as the life expectancy increases. Wearable medical services will be a major part of helping elderly people decrease the frequency of hospital medical treatments and increase their life expectancy.

Among the various medical conditions, heart disease has the highest mortality rate. ECGs are widely used to detect heart disease. Wearable ECGs are a new technology that extends ECG detection from hospitals to daily life. However, patients move about during the course of their daily life, and the weak ECG signal will be affected by motion artefacts, leading to an incorrect estimation of ECG features and triggering unnecessary warnings. Recently, many studies have been published to find the origins of motion artefacts and methods of suppressing motion artefacts.

In this chapter, we will review these studies. First, we will discuss the source of motion artefacts from the view of anatomical and circuit models. Second, we will describe the detection of motion artefacts from the electrical hardware system domain. Finally, we will introduce a method for suppressing motion artefacts using an adaptive filter.

2 The Origin of Motion Artefacts

Tam and Webster [1] found that the amplitude of the deformation potential decreases when the stratum corneum is scraped away. They concluded that the major source of motion artefacts is the skin/paste interface. The magnitude of the change in skin potential will be significantly affected by the degree of skin abrasion.

The skin is made up of three layers: the epidermis, the dermis and the subcutaneous layer. The stratum corneum is the surface layer, which is composed of dead cells. The stratum granulosum and the stratum basale are located below the surface layer, forming the layers of the epidermis. The dermis is located underneath the stratum basale. Connective tissue, elastic tissue and living cells make up the rest of the dermis.

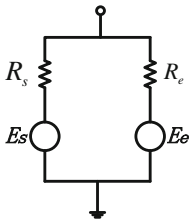

Fig. 1 Edelberg's skin model [1]

Many skin models have been suggested in the literature, but Edelberg's skin model is the most widely accepted.

In Fig. 1, E_e is the potential across the epidermis barrier membrane, wherein the magnitude of E_e is a function of the composition of the electrode paste. R_e is the total series resistance of the epidermis. Because $|E_e| < |E_s|$, skin abrasion decreases R_e such that $R_e \ll R_s$, thereby resulting in a less negative skin potential. Additionally, the smaller variation in R_e caused by skin deformation leads to a much smaller variation in skin potential.

E_s is the potential across the sweat duct membrane at the layer of the stratum basale. E_s is variable because of the diverse salt concentration in sweat. R_s is the total resistance in the sweat duct, for which the magnitude is determined by the height of the column of saline in the sweat duct and the permeability of the sweat duct wall. Sweat gland activity in response to sympathetic activation increases the sweat column, decreasing R_s. Hydration of the stratum corneum also has an effect on reducing R_e. The long settling time of the offset potential after applying the electrodes is the result of the wetting of the stratum corneum over time, which is caused by sweat and the interaction between the sweat and the paste. E_{skin} can be expressed as shown in Eq. (1):

$$E_{skin} = E_e + R_e\left(\frac{E_s - E_e}{R_s + R_e}\right) \tag{1}$$

A variation in the skin potential results from changes in any of the four parameters. Thus, the net variation in the offset potential of the recording system is the sum of the individual variations in skin potential under each electrode. Diagnostic ECG measurements usually use a lower cut-off frequency amplifier of 0.05 Hz. Very slow variations in the offset potential cause a negligible drift in the baseline. In contrast, rapid variations in the offset potential $(d_{E_{offset}}/d_{time})$ lead to obvious motion artefact problems. For example, when $d_{E_{offset}}/d_{time} = 2\,\text{mV}/\,\text{min}$, the baseline varies by 0.1 mV. The lower cut-off frequency is 0.5 Hz or higher in the monitoring mode, so that even greater variations in the offset potentials are tolerable.

The metabolic process of different skin layers causes ion diffusion [2].

Thakor and Webster [3] hypothesized that the metabolic process is the result of the differences in metabolic activity between the dead cells located in the outer layer of the skin and the cells located in the inner layer.

When mechanically stretching the skin, the skin potential V increases by several millivolts. Thakor and Webster explained the origin of this motion artefact ΔV as a reduction in the extracellular channel resistance Z.

The foundation of this model is based on two hypotheses. First, they hypothesized that the skin potential arises from a constant current source called 'injury current', which is generated by the difference in metabolic activity between the dead cells of the stratum corneum and the viable cells of the inner layers of the skin. Second, they hypothesized that this injury current flows through the extracellular channels, generating a negative DC voltage that drops from the inside to the outside of the skin. The skin potential is shown in Eq. (2):

$$V = (-R_{\mathrm{I}})(R_{\mathrm{m}})(I)/(R_{\mathrm{t}} + R_{\mathrm{C}} + R_{\mathrm{m}}) \tag{2}$$

where R_{C} is the impedance of the stratum corneum, R_{t} is the impedance of the transitional region shunted by the current I of negatively charged ions and $R_{\mathrm{m}}$is the impedance of the measuring device. With $R_{\mathrm{m}} \gg (R_{\mathrm{c}} + R_{\mathrm{t}})$, the following expression can be obtained:

$$V = (-R_{\mathrm{t}})(I) \tag{3}$$

Because of the first hypothesis (I is constant), the only way to obtain variations in V when stretching the skin is to assume that R_{t}alters $V : \Delta V = (-\Delta R_{\mathrm{t}})(I)$

In the report by Talhouet and Webster, their model is not perfect because of the values of ΔZ that can be positive or negative at high values of Z and the difference in shape between ΔV and ΔZ.

They explained the increase in Z in their study. They assumed that R_{c} increased more than R_{t} decreased when stretching the skin. The different behaviour between R_{c} and R_{t} could be clarified by the geometrical configuration of the skin cells. The decrease in R_{t} can be explained by the geometrical arrangement of the stratum basale and stratum granulosum, as stretching the skin causes the extracellular channels to increase in diameter. The stratum corneum cells fit together when they are displaced horizontally so that the horizontal channels between the cells form the main resistance pathway for ions.

When stretching the skin, the length of the current pathway increases, and the cross-sectional area of the current pathway decreases. This phenomenon occurs because the cells that are linked together by tight junctions or gap junctions can elongate under stretching.

They also suppose that the diffusion of Na+ ions across the proximal side of the membrane is increased by stretching a cell in the transitional layer. This action will cause the interior potential of the cell to be more positive. Viscoelastic stretching and relaxation of the cell membrane could occur with long time constants and cause variations in the skin potential with long time constants.

Burbank and Webster [4] studied the artefact potential amplitude and strain dependence as a function of the stretching force and time.

The skin under an electrode site was stretched for a time t with a repetition period r. A relationship was then defined between the resulting artefact potential amplitude A, which was defined as the maximum change in skin potential during skin stretching in relation to that just before skin stretching, and the relaxation time $(r - t)$ before stretching. The fitted equation appeared to be quite good, as shown in Eq. (4):

$$A = A_0\left(1 - e^{-(r-t)/\tau}\right) \tag{4}$$

where A_0 is the maximum artefact and τ is the time constant of the system. The time constant is 26 s, which is much longer than the electrical time constant of the skin at low frequencies, which is normally approximately 0.1 s.

By changing the stretching force and simultaneously monitoring the skin strain, potential and impedance, they compared the relationship between the strain and stretching force and the artefact potential. They increased the stretched mass from 0 to 1 kg and then decreased it again to 0 kg at a uniform rate of 30 g/s. The impedance was very nearly stable during this cycle. Although the strain was a nonlinear function of the stretched mass, it had a small time dependence or 'creep'. However, as discussed above, the artefact potential showed a very obvious time dependence.

3 The Detection of Motion Artefacts

Hamilton et al. [5] reported a system for evaluating and comparing motion artefact removal with sensors and impedance. A sinusoidal current was applied to an active electrode pair. They used the series resistances on the secondary side of the transformer to limit the current between the electrodes to 1 μA with a 1 V peak-to-peak output. Low-pass and high-pass filters were used to separate the impedance signal from the ECG and artefact signals. The low-pass filter and high-pass Butterworth filter had cut-off frequencies of 50 and 100 Hz, respectively. They applied an envelope detector to monitor the amplitude of the impedance signal after separating the motion artefact signal from the impedance signal. The final envelope signal was scaled to the impedance between the two electrodes. A bandpass filter with lower cut-off frequencies of 0.16 Hz and upper cut-off frequencies of 106 Hz removed the DC level of the impedance from the envelope signal. They converted the motion artefact signal to a digital signal with 10-bit resolution and a 120 Hz sample rate.

Spinelli et al. reported a simple direct method to measure the unbalance at power line frequency [6]. The external resistors R_c ensure a well-known common-mode input impedance. To apply the method to three electrode amplifiers, the third (right leg) electrode must be disconnected.

They measured actual skin-electrode impedances. Two plate electrodes (12 cm^2 in area) were placed on the right and left inner arms of the patient (ECG lead I). In the next experiment, they measured the imbalance in the electrode impedance between a

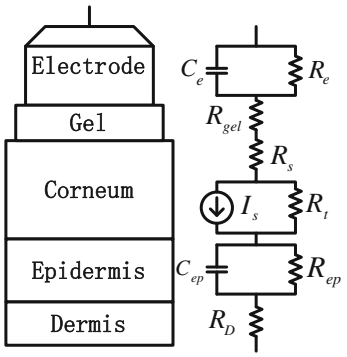

Fig. 2 Circuit representation of the skin-electrode interface and changes in electrical properties under motion artefacts [7]

plate electrode (right arm) and a cup electrode (left arm), which showed a significantly larger discrepancy than those obtained for two similar electrodes placed on the same location. The main interference voltage is expressed as shown in Eq. (5):

$$V_{\mathrm{D.EMI}} = V_{\mathrm{CM}} \frac{\Delta Z_{\mathrm{E}}}{Z_{\mathrm{C}}} \tag{5}$$

where ΔZ_{E} is the electrode-skin imbalance, Z_{C} is the average common-mode input impedance and V_{CM} is the patient common-mode voltage.

Romero et al. [7] reported an application-specific integrated circuit (ASIC) for monitoring three-lead ECG signals and one-channel skin-electrode impedance or electrode-tissue impedance (ETI). To calculate the ETI, they injected an AC signal and measured the voltage induced by the ETI. To avoid any interference with the ECG signal, the frequency of the AC signal needs to be outside the ECG range. They used a square-wave current at 2 kHz with a known amplitude. With the model shown in Fig. 2, concurrent measurement of the resistive and capacitive components was required for accurate measurement of the ETI information.

By demodulating the impedance signal with an in-phase frequency $f(0)$ and a quadrature-phase frequency $f(90)$, they separated the resistive and capacitive components.

The monitoring system synchronously measured the electrode-skin impedance and the ECG signal. The monitoring system comprised an instrumentation amplifier (IA), a ripple filter, a programmable gain amplifier and a bandwidth controllable low-pass filter. The injected AC signal modulates and measures the ETI. An instrumentation amplifier (IA) amplifies the resulting voltage and demodulates the voltage with in-phase and quadrature-phase chopper clocks. The output signal of the IA (in-phase, IMPI; quadrature-phase, IMPQ) will be filtered by a low-pass filter and amplified by a programmable gain amplifier (PGA) that has four different gains. The output signals are a measurement of the complex ETI.

The design of the current stimulation block should pay attention to the following two points. The first is the output impedance of the current generator, which may reduce the total input impedance of the ECG readout channel. The second is the DC

component of the AC source, which may further amplify the effect of motion artefacts at the output of the ECG channel. Therefore, they use a chopper-stabilized AC source. To set the mean value of the stimulation current to zero, they double the frequency of the AC signal to perform chopper stabilization. The continuous time impedance monitoring channels demodulate the resulting AC voltage over the electrode-skin interface. The ECG and impedance signals can be separated in the frequency domain with a low-pass filter.

Ottenbacher et al. [8] reported a method for detecting motion artefacts by the simultaneous measurement of electrode-skin impedance with an ECG signal. A sinusoidal current of 400 Hz was injected at the same electrodes between which the potential was measured.

High- and low-pass filters separated the impedance and potential signal, respectively. A dual lock-in amplifier reconstructed the impedance signal. They used a very small current of <1 μA to measure over a very wide range of electrode impedances (500 Ω to 1 MΩ) and avoid high filter orders to separate the ECG and impedance signal. They performed the experiments on the forearm of a test subject. They used tape to attach a fixed reference electrode with gel near the subject's elbow. In one experiment, they pressed (F_p) the dry measurement electrode against the subject's arm. In a second experiment, they stretched (F_s) the skin under the dry electrode while they pressed the electrode against the subject's arm with a small weight.

The results in their study showed that the force, potential and electrode-skin impedance depended on time (units: s). The potential is on the order of several mV, and the impedance decreases correspondingly when pressing the electrode against the subject's skin. When stretching the skin under the electrode, the potential exhibits an increase in the mV range, and the impedance also increases. The variation in the impedance is diverse from subject to subject and greatly relies on the humidity/sweat on the subject's skin. They assumed that the varying area of the electrode-skin contact, instead of the impedance change in the skin, leads to impedance changes. They also found a very good correlation between force, electrode impedance and electrode-skin potential. Nevertheless, it must be noted that actual movements are changing constantly with simultaneous stretching and pressing.

A measurement is made with two dry electrodes on the chest. The correlation between the impedance signal and the ECG signal against time is depicted. The amplitude of the peaks has less information because it depends on how large artefacts contaminate the signal. However, the peak reveals good correlation between the two signals.

Oberg [9] reported a method to monitor the skin-electrode contact. They added an AC voltage source with the potential U6 between the noninverting input of IC4 and the ground.

The ECG signal U_3 is independent of U_5 because of Eq. (6):

$$U_3 = (U_2 - U_1)\left(2 \cdot \frac{R_1}{R_2} + 1\right) \tag{6}$$

The feedback loop from the output to the inverting input passes all contact points A, B and C. Hereafter, assume that the impedances Z_a, Z_b and Z_c are no longer negligible and that the impedances Z_a and Z_b are equal. Z_t is the sum of the impedances between the ground electrode and one of the inputs, e.g. $Z_t = Z_a + Z_c$. A potential divider is formed by the impedance Z_t together with the input resistance R_{in} of the amplifier. In this case, the expression for U_5 will be Eq. (7):

$$U_5 = U_6\left|1 + \frac{2Z_t}{R_{\text{in}}}\right| \tag{7}$$

This means that U_5 is dependent on Z_t. If the ground electrode is completely loosened, i.e. $Z_t = \infty$, then U_5 goes to infinity. If the impedance $Z_a \neq Z_b$ by an amount Z_{diff}, then U_5 can be expressed as shown in Eq. (8):

$$U_5 = 2U_6\left|1 - \frac{1}{\frac{Z}{R_{\text{in}}} + 2}\right| \tag{8}$$

U_5 relies on the contact impedance. If one of the electrodes at points A or B becomes loose, Eq. (9) is obtained:

$$U_5 = 2U_6 \tag{9}$$

Hence, if we introduce a voltage source at the noninverting input of IC4, then the output voltage depends on the contact impedances Z_a, Z_b and Z_c. If we compare U_5 and U_6, we can decide whether the impedance in some of the contact points is too large.

Degen and Jackel [10] reported a new method that allows continuous monitoring of electrode-skin impedance. Each channel is preceded by a protection circuit, which limits the maximal current through the body to 50 μA. They applied the method to a three-electrode ECG without the additional reference electrode. The operations of the measurement circuit are explained hereafter. The driven right leg (DRL) loop will be forced by any voltage appearing at the positive input of the DRL op-amp. The system reacts in such a way that the differential input voltage of the DRL op-amp is again zero. In the case of a sinusoidal signal, this phenomenon occurs when the bandwidth of the DRL loop is larger than the signal frequency. Therefore, at the input of the instrumentation amplifier (INA), the sinusoidal voltage V_{add} appears as a common-mode voltage. The Common Mode Rejection Ratio (CMRR) of the INA rejects this additional common-mode voltage, except for the part converted to a differential signal by the electrode-skin impedance mismatch (potential divider effect) and amplified by the differential mode gain of each INA. This part is superimposed on the corresponding output voltage V_{outi}. If we exclude all other

signal sources, e.g. power line interference and bioelectric signals, the residual of the additional common-mode signal V_{addi} at electrode i can be calculated.

The input impedance mismatch is expressed in Eq. (10):

$$\Delta Z_{\text{el}} \Delta j\omega \overline{C_{\text{in}}} G_{\text{DM}} \approx \frac{V_{\text{addi}}}{V_{\text{add}}} \tag{10}$$

They applied a sinusoidal force to the signal electrode '*I*'. Moreover, they measured both the norm of the residual voltage |*V* _ addi| and the bioelectric recording. A strong relationship was found between the measured impedance mismatch and the baseline variation.

Bertrand et al. [11] reported that the prediction of motion artefacts at one electrode can be further improved by incorporating impedance measurements at other electrodes in EEG recording.

Comert and Hyttinen [12] reported a simultaneous measurement of impedance at eight current frequencies during the application of controlled motion to the electrode under the mounting force of the monitored electrode. They found that the motion is not reflected by the different frequencies of impedance measurements. The best correlation between impedance and the applied motion appeared when the impedance current frequencies were greater than 11 kHz. The impedance signal correlated well with the applied motion; however, impedance had a lower correlation to the actual motion artefact signal.

Zhang et al. [13] reported an approach that injects an additional common-mode signal through the reference electrode to simultaneously measure the electrode-tissue impedance and ECG signal. To suppress the MA in a WECG, a reference signal that has a high correlation with the MA and a low correlation with the WECG is required by an adaptive filter (AF). Figure 3 shows that the reference signal for the AF can be generated by the multichannel electrode-tissue impedance (MC-ETI) detection approach without any additional sensors.

A 1 kHz AC voltage is forced by amplifier A2 through the driven right leg circuit and electrode Z_{LA}. Two current paths flow through the body. One flows through Z_{LA}, Z_{RA}, and Z_{in} to the ground, and the other flows through Z_{LA}, Z_{LL}, and Z_{in} to the ground. The input impedance of the instrument amplifier is Z_{in}. LL is located several millimetres below the left breast, RA is located several millimetres below the right collarbone and LA is located several millimetres below the left collarbone.

When the electrode movement leads to the variation in Z_{LA}, Z_{LL}, and Z_{RA}, the divided voltages v_{LL} and v_{RA} will vary simultaneously. These voltages are differentially amplified by A3 and A4. A3 and A4 generate two AC voltages $v_{\text{ETI\_LL}}$ and $v_{\text{ETI\_RA}}$. Moreover, the WECG signal v_{ECG} is detected by A1. These voltages are sampled by an analogue-to-digital converter (ADC) with an 80 kHz sampling rate. The digital data are transported to the PC. The DC component from $v_{\text{ETI\_LL}}$ and $v_{\text{ETI\_RA}}$ is extracted by a digital lock-in amplifier. The MC-ETI signal can be calculated by Eqs. (11) and (12):

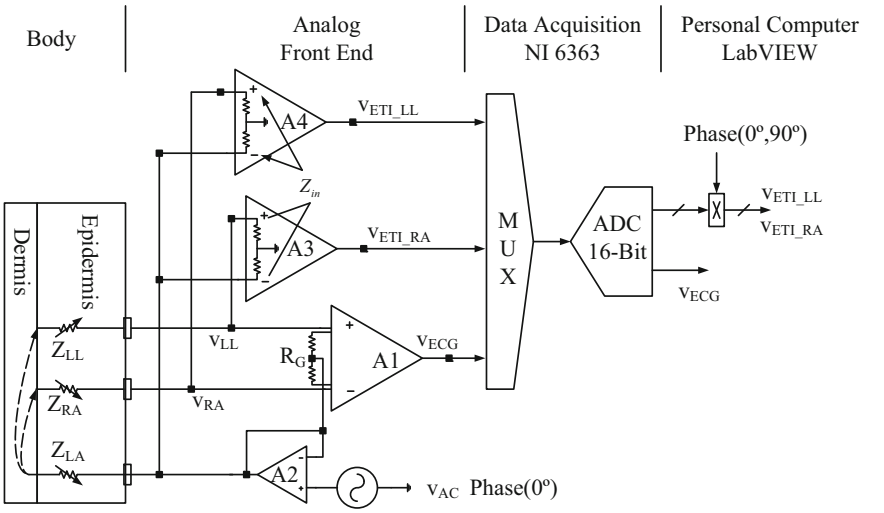

Fig. 3 MC-ETI detection approach [13]

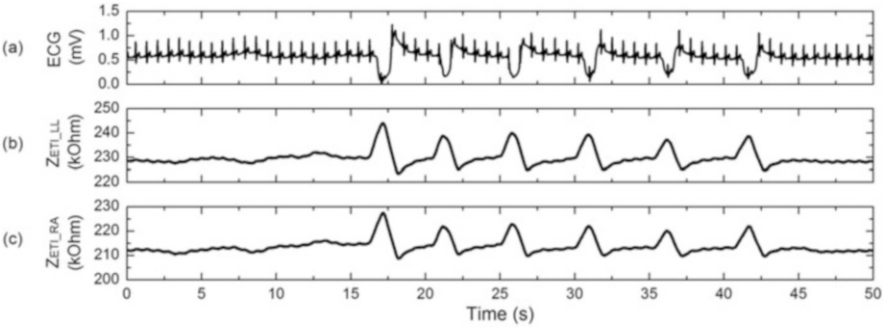

Fig. 4 (**a**) ECG signal and (**b**, **c**) MC-ETI signals [13]

$$Z_{\text{ETI\_LL}} \approx Z_{\text{LL}} \approx (Z_{\text{in}} - 2Z_{\text{LA}})V_{\text{ETI\_LL}}/2V_{\text{AC}} - 2Z_{\text{LA}} \tag{11}$$

$$Z_{\text{ETI\_RA}} \approx Z_{\text{RA}} \approx (Z_{\text{in}} - 2Z_{\text{LA}})V_{\text{ETI\_RA}}/2V_{\text{AC}} - 2Z_{\text{LA}} \tag{12}$$

The MC-ETI generating WECG and reference signals in the motionless and motion state with a period of 50 s are shown in Fig. 4. In the motion state (15–45 s), the WECG and MC-ETI signals have a high correlation. In the motionless state (0–15 and 45–50 s), these signals are stable but have low correlation.

The correlation between the MC-ETI and WECG signals over the time offset is shown in Fig. 5. There is no obvious peak in the black curve at zero time offset in the motionless state. However, in the motion state, the peak in the red curve at zero time offset shows a good correlation between these two signals.

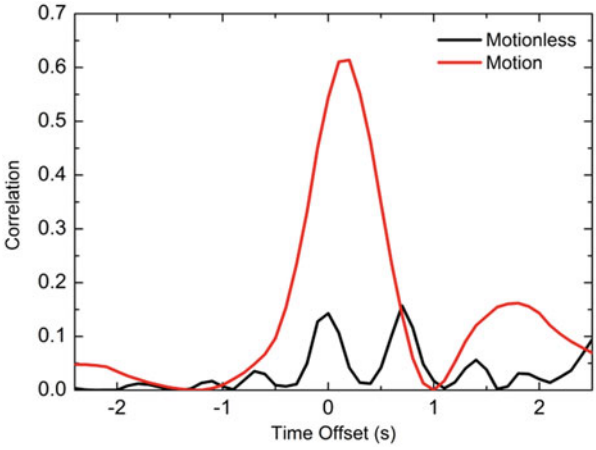

Fig. 5 Correlation as a function of the time offset [13]

4 Suppression of Motion Artefacts with an Adaptive Filter

With digital signal processing, the motion artefacts in ECG signals can be suppressed. Many studies have been reported on artefact suppression, which have mostly focused on two methods: blind source separation (BSS) and adaptive filtering.

Changli [14] reported that BSS has found many applications, including digital image processing, speech signal processing, medical signal processing, geophysical signal processing, communication signal processing and remote sensing image processing. When the mixing process and the original signals are unknown, BSS tries to decompose the observed sensor signals to obtain the unmixed source signals. However, given some assumptions, BSS has had great success, and many novel and effective methods have emerged.

Sweeney et al. [15] reported that, as a branch of BSS, independent component analysis (ICA) can separate different components from the source signals by defining them as statistically independent components.

Romero [16] reported that motion artefacts and ECG signals are statistically independent, so they can be separated by ICA.

However, ICA is restricted by data redundancy. ICA requires several independent sensors and cannot be used in one-channel sensor system, and it is necessary to ensure that the signal from each sensor is uncorrelated with other sensor signals. The large computational cost makes ICA very difficult to implement for real-time low-power applications.

The adaptive filter algorithm can automatically change its filter parameters and is widely used in the signal processing field.

Thakor and Zhu [17] reported that the adaptive filtering technique is useful in many biomedical applications. One simple but important application is in 60 Hz power line interference cancellation.

The adaptive filter has a low computational cost and high reliability, so it is very suitable for real-time low-power applications.

Tong et al. [18] reported that an adaptive filter using electrode motion as the reference signal can reduce motion artefacts. They measured electrode motion with two custom-developed sensors: anisotropic magnetoresistive (AMR) and accelerometer (ACC) sensors.

A two-axis AMR sensor was oriented parallel to the body surface, and a three-axis ACC sensor was developed using two dual-axis ACC chips.

Raya and Sison [19] reported using an accelerometer as a source of noise reference. Least mean squares (LMS) and recursive least squares (RLS) adaptive filter algorithms were used. They claimed that the major kinematic acceleration component during human movement is usually found in the vertical direction. Their adaptive filter can effectively reduce motion artefacts in stress ECGs.

Hamilton and Curley [20] reported that adaptive removal of motion artefacts can be 12.5 dB by using a skin stretching signal derived from sensors mounted on a foam electrode.

The most significant artefacts generated by skin stretching can still be adaptively removed. However, their sensors each cost approximately $600 because of the integrated stretching sensor.

Hamilton et al. [5] reported using a variable step size LMS (VSS-LMS) adaptive filter to remove motion artefacts in ECG signals:

$$\widehat{s}[n] = s[n] - \sum_{0}^{i} w_i n[n-i] \tag{13}$$

$$w_i^* = w_i + \beta \widehat{s}[n] \cdot n[n-i] \tag{14}$$

$$\beta[n] = \frac{a}{\sum_{i=1}^{200} \frac{|s[n-i] \cdot n[n-i]|}{200}} \tag{15}$$

where $s[n]$ represents the nth sample of ECG corrupted by noise, $n[n]$ represents the nth sample of the skin impedance or skin stretching signal and $\widehat{s}[n]$ is the nth estimation of the signal without motion artefacts. Note that 'w' represents filter coefficients that are iteratively updated after each sample by Eqs. (14) and (15), where 'a' is $8.44 \times 10^{-11}\ V^2$.

Wen-Ching et al. [21] reported using the normalized least mean squares (NLMS) adaptive filter algorithm to suppress the motion artefact from the primary input of ECGs. The 120-order finite impulse response (FIR) filter was adaptively adjusted by the NLMS with a 0.05 adaptive step size. They used the ACC signals and strain gauge (SG) signals as reference signals. They analysed the correlation between the ECG and ACC signals and the correlation between the ECG and SG signals. The higher one was chosen as the master reference, whereas the lower one was chosen as the slave reference.

Hyejung et al. [22] reported a two-stage cascade LMS adaptive filter for an ECG monitoring system. The first LMS stage consisted of analogue feedback, which

prevents signal saturation to reduce the input dynamic range. This approach employs a high-pass filter, which mainly targets the baseline wandering suppression to prevent signal saturation. An LMS algorithm with an adaptive step size is introduced and employed in the second LMS stage to remove the remaining motion artefact. The adaptive step size algorithm can achieve fast convergence to quickly track large sudden motion artefacts while preventing the distortion of the ECG component.

They reported a proposed LMS algorithm with adaptive step size control. The difference between their algorithm and the standard LMS algorithm is that they integrated an adaptive step size control block. The step size is updated to be large at a high signal-to-noise ratio (SNR) and small at a low SNR. The variation in both the reference signal $\sigma_{x(n)}$ and the input signal $\sigma_{d(n)}$ proportionally controls the step size adaptation function $\mu'(n)$, as shown in (16) and (17):

$$c(n) = \sigma_{x(n)} \cdot \sigma_{d(n)}/p \tag{16}$$

$$\mu'(n) = \begin{cases} \mu_0, 0 \leq c(n) < \mu_0 \\ c(n), \mu_0 \leq c(n) < 0.9 \\ 0.9, 0.9 \leq c(n) \end{cases} \tag{17}$$

where 'σ' is the standard deviation (STD) of the signal during the half cycle of the heart rate and 'p' is the experimentally determined constant, which sets the $\mu'(n)$ range between 0 and 1.

Romero et al. [23] reported the performance of different implementations of adaptive filter (AF) algorithms in the context of motion artefact reduction in ECG signals.

They used the LMS algorithm with the accelerometer as a reference and recursive least squares (RLS), convex AF and LMS sign-error with the skin-electrode impedance (SEI) as a reference.

Zhang et al. [24] reported a feed forward combined adaptive filter (FFC-AF) which is consisted of two separate AFs (one fast convergence speed AF 'FCS-AF' and one high convergence accuracy AF 'HCA-AF') and one combination AF. The parameter combination varies with the estimation of the reference signal stationary. Figure 6 describes the structure of the FFC-AF, and the corresponding equation is shown in Eq. (18):

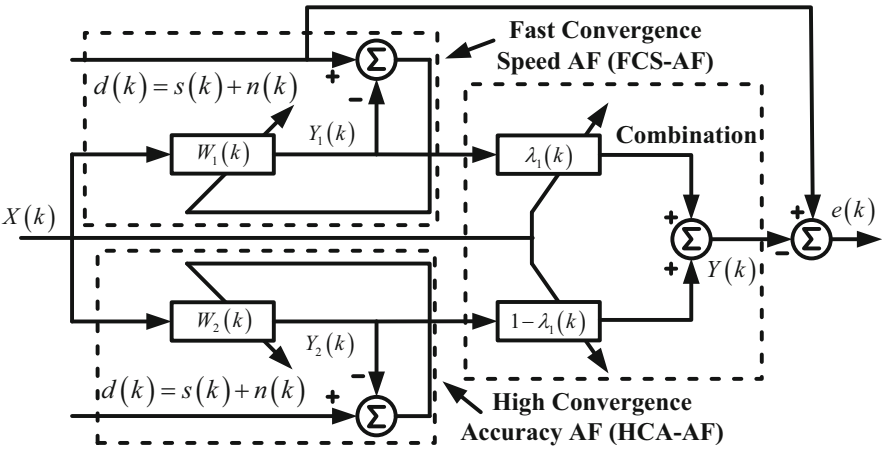

Fig. 6 FFC-AF block diagram [24]

$$\begin{cases} Y(k) = \lambda_1(k) \cdot Y_1(k) + [1 - \lambda_1(k)] \cdot Y_2(k) \\ Y_{1,2} = W_{1,2}^T(k) \cdot X(k) \\ \lambda_1(k) = \dfrac{2}{1 + \exp[-p \cdot \delta(k)]} - 1 \\ \delta(k) = \dfrac{1}{n} \sum\limits_{i=0}^{n-1} \left(x_i - \dfrac{1}{n} \sum\limits_{i=0}^{n-1} x_i \right)^2 \\ p = \begin{cases} p_{\text{non}} & \delta(k) \geq \delta_{\text{th}} \\ p_{\text{sta}} & \delta(k) < \delta_{\text{th}} \end{cases} \end{cases} \tag{18}$$

where $X(k)$ is the reference signal of two separate filters. Parameter $\lambda_1(k)$ is the combination weight of the filter output, which ranges between 0 and 1. At each iteration, $\lambda_1(k)$ updates its value according to the stationary degree of $X(k)$ by its variance $\delta(k)$. When $X(k)$ is in the stationary state, $\delta(k)$ is lower than the threshold δ_{th}, which means that AF is in the motion artefact (MA)-free state. Then, $\lambda_1(k)$ will remain at approximately 0 to increase the weight of the high convergence accuracy AF (HCA-AF) output. When $X(k)$ is in the nonstationary state, $\delta(k)$ is larger than δ_{th}, which means that AF is in the MA state. Then, $\lambda_1(k)$ will be maintained at approximately 1 to increase the weight of the fast convergence speed AF (FCS-AF) output.

Figure 7 shows the AF results, in which one triangle represents one QRS beat detection.

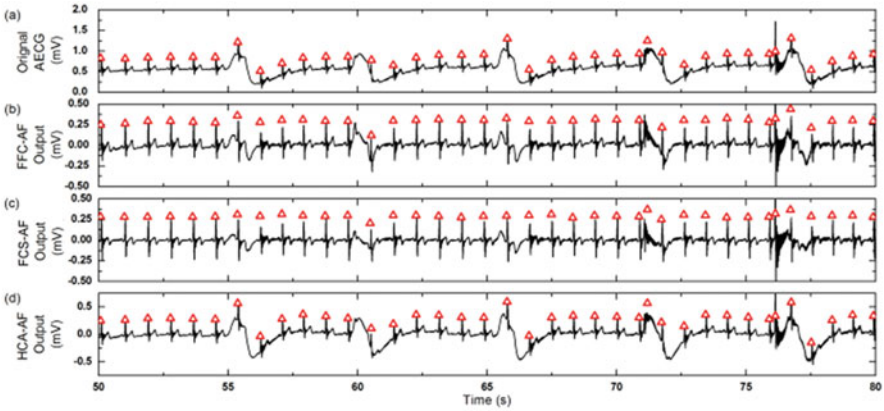

Fig. 7 Adaptive filter results and QRS beat detection [24]

5 Conclusion

With the development of wearable medical technology, an increasing number of patients with wearable healthcare monitoring devices are moving home from the hospital. ECGs are important vital tests, which are the basis of wearable healthcare monitoring.

This chapter illustrates the origins of motion artefacts in ECG signals. Interesting experimental results are introduced to describe the physical reasons for the observed motion artefacts. Moreover, circuit models are provided to qualitatively explain the motion artefacts. Then, we describe the detection of motion artefacts. Several electrical circuit architectures are provided from single-channel skin-electrode impedance measurements and multichannel skin-electrode impedance measurements. Finally, we summarize the application of an adaptive filter in motion artefact suppression. The LMS, NLMS, VSS-LMS, cascade LMS and FFC LMS algorithms are discussed. These algorithms form the foundation for developing an ECG system without motion artefacts.

References

1. Tam, H., Webster, J.G.: Minimizing electrode motion artifact by skin abrasion. IEEE Trans. Biomed. Eng. **BME-24**(2), 134–139 (1977)
2. Talhouet, H., Webster, J.G.: The origin of skin-stretch-caused motion artifacts under electrodes. Physiol. Meas. **17**(2), 81–93 (1996)
3. Thakor, N., Webster, J.: The origin of skin potential and its variations. Proc. Ann. Conf. Eng. Biol. Med. **20**, 212 (1978)
4. Burbank, D.P., Webster, J.G.: Reducing skin potential motion artifact by skin abrasion. Med. Biol. Eng. Comput. **16**(1), 31–38 (1978)

5. Hamilton, P. S., et al.: Comparison of methods for adaptive removal of motion artifact. In: Computing in Cardioliology, vol. 27, pp. 383–386 (2000)
6. Spinelli, E.M., et al.: A practical approach to electrode-skin impedance unbalance measurement. IEEE Trans. Biomed. Eng. **53**(7), 1451–1453 (2006)
7. Romero, I., et al.: Motion artifact reduction in ambulatory ECG monitoring: an integrated system approach. In: Proceedings of the 2nd Conference on Wireless Health, San Diego, California, 2011 (2011)
8. Ottenbacher, J., et al.: ECG electrodes for a context-aware cardiac permanent monitoring system. In: Magjarevic, R., Nagel, J.H. (eds.) World Congress on Medical Physics and Biomedical Engineering 2006, vol. 14, pp. 672–675. Springer, Berlin (2007)
9. Oberg, T.: A circuit for contact monitoring in electrocardiography. IEEE Trans. Biomed. Eng. **BME-29**(5), 361–364 (1982)
10. Degen, T., Jackel, H.: Continuous monitoring of electrode skin impedance mismatch during bioelectric recordings. IEEE Trans. Biomed. Eng. **55**(6), 1711–1715 (2008)
11. Bertrand, A., et al.: Motion artifact reduction in EEG recordings using multi-channel contact impedance measurements. In: Biomedical Circuits and Systems Conference (BioCAS), pp. 258–261 (2013)
12. Comert, A., Hyttinen, J.: Impedance spectroscopy of changes in skin-electrode impedance induced by motion. Biomed. Eng. Online. **13**, 149 (2014)
13. Zhang, H., et al.: A multi-channel electrode tissue impedance detection approach for motion artifact suppression in ambulatory electrocardiography. In: Computing in Cardiology Conference (CinC), pp. 117–120 (2015)
14. Changli, L.: Study on Some Algorithms for Blind Source Separation and Their Applications. Xidian University, Xian (2010)
15. Sweeney, K.T., et al.: Artifact removal in physiological signals—practices and possibilities. IEEE Trans. Inf. Technol. Biomed. **16**(3), 488–500 (2012)
16. Romero, I.: PCA and ICA applied to noise reduction in multi-lead ECG. In: Computing in Cardiology, pp. 613–616 (2011)
17. Thakor, N.V., Zhu, Y.-S.: Applications of adaptive filtering to ECG analysis: noise cancellation and arrhythmia detection. IEEE Trans. Biomed. Eng. **38**(8), 785–794 (1991)
18. Tong, D., et al.: Adaptive reduction of motion artifact in the electrocardiogram. In: 24th Annual Conference and the Annual Fall Meeting of the Biomedical Engineering Society EMBS/BMES Conference, vol. 2, pp. 1403–1404. IEEE (2002)
19. Raya, M.A.D., Sison, L.G.: Adaptive noise cancelling of motion artifact in stress ECG signals using accelerometer. Eng. Med. Biol. **2**, 1756–1757 (2002)
20. Hamilton, P.S., Curley, M.G.: Adaptive removal of motion artifact. IEEE Eng. Med. Biol. Soc. **1**, 297–299 (1997)
21. Wen-Ching, L., et al.: Adaptive reduction of motion artifact in a portable ECG system. Sensors. 704–707 (2010)
22. Hyejung, K., et al.: Motion artifact removal using cascade adaptive filtering for ambulatory ECG monitoring system. In: Biomedical Circuits and Systems Conference (BioCAS), pp. 160–163 (2012)
23. Romero, I., et al.: Adaptive filtering in ECG denoising: a comparative study. In: Computing in Cardiology (CinC), 2012, pp. 45–48 (2012)
24. Zhang, H., et al.: Motion artifact suppression in ambulatory ECG with feed forward combined adaptive filter. In: 2016 Computing in Cardiology Conference (CinC). IEEE (2016)

Part IV
Latest Techniques for Machine Leaning

Data Augmentation for Deep Learning-Based ECG Analysis

Qing Pan, Xinyi Li, and Luping Fang

Abstract Deep learning has become the technology that gets the most attention in recent years owing to its admirable performance compared to the conventional methods in a series of tasks. Though its application in electrocardiogram (ECG) analysis has enhanced the understanding and the applicability of many disease diagnosis in clinic, lack of annotated data hampers the deep learning-based ECG analysis as large amount of data is required for a well-performed deep learning model. Data augmentation, which refers to the procedure that enriches the dataset by introducing unobserved samples, plays an important role in this respect. Despite the successful usage of data augmentation in the image-based deep learning analysis, its application in one-dimensional physiological signals, such as ECG, is still limited. In this chapter, we summarize the data augmentation methods applicable for ECG analysis and examine their performance on a task for detecting atrial fibrillation (AF).

Keywords Data augmentation · Electrocardiogram · Deep learning · Atrial fibrillation

1 Introduction

Analysis of electrocardiogram (ECG) plays a significant role in diagnosis and screening of cardiac diseases. It allows detecting the occurrence of arrhythmia, which is often responsible for sudden cardiac death and largely related to a series of cardiovascular diseases [1–3]. In addition, ECG analysis locates the R peak in the ECG waveform to provide the basis for the analysis of R-R interval pattern, which contains important information about the autonomic nervous system [4]. ECG

Q. Pan (✉) · X. Li · L. Fang (✉)
College of Information Engineering, Zhejiang University of Technology, Hangzhou, People's Republic of China
e-mail: pqpq@zjut.edu.cn; flp@zjut.edu.cn

C. Liu, J. Li (eds.), *Feature Engineering and Computational Intelligence in ECG Monitoring*, https://doi.org/10.1007/978-981-15-3824-7_6

analysis also contributes to other related fields, such as biometric identification [5] and emotion assessment [6].

Computerized algorithms have been applied for ECG analysis for decades. A lot of efforts have been made in locating the QRS complex, recognizing various heartbeat types, detecting arrhythmia, and so on [7, 8]. For example, the most famous Pan–Tompkins algorithm for QRS complex detection [9] automatically adjusts thresholds and parameters periodically to adapt to such ECG changes such as QRS morphology and heart rate, which makes the detection more effective and saves much human cost. Wavelet analysis is widely used in the detection of QRS complexes [10] and the onsets and offsets of P- and T-waves [11], recognizing and describing isolated heartbeats [12]. With the help of these algorithms, automatic analysis of ECG becomes possible and increases the efficiency of cardiologists greatly. It also eliminates the intra- and interobserver variability in recognition and diagnosis.

Machine learning has shown great advantages and potentials in physiological signal processing. Zhao and Zhang [13] applied support vector machine (SVM) with Gaussian kernel to classify six heart rhythm types, and the accuracy of recognition has reached 99.68%. Emanet [14] used random forest to classify five types of ECG beats with the accuracy of 99.8% and increased the speed. Li et al. [15] also used a SVM to classify the ventricular fibrillation and tachycardia, and the validation accuracy has reached to 96.3%.

Although machine learning methods have got many achievements in the field of ECG processing, conventional machine learning techniques require complicated engineering and domain expertise to extract suitable features and design a feature extractor. However, representing the raw data with appropriate vectors and building an effective machine learning system [16] are always time consuming and difficult.

By contrast, the emergence of deep learning, which is capable of extracting the features by the model itself, overcomes the current challenges in the machine learning field and, therefore, boosts its application in many fields, such as image processing [17–19] and speech recognition [20–22]. Since deep learning has got great success in large-scale automatic speech recognition, people start to realize its great potential in other sequence signal processing fields, such as biomedical signal processing. As ECG is a kind of time series with specific shape morphological features, deep learning has become an important approach that can benefit in promoting the development of automated ECG analysis.

2 Deep Learning Models for ECG Analysis

In supervised learning tasks, deep learning network models can usually be divided into two types: convolutional neural network (CNN) and recurrent neural network (RNN). CNN is a kind of neural network specially used to process data with similar grid structure, which uses convolution instead of matrix multiplication at least in one layer of the network [23]. Because of its excellent performance on the processing of

entire visual fields, it is widely used for image classification. RNN is a kind of neural network used to process sequential data $\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(\tau)}$, which can process sequences of variable length. So they have efficiency on the tasks such as unsegmented handwriting recognition [24] or speech recognition [25, 26]. Long short-term memory (LSTM) networks, which were put forward by Hochreiter and Schmidhuber in 1997 [27], have become the main method of prediction and classification of time sequences.

In the past few years, the application of deep learning models for ECG analysis increases rapidly. Both CNN and RNN models are used for classifying beat type, locating QRS complex, and detecting arrhythmia. The most representative study is carried out by Andrew Ng et al. They develop a 34-layer CNN model to beat the certified cardiologists in detecting a wide range of heart arrhythmias from ECGs recorded with a single-lead wearable monitor [28]. Ubeyli presented the input to a RNN classifier by composing feature vectors and Lyapunov exponents of each ECG beat in classifying ECG signals [29]. This method has successfully classified four different beat types. Chauhan and his group used deep LSTM networks to detect abnormal and normal signals in ECG data [30]. The performance of the depth detection system based on LSTM is 96.45%. In addition, the authors have examined the type of abnormal signal detected for which type they are successful. Rahhal et al. proposed a method of ECG data classification based on depth analysis [31]. In the learning attribute stage, the sparse constraint method is used to superimpose the autoencoder to remove noise. These learning features were classified by deep neural network (DNN) structure. Recently, a model specially for atrial fibrillation (AF) detection was put forward [32], which added an attention network to visualize which regions of the signal are important while there is an underlying AF arrhythmia in the signal. So, the model can automatically extract features from the focused parts of the signal. It also transformed the signal into spectrograms and fed the images to a CNN to extract more features.

3 Data Augmentation Approaches for ECG Analysis

Success of ECG analysis based on deep learning relies on rich annotated dataset. Although large amount of ECG recordings is available nowadays thanks to fast developing digital medical devices, constructing a high-quality annotated ECG dataset for deep learning model remains challenging. On one hand, it is demanding for the experienced cardiologists to annotate the recordings. Therefore, the annotated data are insufficient in quantity and diversity. On the other hand, the incidence of abnormal cardiac events is much lower than that of normal beats, resulting in a highly imbalanced dataset. If not properly treated, the imbalance will lead to biased classification results using deep learning models.

Data augmentation may shed a light to the problem. It refers to a procedure that enriches the dataset by introducing unobserved samples [33]. Many successful deep learning applications benefit from efficient data augmentation approaches. For

instance, in image-based supervised learning tasks, simple parametric transformations such as rotation and scaling are usually used to perform data augmentation [34]. While data augmentation has been widely used in many image-based deep learning applications [35–37], its effect on sequential physiological signals, such as ECG, has not been extensively explored yet. In this chapter, we summarize the currently available data augmentation methods for ECG and demonstrate their application on deep learning-based ECG analysis. As ECG can be processed in either episode-based form or beat-by-beat manner for different research purposes, this chapter introduces the data augmentation methods for ECG records and ECG beats, respectively.

3.1 Data Augmentation for ECG Records

3.1.1 Window Slicing (WS)

Window slicing (WS) for time series data augmentation was first proposed by Cui et al. [38]. It follows the assumption that a slice of the original recording holds the same label as the original one. For an ECG recording T consisting of a set of data points $\{t_1, t_2, \ldots t_i, \ldots, t_n\}$, a slice of the original time series can be denoted as S_i, which is a subset $\{t_i, \ldots, t_{i+s-1}\}$ extracted from the original recording T, requiring $0 < i \leq n - s + 1$, where s is the width of the slicing window. The window is slid on the original recording to generate a set of new recordings $\{S_1, S_2, \ldots, S_m\}$, where all the new recordings in the set have the same length s, and the same label as their original recording Tdoes. Figure 1 illustrates the WS method.

Selection of the window size is critical for the WS method. Selecting a small window size may render loss of the discriminative information of an ECG recording. Oppositely, a large size leads to too much overlapping, which results in high

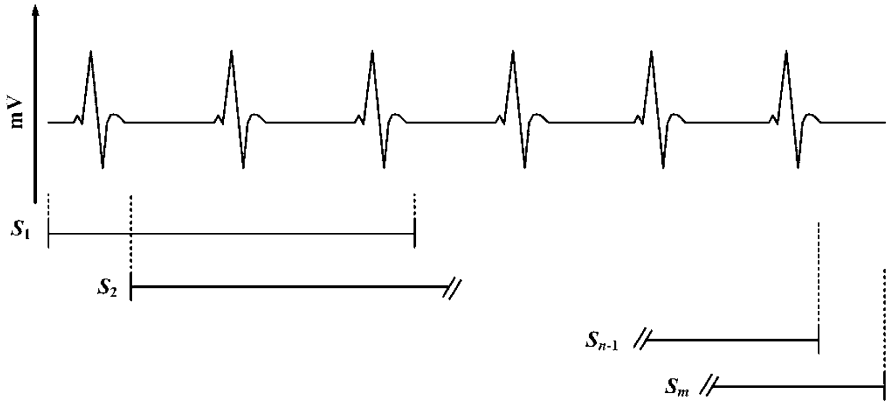

Fig. 1 Illustration of window-slicing method. With a sliding window of fixed width, cut the original signal into m slices

similarity among augmented samples. The selection of an appropriate window size should be the focus because cutting the time series arbitrarily tends to destroy temporal correlation in the data [39, 40].

The step length is another parameter. It controls the number of sliced snippets, which depends on how many new samples are needed. Also, it determines the similarity among the sliced snippets.

When training the deep learning model, all the augmented slices are considered as independent training samples. At the testing stage, because the training samples are shortened compared to the original ones, the testing samples need to be cut into several snippets (odd number in general) using the same window for data augmentation. Then, the classifier is applied to each snippet, and a majority voter is utilized to make the prediction.

WS method is especially suitable for CNN models, which require all the samples to have equal length. It has been widely used in many experiments. It has achieved excellent performance on University of California, Riverside Home Time Series Classification Archive, when using multiscale CNN approach as the training model [38], and the superior performance is ascribed to the WS applied to the dataset [41]. The greatest advantage of WS is while retaining most of the original information, increasing the number and diversity of data.

3.1.2 Permutation

In mathematics, permutation is an act which arranges the elements of a set into a sequence. If the set is already ordered, its elements shall be rearranged by this act. Permutation was firstly proposed as a data augmentation method in [42]. It is a simple way to randomly perturb the temporal location of the events.

Permutation can be used in two ways. On the one hand, it can be used in a whole recording. It is carried out by dividing each sample into N_p subsegments with the same length firstly. A set was defined with the subsegments from one sample. Then, all the subsegments in one set were randomly permuted and assemble the perturbed segments to generate a new recording with the same label of the original signal. This operation should be repeated t_C times, which is the parameter that could be used for balancing different categories. If we want to balance the number of samples in each category to N_{max}, the t_C of class C with N_C samples can be calculated as Eq. (1):

$$t_C = \left[\frac{N_{max}}{N_C}\right] \tag{1}$$

where $[x]$ indicates the rounding value of x and N_{max} represents the number of samples after this operation.

In addition, the number of samples in the dataset can be increased to M ($M = \lambda N_{max}$) by modifying the Eq. (2) as:

$$t_C = \left\lfloor \frac{M}{N_C} \right\rfloor = \left\lfloor \frac{\lambda N_{\max}}{N_C} \right\rfloor \tag{2}$$

The number of subsegments N_p was defined to satisfy:

$$N_p! > t_C > (N_p - 1)! \tag{3}$$

in order to guarantee that the permutation does not produce repeating samples.

On the other hand, permutation can be used in together with the WS method to randomly perturb the temporal location of within-window slices. This method can provide more diversities because it has two parameters: the width of the window and the permuting times. The permutation method combined with WS was tested on wearable sensor data for Parkinson's disease monitoring [42], and the result indicated that permutation has an improvement on the classification of time series.

3.1.3 Concatenating and Resampling Method

Although permutation method increases the diversity of the dataset, slicing the ECG episodes and randomly perturbing the events may destroy the orders and morphologies of the heartbeats. To deal with this problem, a slicing based on characteristic points and resampling method was proposed [43], aiming to increase the diversity of the samples, balance the number of samples among the classes, and reserve the physiological information at the same time. The diagram of the strategy is shown in Fig. 2.

First, the Pan–Tompkins QRS detector [9] is used to locate the QRS complex of each heartbeat. Then, the location of the notch between Q-wave and R-peak (starting point of the QRS) is determined as the characteristic point. All the characteristic points are represented as $q_1, q_2, q_3 \ldots q_n$. The episode between q_1 and q_n is defined as the selected sequence (SS).

Second, the SS is duplicated and concatenated to the original series. Finally, it uses a sliding window of size l_w to generate t_C new samples from the reconstructed series. $t_C$is calculated by Eq. (1) or Eq. (2). The generated segments are identified as $S_i (i = 1, 2, \ldots, t_C)$. To ensure that S_i have the same annotation with the original recording, l_w should be greater than or equal to the length of the original ECG recording.

This method duplicates the subsequence, which is sliced from the original data based on the characteristic points, to generate a series with the same pattern fragment. The duplication makes full use of the fragment with specific characters, which can help the deep learning models to learn the determinant features more effectively. Then, as mentioned above, the WS method is applied to generate more samples.

As the concatenation may disrupt the patterns of R-R interval of the raw signal, more strict concatenation making use of the information of heartbeat dynamics [44] could be developed for better quality of data augmentation.

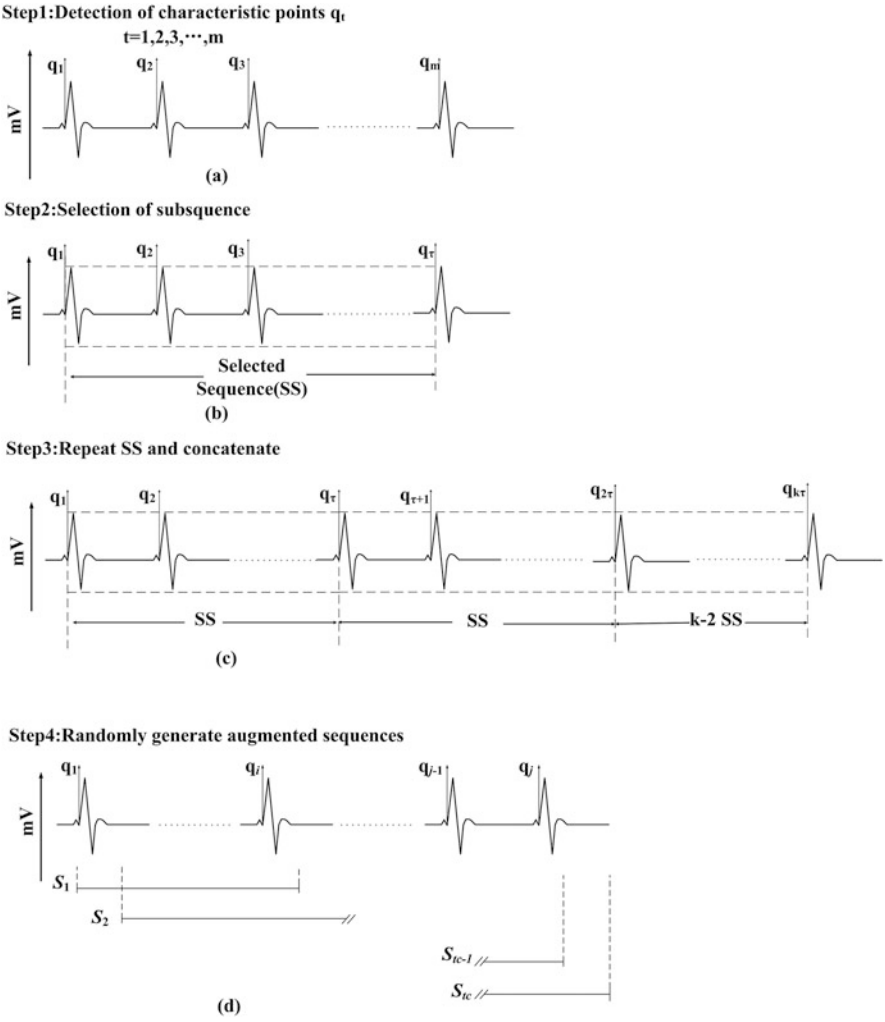

Fig. 2 Illustration of concatenating and resampling method. (**a**) Step 1: detect the characteristic points as starting points, which are denoted as $q_1, q_2, q_3, \ldots q_m$; (**b**) Step 2: select a subsequence SS $\{q_1, q_2, q_3, \ldots q_\tau\}$; (**c**) Step 3: duplicate the SS and concatenate it with itself; (**d**) Step 4: use a sliding window of size l_w to slice the sequence which was generated in Step 2 into t_C slices

3.1.4 Window Warping (WW)

Le Guennec et al. introduced a more time-series-specific method based on WS called window warping (WW) [45], which aims at increasing the diversity of data as well as the amount. It randomly selects a slice of the time series using a sliding window and warps it by squeezing or dilating. When applied to the ECG signal, WW can be illustrated as in Fig. 3.

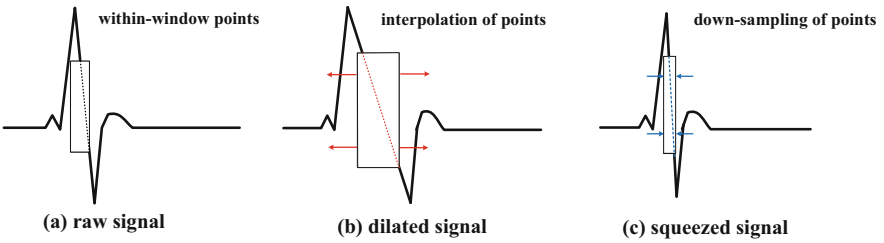

Fig. 3 Illustration of window warping. (**a**) The raw data with a sliding window, the data points within the window represent the segment which will be warped; (**b**) the signal after dilating the data points by interpolation; (**c**) the signal after squeezing by down sampling

Like WS, the width of the sliding window is a parameter of this method. As mentioned above, the WS method can be adopted in order to ensure that subsequences from multilength original series to have same length for training the network. The other parameter is the warping ratio, which controls the morphology of the generated signal. A proper warping ratio guarantees that the augmented sample remains physiological and holds the same label as the original one.

The selection of the warping area is critical for the performance of the WW method, because for an ECG recording, the time scale has significant physiological meanings. Different warping areas may result very different interpretations. Therefore, the WW method is sometimes two-sided. It may introduce nonredundant samples to the dataset with proper selection of parameters. However, if the warping area or the warping ratio is not carefully selected, the generated new samples may destroy the physiological significance, which will impair the deep learning-based analysis.

3.2 Data Augmentation for ECG Beats

There are about 100,000 heartbeats every day for one person. Therefore, the amount of heartbeats is usually sufficient for a beat-wise deep learning-based analysis. For example, in the widely used Massachusetts Institute of Technology-Beth Israel Hospital (MIT-BIH) Arrhythmia Database, 48 ECG recordings lasting about 30 min contribute 110,159 beats in total. However, as the incidence of abnormal beats is much lower than the normal ones, the dataset is often highly imbalanced. In addition, such a dataset is suffering from the low diversity as the number of subjects is small.

Currently, most studies are only simply removing samples from the prominent category for balancing the dataset. Acharya et al. proposed a method to generate synthetic data [46] by varying the standard deviation and mean of Z-score calculated

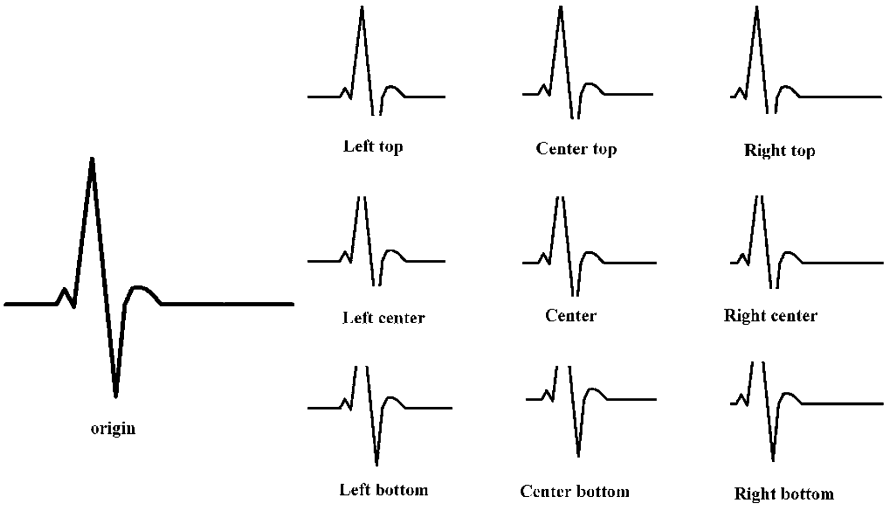

Fig. 4 Illustration of the cropping data augmentation [47]

from the original normalized ECG signals. The data augmentation introduces high diversity to the dataset and, therefore, leads to that the trained CNN achieved an accuracy of 94.03% and 93.47% in the diagnostic classification of heartbeats in original and noise-free ECGs, respectively. However, the available data augmentation methods for ECG beats are limited.

Despite the fact that ECG is intrinsically a one-dimensional (1D) time series, it can also be represented as two dimensional (2D) if it is plotted as an image, to allow utilization of the powerful 2D CNN models. Jun et al. firstly proposed an ECG arrhythmia classification method using a deep 2D CNN with grayscale ECG images [47]. They plotted each heartbeat in a single image. In order to reduce overfitting and maintain a balanced distribution between classes, they applied cropping, which is the most widely used method for image augmentation. Every ECG beat was augmented with nine different cropping methods: left top, center top, right top, center left, center, center right, left bottom, center bottom, and right bottom (Fig. 4).

The cropping size is an adjustable parameter in the data augmentation. The suitable size of cropping is essential for guaranteeing the discriminative information for each class is kept. Then, these augmented images were resized to the original size and fed into the network. This method transformed the 1D sequence into 2D images, so that it can use the CNN models to classify the images, which have much faster training speed than RNN models. It also has a great contribution in balancing the ECG dataset and increases the diversity as well.

3.3 Other Augmentation Methods

In addition to the methods mentioned above, several other techniques are available for augmenting physiological time series signals. Jittering is a maneuver that randomly adds jitter on the ECG recordings to simulate the noise [21]. Scaling multiplies every data point of the series by a random scalar to compress or stretch the magnitude. Rotating turns the signal upside down. Cropping randomly removes a part of the recording. The applicability of these data augmentation methods, however, is weak in the ECG. Jittering adds redundant noise to the recordings which are already removed from noise and baseline. Other methods have changed the morphologies of recordings which have physiological significance. Le Guennec et al. proposed a method for data augmentation that is a mixture between various datasets [45]. It requests that the chosen datasets share similar time-series length. They used this method to learn the convolutional part of their model in an unsupervised manner, which reached an excellent result.

4 Applications on Atrial Fibrillation Detection

In this section, data from the PhysioNet/Computing in Cardiology (CinC) Challenge 2017 dataset were used to illustrate the application of various data augmentation methods.

The dataset contains 8528 single-lead ECG recordings with different lengths (9–61 s), which are divided into four types: (1) normal sinus rhythm (NSR), (2) AF, (3) other abnormal rhythms (OAR), and (4) noise (NO, too noisy to classify). Note that the OAR class contains signals with premature ventricular contraction (PVC), premature atrial contraction (PAC), and other abnormal beats. The labels represented the characters of each recording instead of each single beat. All the recordings were collected using the AliveCor device (AliveCor Inc., USA), which were sampled at 300 Hz. The distribution of data is shown in Table 1, and the ECG waveform of each class is shown in Fig. 5.

The effect the data augmentation methods mentioned in Sects. 3.1.1–3.1.3 were evaluated on a two-layer LSTM network, which is a widely used RNN architecture. The structure of each cell is show as Fig. 6.

Table 1 Detailed information of the dataset

Type	# Recording	Time length (s)				
		Mean	SD	Max	Median	Min
NSR	5154	31.9	10.0	61.0	30.0	9.0
AF	771	31.6	12.5	60.0	30.0	10.0
OAR	2557	34.1	11.8	60.9	30.0	9.1
NO	46	27.1	9.0	60.0	30.0	10.2

NSR normal sinus rhythm, *AF* atrial fibrillation, *OAR* other abnormal rhythms, *NO* too noisy to classify

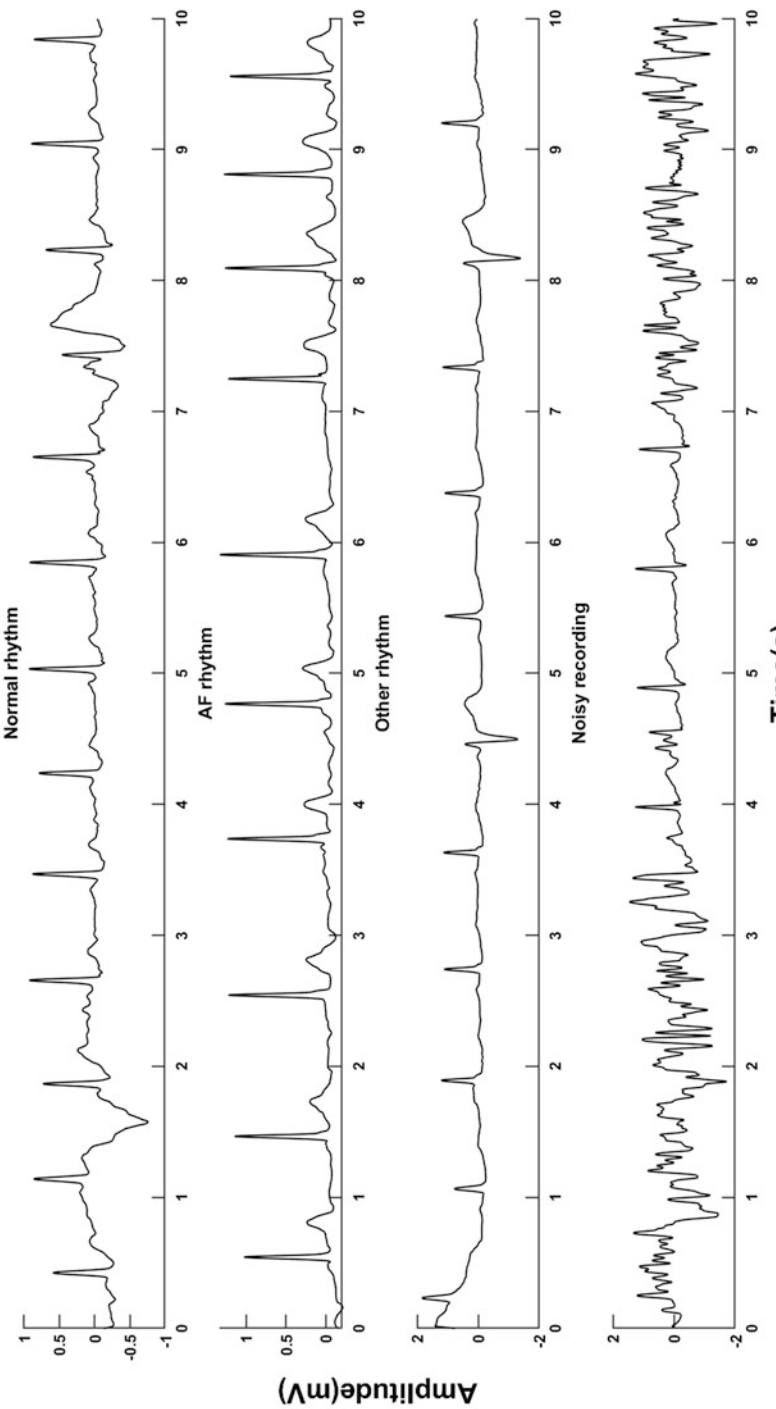

Fig. 5 Examples of four types of ECG waveforms (lasting 10 s). From top to bottom, they are NSR, AF, OAR, and NO [48]

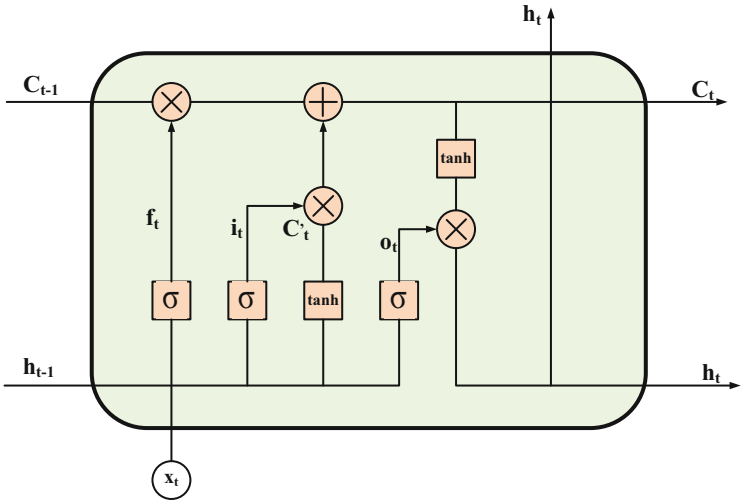

Fig. 6 The structure of each cell of LSTM

Each cell of LSTM consists of a forget gate, an input gate, an output gate, and cell states C_t and C_{t-1}.

The forget gate is used to control whether to forget the hidden cell state of the upper layer with a certain probability. Its activation at time t is defined as Eq. (4).

$$f_t = \sigma(\mathbf{W}[\mathbf{x}_t, \mathbf{h}_{t-1}] + \mathbf{b}_\mathrm{f}) \tag{4}$$

The input gate is responsible for processing input of the current sequence position. Its activation at time t is defined as Eq. (5).

$$i_t = \sigma(\mathbf{W}[\mathbf{x}_t, \mathbf{h}_{t-1}] + \mathbf{b}_i) \tag{5}$$

Finally, the output gate updates the cell state as Eq. (6).

$$o_t = \sigma(\mathbf{W}[\mathbf{x}_t, \mathbf{h}_{t-1} + \mathbf{b}_\mathrm{o}]) \tag{6}$$

Among Eqs. (4), (5), and (6), $\mathbf{x}_t$ represents the input of the sequence at time t. $\mathbf{C}$ represents the memory at each time step. $\mathbf{h}_{t-1}$ is the activation of the previous state. $\mathbf{W}$ is the weight vector, and $\mathbf{b}$ is the bias vector.

The flowchart of the evaluation is shown in Fig. 7.

The number of recordings in each category was augmented to 10,000. Then, the augmented dataset was split into a training and test set. Ninety percent of the recordings were selected as the training set and the rest 10% as the test set. The training set contained a validation set with 30% of data, which was used for early stopping, and the remaining data were used for network training. The performance of each method was examined by applying a fivefold cross-validation.

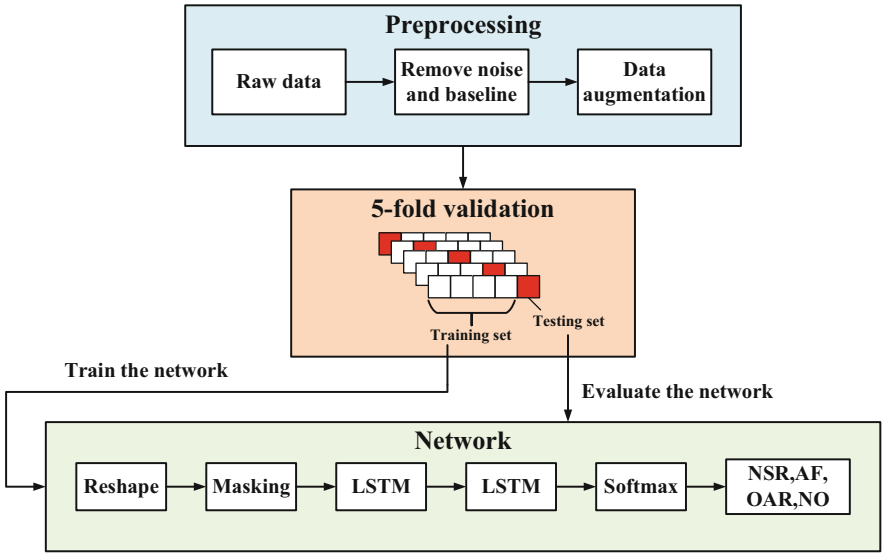

Fig. 7 Flowchart of the evaluation approach

We used cross-entropy as the loss function and applied a training optimization strategy and an early stopping criterion during the training phase. An Adam optimizer [49] was selected for the training process, whose initial learning rate was set as 0.001. The training would be terminated without the loss of validation decreasing ten consecutive epochs. During the training procedure, any model with a decreased loss of validation would be saved for testing. We also applied dropout [50] and recurrent dropout [51] strategies during the training procedure to avoid overfitting.

To evaluate the power of the proposed data augmentation methods for AF detection, specificity and sensitivity were calculated, which are defined as:

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}} \tag{7}$$

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{8}$$

where TP (true positive), TN (true negative), FP (false positive), and FN (false negative) indicate the number of recordings correctly labeled as AF, correctly identified as non-AF, incorrectly labeled as AF, and unrecognized as AF should be identified as AF, respectively.

F_1 score was used as the evaluation index following the rules of the PhysioNet/CinC challenge 2017 [48]. The F_1 score for class C is given as:

$$F_{1C} = \frac{2 \times \mathrm{TP}}{2 \times \mathrm{TP} + \mathrm{FP} + \mathrm{FN}} \tag{9}$$

The multiclass F_1 score is the weighted average of the F_1 score for the three classes "NSR," "AF," and "OAR" following the rule of the PhysioNet/CinC challenge 2017.

$$F_1 = \frac{F_{1N} + F_{1A} + F_{1O}}{3} \tag{10}$$

The accuracy for all classes is calculated as:

$$\mathrm{Accuracy} = \frac{\mathrm{TP_{all}}}{\mathrm{TP_{all}} + \mathrm{FN_{all}}} \tag{11}$$

Confusion matrix was used to demonstrate the performance of various methods more intuitively based on the best performance in the fivefolds.

4.1 Results with Raw Dataset

Table 2 shows the F_1 scores and accuracy based on raw data, and Table 3 shows the specificity and sensitivity of AF.

As seen in Table 2, the performance of the model is low without data augmentation mainly due to the imbalanced and insufficient dataset. The F_1 score of AF is much lower than other two classes. The imbalanced dataset leads the model to classify most testing samples to the prominent categories, which resulted in confusing accuracy. It needs to be noted that the variance of accuracy is high, which means the model is not robust. The low specificity of AF indicates that most AF samples have not been identified correctly.

Figure 8 presents the confusion matrix of the best fold achieved by the model based on the raw data. The diagonal denotes the number of recordings correctly

Table 2 F_1 scores and accuracy based on raw data

Index	Mean ± SD
F_{1N}	0.701 ± 0.076
F_{1A}	0.080 ± 0.147
F_{1O}	0.427 ± 0.009
F_1	0.403 ± 0.065
Accuracy (%)	58.030 ± 5.300

Table 3 Specificity and sensitivity of AF

Index	Mean ± SD
Specificity (%)	7.792 ± 14.944
Sensitivity (%)	98.737 ± 2.526

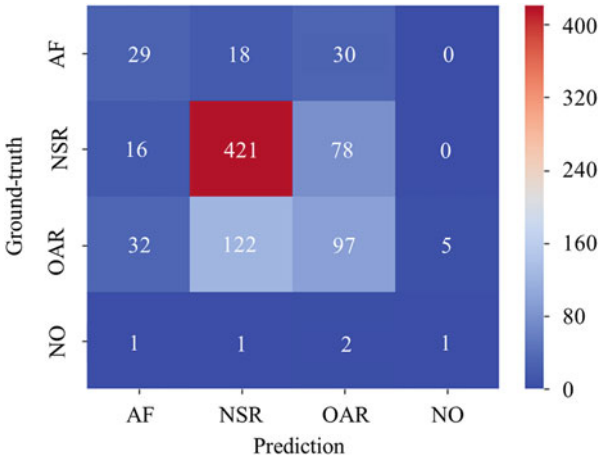

Fig. 8 Confusion matrix achieved by the raw dataset

Table 4 F_1 scores and accuracy based on WS

Index	Mean ± SD
F_{1N}	0.836 ± 0.063
F_{1A}	0.587 ± 0.177
F_{1O}	0.563 ± 0.231
F_1	0.662 ± 0.157
Accuracy (%)	73.505 ± 10.475

Table 5 Specificity and sensitivity of AF

Index	Mean ± SD
Specificity (%)	76.623 ± 4.107
Sensitivity (%)	95.747 ± 0.866

labeled, while the other cells denote the number of misclassified recordings. As shown in the figure, more samples were misclassified than the number of correctly classified samples for AF, OAR, and NO, suggesting that the model gives an inferior performance based on the raw dataset.

4.2 Results with Window Slicing (WS)

Tables 4 and 5 show the result of the network when using WS to augment the dataset. Every ECG recording was sliced into a shorter one with the sliding window size being fixed at 90% of the original recording length.

Comparing with the results shown in Sect. 4.1, it is clear that the F_1 scores of all the categories have improved significantly. The size of the dataset is enlarged, which enables the model to learn more features and conduct classification more accurately. And the number of each samples of each category has been set to a fixed one, so the model can learn features of each category fairly to avoid biased results.

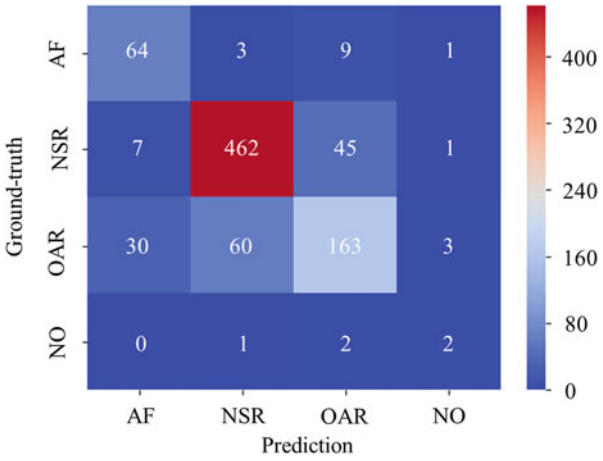

Fig. 9 Confusion matrix achieved by WS method

Table 6 F_1 scores and accuracy based on permutation

Index	Mean ± SD
F_{1N}	0.851 ± 0.015
F_{1A}	0.635 ± 0.038
F_{1O}	0.624 ± 0.024
F_1	0.703 ± 0.022
Accuracy (%)	75.967 ± 1.750

As shown in Table 5, the specificity of AF has been increased largely than using raw data, indicating the importance of data augmentation. As shown in Fig. 9, the accuracy of model prediction has been significantly improved. Only 13 AF samples are misclassified, compared to 48 based on raw data. Meanwhile, more testing samples of other categories have been classified correctly, suggesting the effectiveness of the WS method.

4.3 Results with Permutation

In order to balance the dataset, each recording was sliced into different snippets according to the full arrangement formula (12).

$$A_n^n = M \tag{12}$$

In order to reach the $M = 10,000$, n is set as 4 for AF, 3 for NSR and OAR, and 6 for NO. Then, these segments were randomly permutated and arranged to generate new signals. Tables 6 and 7 show the results of permutation.

It is observed that permutation performed better than WS. This is because in addition to increasing the dataset size, permutation also increases the diversity,

Table 7 Specificity and sensitivity of AF

Index	Mean ± SD
Specificity (%)	67.013 ± 7.189
Sensitivity (%)	95.644 ± 1.272

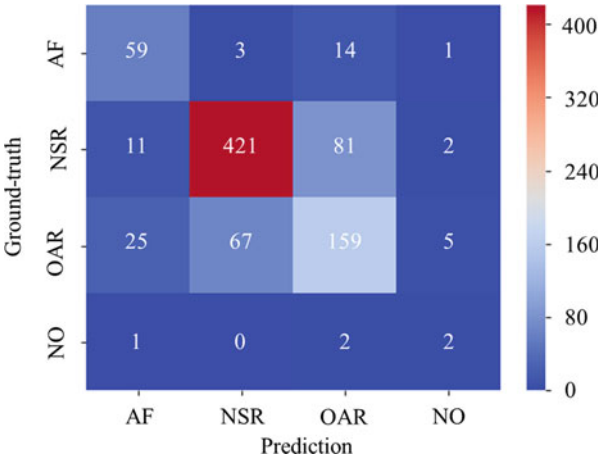

Fig. 10 Confusion matrix achieved by permutation method

Table 8 F_1 scores and accuracy based on concatenating and resampling method

Index	Mean ± SD
F_{1N}	0.883 ± 0.007
F_{1A}	0.715 ± 0.032
F_{1O}	0.696 ± 0.034
F_1	0.765 ± 0.023
Accuracy (%)	80.750 ± 1.605

which can improve the generalization ability of the model. It can be seen in Table 7 that the detection accuracy of AF is further improved compared to using WS.

Figure 10 indicates that permutation also increases the number of correctly classified. Although the correct values are lower than WS methods, it should be noted that the variance is much lower as well, because the permutation method generates more samples with more diverse morphologies, which helps the model to learn better.

4.4 Results with Concatenating and Resampling Method

Tables 8 and 9 demonstrate the performance of the model when using the dataset augmented by concatenating and resampling method.

Table 8 shows that the F_1 scores and accuracy are all improved compared to the results using WS and permutation. This method retains more original information than randomly segmenting the original signal which may destroy its morphology.

Table 9 Specificity and sensitivity of AF

Index	Mean ± SD
Specificity (%)	83.117 ± 5.631
Sensitivity (%)	95.103 ± 1.015

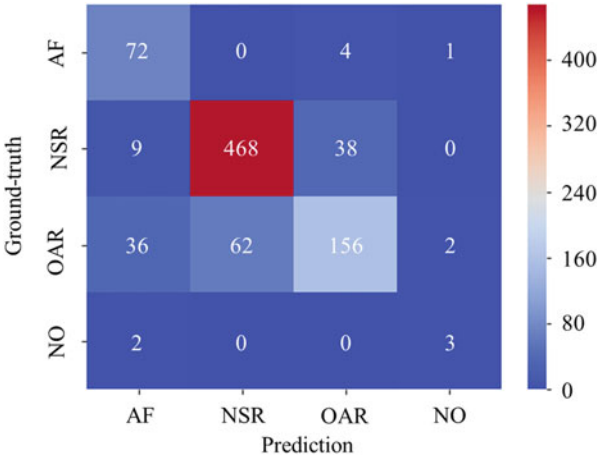

Fig. 11 Confusion matrix achieved by concatenating and resampling method

Repeating and concatenating the snippet can make the feature more prominent, which also helps the model to learn more detailed features. So, it has a slight improvement on the results compared with the two methods mentioned above.

Table 9 shows that this method also gives better sensitivity and specificity. Figure 11 presents the confusion matrix of this method. Only five testing samples of AF are misclassified. And the total correct values are greater than those of other methods.

5 Discussions and Conclusions

Because of the complexity of unstructured signal and the disturbance of noise affecting the patterns, conventional machine learning has great limitations in ECG observation. Deep neural networks have shown to be very powerful to learn the complex patterns in the ECG. However, lack of large amount of data with balanced distribution and high diversity challenges its successful application in ECG analysis.

Data augmentation is an effective approach to deal with the problem. It has become an inevitable processing in image-based deep learning applications, whereas its application in time-series physiological signals remains limited. We have summarized several data augmentation methods for deep learning-based ECG analysis and demonstrated that data augmentation can improve the performance of deep learning models greatly in detecting arrhythmia. In spite of these achievements, the currently available and effective data augmentation methods are still limited. In

particular, for beat-wise ECG analysis, data augmentation is not widely used because it seems that there are "enough" number of beats in the datasets. However, the diversity of the beats is relatively low as a large number of samples are from limited number of patients. Therefore, although the algorithms can obtain superior performance on the public dataset, their generalization ability is not satisfactory.

Recently, more researches lay their attention on GAN (Generative Adversarial Network), which is an unsupervised learning method by allowing two neural networks to contest with each other [52]. As the GAN model is capable of reproducing samples with realistic properties, researchers start to use it as a tool for data augmentation based on existing datasets, which has been proved feasible and potential on time series [53, 54].

In summary, more data augmentation approaches especially designed for ECG analysis are required to deal with the shortage of annotated data with high quality. With the help of more advanced data augmentation methods, we believe that deep learning may play a more significant role in ECG analysis.

Acknowledgments This study was supported by the National Natural Science Foundation of China (Grant 31870938) and Zhejiang Provincial Key Laboratory of Communication Networks and Applications.

References

1. De Bacquer, D., De Backer, G., Kornitzer, M., et al.: Prognostic value of ECG findings for total, cardiovascular disease, and coronary heart disease death in men and women. Heart. **80**, 570–577 (1998)
2. Juergens, K.U., Grude, M., Fallenberg, E.M., et al.: Using ECG-gated multidetector CT to evaluate global left ventricular myocardial function in patients with coronary artery disease. Am. J. Roentgenol. **179**, 1545–1550 (2002)
3. Auer, R., Bauer, D.C., Marques-Vidal, P., et al.: Association of major and minor ECG abnormalities with coronary heart disease events. JAMA. **307**, 1497–1505 (2012)
4. Billman, G.E.: Heart rate variability—a historical perspective. Front. Physiol. **2**, 86 (2011)
5. Odinaka, I., Lai, P.-H., Kaplan, A.D., et al.: ECG biometric recognition: a comparative analysis. IEEE Trans. Inf. Forensics Secur. **7**, 1812–1824 (2012)
6. Agrafioti, F., Hatzinakos, D., Anderson, A.K.: ECG pattern analysis for emotion detection. IEEE Trans. Affect. Comput. **3**, 102–115 (2011)
7. Jalaleddine, S.M., Hutchens, C.G., Strattan, R.D., et al.: ECG data compression techniques—a unified approach. IEEE Trans. Biomed. Eng. **37**, 329–343 (1990)
8. Pietka, E.: Feature extraction in computerized approach to the ECG analysis. Pattern Recogn. **24**, 139–146 (1991)
9. Pan, J., Tompkins, W.J.: A real-time QRS detection algorithm. IEEE Trans. Biomed. Eng. **32**, 230–236 (1985)
10. Li, C., Zheng, C., Tai, C.: Detection of ECG characteristic points using wavelet transforms. IEEE Trans. Biomed. Eng. **42**, 21–28 (1995)
11. Sahambi, J., Tandon, S., Bhatt, R.: Using wavelet transforms for ECG characterization. An on-line digital signal processing system. IEEE Eng. Med. Biol. Mag. **16**, 77–83 (1997)
12. Senhadji, L., Carrault, G., Bellanger, J., et al.: Comparing wavelet transforms for recognizing cardiac patterns. IEEE Eng. Med. Biol. Mag. **14**, 167–173 (1995)

13. Zhao, Q., Zhang, L.: ECG feature extraction and classification using wavelet transform and support vector machines. In: 2005 International Conference on Neural Networks and Brain, vol. 2, pp. 1089–1092 (2005)
14. Emanet, N.: ECG beat classification by using discrete wavelet transform and Random Forest algorithm. In: 2009 Fifth International Conference on Soft Computing, Computing with Words and Perceptions in System Analysis, Decision and Control, pp. 1–4 (2009)
15. Li, Q., Rajagopalan, C., Clifford, G.D.: Ventricular fibrillation and tachycardia classification using a machine learning approach. IEEE Trans. Biomed. Eng. **61**, 1607–1613 (2013)
16. LeCun, Y., Bengio, Y., Hinton, G.: Deep learning. Nature. **521**, 436 (2015)
17. Krizhevsky, A., Sutskever, I., Hinton, G.E.: Imagenet classification with deep convolutional neural networks. Adv. Neural Inf. Proces. Syst. **25**, 1097–1105 (2012)
18. Farabet, C., Couprie, C., Najman, L., et al.: Learning hierarchical features for scene labeling. IEEE Trans. Pattern Anal. Mach. Intell. **35**, 1915–1929 (2012)
19. Tompson, J.J., Jain, A., LeCun, Y., et al.: Joint training of a convolutional network and a graphical model for human pose estimation. Adv. Neural Inf. Proces. Syst. 1799–1807 (2014)
20. Mikolov, T., Deoras, A., Povey, D., et al.: Strategies for training large scale neural network language models. In: 2011 IEEE Workshop on Automatic Speech Recognition & Understanding, pp. 196–201 (2011)
21. Hinton, G., Deng, L., Yu, D., et al.: Deep neural networks for acoustic modeling in speech recognition. IEEE Signal Process. Mag. **29**, 82 (2012)
22. Sainath, T.N., Mohamed, A.-R., Kingsbury, B., et al.: Deep convolutional neural networks for LVCSR. In: 2013 IEEE International Conference on Acoustics, Speech and Signal Processing, pp. 8614–8618 (2013)
23. Hubel, D.H., Wiesel, T.N.: Receptive fields and functional architecture of monkey striate cortex. J. Physiol. **195**, 215–243 (1968)
24. Graves, A., Liwicki, M., Fernández, S., et al.: A novel connectionist system for unconstrained handwriting recognition. IEEE Trans. Pattern Anal. Mach. Intell. **31**, 855–868 (2008)
25. Sak, H., Senior, A., Beaufays, F.: Long short-term memory recurrent neural network architectures for large scale acoustic modeling. In: Fifteenth Annual Conference of the International Speech Communication Association (2014)
26. Li, X., Wu, X.: Constructing long short-term memory based deep recurrent neural networks for large vocabulary speech recognition. In: 2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 4520–4524 (2015)
27. Hochreiter, S., Schmidhuber, J.: Long short-term memory. Neural Comput. **9**, 1735–1780 (1997)
28. Hannun, A.Y., Rajpurkar, P., Haghpanahi, M.: Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network. Nat. Med. **25**, 65–69 (2019)
29. Übeyli, E.D.: Recurrent neural networks employing Lyapunov exponents for analysis of ECG signals. Expert Syst. Appl. **37**, 1192–1199 (2010)
30. Chauhan, S., Vig, L.: Anomaly detection in ECG time signals via deep long short-term memory networks. In: 2015 IEEE International Conference on Data Science and Advanced Analytics (DSAA), pp. 1–7 (2015)
31. Al Rahhal, M.M., Bazi, Y., AlHichri, H., et al.: Deep learning approach for active classification of electrocardiogram signals. Inf. Sci. **345**, 340–354 (2016)
32. Mousavi, S., Afghah, F., Razi, A., et al.: ECGNET: learning where to attend for detection of atrial fibrillation with deep visual attention. In: IEEE-EMBS International Conferences on Biomedical and Health Informatics (IEEE BHI) (2019)
33. Van Dyk, D.A., Meng, X.L.: The art of data augmentation. J. Comput. Graph. Stat. **10**, 1–50 (2001)
34. Zhao, A., Balakrishnan, G., Durand, F., et al.: Data augmentation using learned transformations for one-shot medical image segmentation. Proc. IEEE Conf. Comput. Vis. Pattern Recognit. 8543–8553 (2019)

35. Perez, L., Wang, J.: The effectiveness of data augmentation in image classification using deep learning. arXiv preprint arXiv:1712.04621 (2017)
36. Mikołajczyk, A., Grochowski, M.: Data augmentation for improving deep learning in image classification problem. In: 2018 International Interdisciplinary PhD Workshop (IIPhDW), pp. 117–122 (2018)
37. Fawzi, A., Samulowitz, H., Turaga, D., et al.: Adaptive data augmentation for image classification. In: 2016 IEEE International Conference on Image Processing (ICIP) pp. 3688–3692 (2016)
38. Cui, Z., Chen, W., Chen, Y.: Multi-scale convolutional neural networks for time series classification. arXiv preprint arXiv:1603.06995 (2016)
39. Fawaz, H.I., Forestier, G., Weber, J., et al.: Deep learning for time series classification: a review. Data Min. Knowl. Disc. **33**, 917–963 (2019)
40. Ramponi, G., Protopapas, P., Brambilla, M., et al.: T-CGAN: conditional generative adversarial network for data augmentation in noisy time series with irregular sampling. arXiv preprint arXiv:1811.08295 (2018)
41. Karim, F., Majumdar, S., Darabi, H., et al.: LSTM fully convolutional networks for time series classification. IEEE Access. **6**, 1662–1669 (2017)
42. Um, T.T., Pfister, F.M.J., Pichler, D., et al.: Data augmentation of wearable sensor data for Parkinson's disease monitoring using convolutional neural networks. arXiv preprint arXiv:1706.00527 (2017)
43. Cao, P., Li, X., Mao, K., et al.: A novel data augmentation method to enhance deep neural networks for detection of atrial fibrillation. Biomed. Signal Process. Control. **56**, 101675 (2020)
44. Yeh, J.-R., Sun, W.-Z., Shieh, J.-S., et al.: Intrinsic mode analysis of human heartbeat time series. Ann. Biomed. Eng. **38**, 1337–1344 (2010)
45. Le Guennec, A., Malinowski, S., Tavenard, R.: Data augmentation for time series classification using convolutional neural networks. In: ECML/PKDD Workshop on Advanced Analytics and Learning on Temporal Data (2016)
46. Acharya, U.R., Oh, S.L., Hagiwara, Y., et al.: A deep convolutional neural network model to classify heartbeats. Comput. Biol. Med. **89**, 389–396 (2017)
47. Jun, T.J., Nguyen, H.M., Kang, D., et al.: ECG arrhythmia classification using a 2-D convolutional neural network. arXiv preprint arXiv:1804.06812 (2018)
48. Clifford, G.D., Liu, C., Moody, B., et al.: AF classification from a short single lead ECG recording: the PhysioNet/computing in cardiology challenge 2017. Comput. Cardiol. (CinC). **2017**, 1–4 (2017)
49. Kingma, J.B.D.P.: Adam: a method for stochastic optimization. In: International Conference on Learning Representations, pp. 1–15 (2015)
50. Srivastava, N., Hinton, G., Krizhevsky, A., et al.: Dropout: a simple way to prevent neural networks from overfitting. J. Mach. Learn. Res. **15**, 1929–1958 (2014)
51. Gal, Y., Ghahramani, Z.: A theoretically grounded application of dropout in recurrent neural networks. Adv. Neural Inf. Proces. Syst. 1019–1027 (2016)
52. Goodfellow, I., Pouget-Abadie, J., Mirza, M., et al. Generative adversarial nets. Adv. Neural Inf. Proces. Syst. 2672–2680 (2014)
53. Nagasawa, T., Sato, T., Nambu, I., et al.: Improving fNIRS-BCI accuracy using GAN-based data augmentation. In: 2019 9th International IEEE/EMBS Conference on Neural Engineering (NER), pp. 1208–1211 (2019)
54. Nikolaidis, K., Kristiansen, S., Goebel, V., et al.: Augmenting physiological time series data: a case study for sleep apnea detection. arXiv preprint arXiv:1905.09068 (2019)

Study on Automatic Classification of Arrhythmias

Runnan He, Yang Liu, and Henggui Zhang

Abstract Electrocardiogram (ECG) signals reveal the electrical activity of the heart and can be used to diagnose heart abnormalities. In the past few decades, ECG signals have been utilized for automatic arrhythmia detection owing to the noninvasive nature and convenience of electrocardiography. However, it is difficult to extract and select reliable features or design robust and generic classifiers because of the complexity and diversity of ECG signals. Consequently, improving the classification rate of arrhythmias still remains a considerable challenge. To resolve this pressing issue, we have proposed a model composed of preprocessing, feature extraction, and classification, where the correct implementation of each part is crucial for final arrhythmia identification. In this chapter, the literature on existing algorithms is comprehensively reviewed according to the aforementioned primary aspects.

Keywords ECG · Arrhythmia · Classification · Machine learning

R. He (✉) · Y. Liu
School of Computer Science and Technology, Harbin Institute of Technology, Harbin, China
e-mail: runnanhe@sina.com

H. Zhang
School of Computer Science and Technology, Harbin Institute of Technology, Harbin, China

Biological Physics Group, Department of Physics and Astronomy, University of Manchester, Manchester, UK

Peng Cheng Laboratory, Shenzhen, China

Pilot National Laboratory of Marine Science and Technology, Qingdao, China

International Laboratory for Smart Systems and Key Laboratory of Intelligent of Computing in Medical Image, Ministry of Education, Northeastern University, Shenyang, China

C. Liu, J. Li (eds.), *Feature Engineering and Computational Intelligence in ECG Monitoring*, https://doi.org/10.1007/978-981-15-3824-7_7

1 Introduction

Cardiovascular diseases are the primary cause of death worldwide. Among them, cardiac arrhythmias are very common conditions that require new approaches for early diagnosis and more efficient treatments to prevent heart-related issues. Cardiac arrhythmia is the general name for several disorders where the heart rhythm is too slow, too fast, or irregular. Electrocardiography, which records the electrical activity of the heart, is the most widely applied technique for heart disease diagnosis [1]. A typical cardiac cycle from one electrocardiogram (ECG) signal is depicted in Fig. 1 [2]. As indicated in Fig. 1, each ECG beat displays morphological and temporal features of different components, such as the QRS complex, which supply useful clinical information that assists the automatic diagnosis of arrhythmias. However, automatic classification of arrhythmias is an enormous challenge owing to the diverse features of ECG signals from different patients under varying conditions [3]. Furthermore, each individual displays a particular ECG signal, and different patients may exhibit distinct ECG shapes for the same disease or similar ECG signals for different diseases. Consequently, it is necessary to analyze each heartbeat to detect arrhythmias. It is time consuming for cardiologists to diagnose many patients by analyzing long ECG signals in an efficient manner, and human errors due to fatigue can also occur during ECG analysis. Therefore, ECG signal interpretation can be improved by employing computer-aided algorithms for automatic arrhythmia detection. Cardiac arrhythmias are characterized by varying in waveform to generate some modes. Using these modes, it is possible to identify the type of arrhythmia. In general, the automatic classification of arrhythmias can be divided into four successive procedures. This chapter presents a comprehensive review of arrhythmia classification algorithms considering all of the key aspects, including preprocessing, feature extraction, feature selection, and classification.

The first step includes filtering of the ECG signal, data segmentation, and balancing. The techniques employed during the filtering directly impact the next

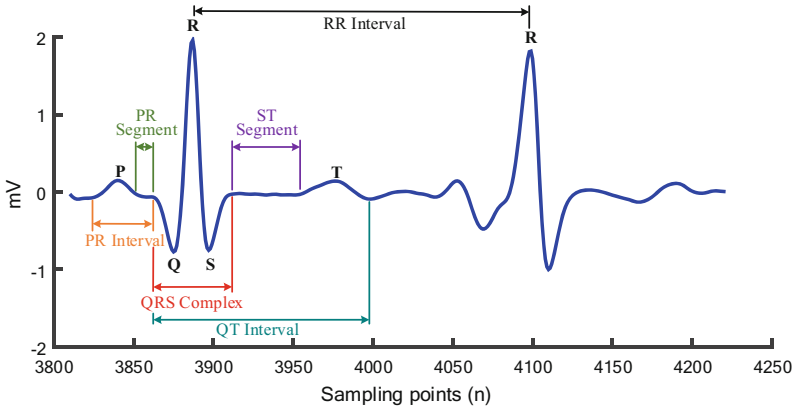

Fig. 1 A typical cardiac cycle from the MIT-BIH Arrhythmia Database [2]

step and should therefore be carefully selected. After denoising, segmentation is typically accomplished using the detected QRS complexes or a specific length depending on the particular requirements. Furthermore, balancing of the dataset is also important for the classification of arrhythmias. In practical applications, imbalanced data is a particularly common problem that requires attention during the construction of classifiers. Therefore, there is still considerable room for improvement in these aspects of classification.

Among the four key steps, the features employed in the final classification phase are vital for building a successful model. Thus, the feature extraction step directly influences the performance and accuracy of the overall sorting process. All of the information extracted from one heartbeat or a certain length can be applied to assign the appropriate category. The features can be extracted from the wave morphology of the ECG or higher-order statistical features obtained using various custom-written algorithms [4, 5]. Recently, deep neural networks (DNNs) have attracted considerable attention for feature extraction and have permitted great achievements in fields such as computer vision [6]. In contrast to conventional methods, DNNs can automatically extract features on the basis of the probability distribution of the dataset. With adequate training samples, the features extracted by a DNN model can be more comprehensive than those extracted by handcrafted methods. Consequently, numerous researchers have applied DNNs to arrhythmia classification over the last several years [7, 8]. Following feature extraction, it requires to employ techniques to reduce the feature dimensionality prior to feature selection [9]. The final step of classification is to identify arrhythmias from the extracted and selected feature sets, which is used to construct models. All of these procedures will be discussed in detail in this paper.

2 Preprocessing

Preprocessing is the first step and consists of filtering and data processing (such as segmentation and balancing). ECG signals are frequently contaminated by various types of noise, such as powerline interference, baseline wander, and muscle noise. Furthermore, imbalanced data is also a common issue that requires balancing by data segmentation to permit the final classification. These aspects are discussed in this section.

2.1 Filtering

The raw ECG signals are fed into a filter to eliminate noise originating from the electrical network, incorrect placement of the electrodes, or even bodily movement. Among the proposed filters, the simplest and most frequently used method is the recursive digital filter response to the finite impulse [10]. Although this method is

effective for reduction of known frequency bands such as powerline interference (50/60 Hz), the frequency bands of other types of noise are typically unknown. Although this problem can be overcome by applying multiple filters for various frequency bands, the casual application of filters such as high-pass and low-pass ones can distort the morphology of ECG signals, which limits the utility of this approach for diagnosing cardiac diseases. Therefore, linear adaptive filters have been proposed for noise removal from ECG signals [11, 12]. However, these filters are not clearly superior to digital filters [13]. In an effort to improve performance, researchers have utilized neural networks to construct adaptive filters, which resulted in significant noise attenuation [14]. In addition, some other methods such as quadratic filters have been applied for signal denoising while reinforcing the QRS complex for the subsequent segmentation [15, 16].

Over the past decade, wavelet transforms (WT) have been utilized in various fields. As a result, numerous filtering algorithms based on WTs have been proposed to solve the problem of denoising. Owing to their ability to preserve ECG signal properties and avoid the loss of significant physiological information, WT is effective for handling diverse types of noise [17–19]. Another study reported a modified WT referred to as the multi-adaptive bionic wavelet transform, which can be employed to remove noise and baseline variation from ECG signals. Compared with conventional WT, this transform exhibits superior performance. Besides wavelet denoising, other algorithms have also been exploited for noise attenuation. For example, nonlinear Bayesian filters have delivered promising results for the reduction of ECG signal noise [20]. Among such filters, the most widely used is the Kalman filter, which employs the parameters of an ECG dynamic model for noise reduction and signal compression, and the extended Kalman Filter in particular has made a huge contribution to data processing owing to its superior performance.

Although techniques for filtering ECG signals are highly developed, the specific choice of method is dependent on the final task. For dividing ECG signals into heartbeats, preliminary filtering may be necessary. In contrast, if the goal is the automatic classification of arrhythmias through deep learning, the noise is likely to improve the identification accuracy and may enhance model robustness.

2.2 Segmentation

In the segmentation stage, there are generally two ways to partition the signal. For heartbeat classification, the QRS complex (R peak) is typically used as the central point, and the signal is cut off a certain length before and after the R peak. After segmentation, several types of physiological information can be obtained by feature extraction, which is considered the foundation of automatic classification. Therefore, the accuracy of QRS complex detection is important as it impacts the following steps and ultimately the final classification.

Many different algorithms for detection of the QRS complex have been proposed over more than three decades [21, 22]. Among them, the Pan–Tompkins algorithm,

Table 1 The training set according to the "Frist label" annotations

Type	Number of recordings	Time length (s)				
		Mean	SD	Min	Median	Max
Normal	918	15.43	7.61	10.00	13.00	60.00
Atrial fibrillation (AF)	1098	15.01	8.39	9.00	11.00	60.00
First-degree atrioventricular block (I-AVB)	704	14.32	7.21	10.00	11.27	60.00
Left bundle branch block (LBBB)	207	14.92	8.09	9.00	12.00	60.00
Right bundle branch block (RBBB)	1695	14.42	7.60	10.00	11.19	60.00
Premature atrial contraction (PAC)	556	19.46	12.36	9.00	14.00	60.00
Premature ventricular contraction (PAC)	672	20.21	12.85	6.00	15.00	60.00
ST-segment depression (STD)	825	15.13	6.82	8.00	12.78	60.00
ST-segment elevated (STE)	202	17.15	10.72	10.00	11.89	60.00
Total	6877	15.79	9.04	6.00	12.00	60.00

Mean, SD, Min, Median, and Max in the header indicate the mean, standard deviation, minimum, median, and maximum of the time lengths of the recordings, respectively [31]

which is based on analysis of the slope of the QRS complex, is widely applied for heartbeat segmentation owing to its simplicity and effectiveness. This algorithm is also considered one of the gold-standard methods. In some similar algorithms, the first- and second-order derivatives are calculated, and a sliding window is then used to determine the location of the QRS complex. Other sophisticated algorithms such as WT and empirical mode decomposition (EMD) have also been investigated for QRS complex detection [23, 24]. Based on traditional machine learning methods, neural networks, support vector machine (SVM) [25], genetic algorithms [26], filter banks, and quad level vector [27] have also been utilized. For the comparison of these heartbeat segmentation algorithms, standard public databases must be employed. In the field of ECG signal processing, the most recommended and utilized database for the analysis of arrhythmias is the MIT-BIH Arrhythmia Database (MITDB) [28]. Other databases such as AHA [29] and CSE [30] have also been employed.

In the classification of arrhythmias, each ECG signal may have a different length that must be divided into fixed lengths, which are convenient for training the classification model. In this respect, the China Physiological Signal Challenge (CPSC) database has been used as an example [31]. As shown in Table 1, each recording is segmented into lengths of 30 s, which ensures the recording integrity while avoiding increasing the computational complexity by padding or truncating the recordings. If the recording length is less than 30 s, the recording is padded to 30 s by assigning zero values at the start. Alternatively, if the length exceeds 30 s, the extra data is truncated after 30 s. It is possible that the processes of padding and truncating may have some influence on the classifier performance; therefore, the

operations can be done in different positions, which may decrease the bias when training the classification model. The details will be discussed in the following subsection.

2.3 Data Augmentation and Balancing

After segmentation, augmentation and balancing of the dataset are also important for the classification of arrhythmias. Data imbalance occurs when the number of a certain class (termed the majority class) in the training set far exceeds the numbers of other classes (termed minority classes). To the best of our knowledge, most of the common datasets, such as the MITDB and CPSC databases, are imbalanced. In clinical diagnosis based on ECG recordings, treatment could be delayed if the minority class is abnormal samples. Furthermore, the majority of classification algorithms are developed under the assumption that the training set is balanced, which will tend to cause the learning algorithm to favor the majority class. This not only impedes convergence during the training phase but also reduces the generalization ability of the classifier. Therefore, it is crucial to address the data imbalance problem by focusing on the minority class. Numerous methods have been adopted to solve this class imbalance, as discussed in detail below [32, 33].

The first method is to modify the distribution of the training data by oversampling or undersampling. The most common approach to oversampling is to simply copy the minority class samples randomly, which is termed random oversampling. However, as this operation can easily lead to overfitting, more advanced techniques have been proposed, such as the synthetic minority oversampling technique (SMOTE) [34] and adaptive synthetic sampling approach for imbalanced learning (ADASYN) [35]. Furthermore, several other SMOTE algorithms have also been proposed in recent years [36–38]. In contrast, in undersampling, the majority samples are randomly removed such that their number is approximately equal to the numbers of other minority classes. However, careful attention must be paid to minimize the loss of useful information during undersampling.

The second method is to adjust the algorithm rather than changing the distribution of the training data. Examples of this approach reported in the literature include cost-sensitive learning and ensemble methods [39, 40]. Cost-sensitive learning can greatly improve classification performance. However, this method is only suitable if the costs of misclassification are known, and it can be very difficult or even impossible to confirm the cost of misclassification in certain areas [41]. Ensemble methods require considerable time to train multiple classifiers and can be impractical when DNNs are employed as the base classifiers. Therefore, new algorithms such as second-order cone programming SVM have been proposed [42]. Some researchers have also combined these two types of methods to develop mixed methods such as EasyEnsemble and SMOTEBoost, which have been demonstrated to more effectively deal with imbalanced data [43, 44].

Recently, generative adversarial networks (GANs) have demonstrated great promise for producing synthetic data [45]. GANs consist of two parts, namely, a generative model and a discriminative model, which compete with one another to obtain optimal outputs. When supplied with training data, the generator makes the noise vector as input and attempts to produce synthetic data similar to the training data. Then, the discriminator attempts to differentiate the training data from the synthetic data obtained from the generator. Therefore, given imbalanced training data, GANs can learn to estimate and change the distribution of the training data.

On the basis of traditional GAN models, a modified GAN named InfoGAN has also been proposed, which exerts a systematic and predictable influence on the outcome [46]. The use of synthetic samples generated by GANs has been gaining prominence with the results obtained in several works, such as those using the data augmentation generative adversarial network (DAGAN) approach, where it was empirically demonstrated that the data generated by DAGAN generally improved the accuracy [47]. In other fields, a deep convolutional GAN (DCGAN) was exploited for data augmentation to improve both liver lesion classification [48] and skin lesion segmentation [49]. An InfoGAN model was also previously employed for heartbeat anomaly detection to produce synthetic images to unbalanced classes [50].

3 Feature Extraction

ECG signals reflect the electrical activity of the heart, and the correct representation of these signals plays important roles in feature extraction and ultimately arrhythmia classification. Various feature extraction methods have been proposed in the literature to generate the unique information regarded as one feature. These features can be extracted in various forms, which can be applied individually or in combination. In this section, the types of features are divided into four categories, namely, waveform, statistical, wavelet, and other features.

3.1 Waveform Features

An ECG signal is mainly composed of the P wave, the QRS complex, and the T wave. Therefore, the most intuitive and fundamental features are the locations, durations, and shapes of these components. Excitation of the heart typically begins in the sinoatrial node and then propagates to the atria. The P wave is a weak, round, broad wave with low voltage amplitude that corresponds to the depolarization of the atria. This wave is an important feature for detecting AF. The QRS complex corresponds to the depolarization of the left and right ventricles. The Q wave is the first downward wave after the P wave. The R wave is the most prominent peak in an ECG signal and is important for the classification of arrhythmias. The S wave is a

downward deflection following the R wave and together with the Q and R waves makes up the QRS complex. Numerous studies have attempted to extract features from the QRS complex, which are calculated from the heartbeat interval referred to as the RR interval [51, 52]. The RR interval is the time between the R peaks of successive heartbeats and represents the duration of a single cardiac cycle. Changes in the RR interval can be used to calculate the heart rate variability, which is related to heart disease. Therefore, the RR interval has great value in arrhythmia classification. The T wave is a long wave with a round, broad peak that corresponds to the process of ventricular repolarization that follows the S wave. Although changes in the T wave are typically quite small, they can be important for diagnosing conditions such as myocardial infarction (MI). In addition to the original waveform characteristics, some researchers have employed other techniques to compress and reconstruct ECG signals, which represent new morphological features used to diagnose arrhythmias [53, 54]. For example, morphological features such as the Hermit function coefficient and temporal features have been extracted from ECG signals to classify heartbeats. Furthermore, the authors have used a random projection matrix to extract morphological features and combine dynamic features for heartbeat identification [55].

3.2 Statistical Features

In the last decade, statistical features have been employed for ECG classification. Among these features, the mean, standard deviation, and other statistical measures have been adopted in numerous studies [56, 57]. In general, these features provide an efficient strategy for analyzing the distribution and complexity of an ECG signal to better represent the temporal sequence. Moreover, higher-order statistics such as skewness and kurtosis are not particularly sensitive to changes in the ECG waveform. The potential change in the ECG signals can be examined through the owning dynamic and nonlinear properties [58]. Therefore, these features can help improve the ability to identify arrhythmias.

3.3 Wavelet Features

Wavelet features are another popular representation for ECG signals. Generally, ECG signals are nonstationary signals, making WT an efficient instrument for their analysis. Numerous studies have employed wavelet features for the classification of ECG signals [59, 60]. Compared with Fourier transforms, WT display strong time-frequency localization performance and rely on a linear transformation to decompose the ECG signals into different scales with a mother wavelet. Therefore, the choice of wavelet and number of signal decomposition levels are crucial for the analysis of ECG signals. The number of decomposition layers is determined on

the basis of the main frequency components. Then, according to the requirements, the corresponding frequency components of the ECG signals are reserved for classification of the wavelet coefficients. The initial operation of a WT is to divide ECG signals into low- and high-frequency components, which are considered approximate and detailed information. The decomposition of the next level is derived from the approximate component of the previous layer. Alternatively, both the approximate and detailed information can be decomposed at each layer to obtain more sub-band contents by using the wavelet packet. Various types of WT are available for extracting wavelet features. In addition to the traditional continuous wavelet transform (CWT), other transforms such as discrete wavelet transform, dual-tree complex WT, and flexible analytic WT have been employed to acquire other wavelet features.

3.4 Other Features

Several other techniques, such as independent component analysis (ICA) and principal component analysis (PCA), have been employed to reduce the dimensionality of features by extracting new coefficients to represent the ECG signals. The ICA technique can be used to statistically separate individual sources from a composite signal similar to the ECG signal, which consists of several action potential sources connected with different arrhythmia types. Therefore, the principle of the ICA for ECG signal classification is to separate the useful action potentials from the noise. The PCA technique is used to separate the sources from the signals based on their energy contributions. Research suggests that noise sources have low energy and are difficult to isolate and that the individual sources isolated by ICA are promising features for ECG classification. Moreover, it has been demonstrated that the combination of these two techniques can provide considerable advantages compared with the use of either technique alone [61]. Another proposed PCA algorithm is kernel principal component analysis, which is superior to PCA for arrhythmia detection owing to its nonlinear structure [62]. In addition to the two techniques discussed above, other feature extraction methods have also been proposed, such as Lyapunov exponents [63], Kolmogorov complexity, autocorrelation analysis, power spectral density [64], and genetic algorithms [65].

4 Feature Selection

Although some studies consider feature extraction and feature selection as two similar works, however, the two parts are in fact different. While feature extraction is regarded as the step to descript the information of ECG signals, feature selection is to use a series of methods for improving the performance of classifiers with the most

representative feature. In addition to enhance the performance of algorithm, it can also reduce the dimension of feature space that saves computational time.

4.1 Deterministic Wrappers

In terms of deterministic wrappers, the most commonly used algorithms are those based on sequential selection to estimate the significance of each feature for the final classification by adding or removing various numbers of features to or from the original feature set to achieve a better criterion value [66]. There are two approaches to sequential selection. The first is forward selection, which begins with an empty set and then adds each feature that maximizes the criterion function to the current set in a step-by-step manner until the predetermined number of features is reached. The second is sequential backward selection, which operates in the reverse direction, starting with the entire feature set and removing features to maximize the criterion function until the predetermined number of features is again reached. However, owing to the nesting effect, these two methods can only afford suboptimal results. Therefore, instead of a single feature, n features can be added to or removed from the current feature set at each step [67]. Another generalization of these two methods is plus-l takeaway-r (PTA), in which one feature is selected to add to the set at each step, and r features are then removed from the set. Although these methods reduce the nesting effect, they still do not provide optimal results.

The sequential floating selection algorithm is an improvement over PTA algorithms, which utilizes a flexible backtracking mechanism. Sequential floating forward selection starts with an empty feature set. After each forward step, backward steps are performed until the criterion function stops increasing, such that this method is considered a dynamic version of PTA [68]. Therefore, in each step of selection, different numbers of features can be added to or removed from the original set until an improved criterion value is attained. Sequential floating backward selection operates in the reverse direction. In an effort to further improve these approaches, several other strategies have also been proposed, such as forward and backward selection SVM with a Gaussian RBF kernel and k-nearest neighbor (KNN)-based sequential forward selection methods for feature selection from ECG signals.

4.2 Randomized Wrappers

Another popular feature selection approach is genetic algorithm (GA), which rely on a probabilistic search strategy inspired by the process of biological evolution [69]. For a given issue, GA can be applied to find the optimal solution in various potential ways. In this approach, each new generation is expected to provide a closer approximation to the optimal solution. In GA, the evolution process typically begins

with an initial population that is either randomly generated or manually defined. In each iteration, the population is regarded as a generation, in which the fitness of every individual is assessed. The fitness is the value of the objective function, which is typically defined as the classification accuracy of ECG signals with the chosen features. The much fitter individuals are stochastically selected from the current population, and the genome of every individual is modified via recombination and random mutation to produce the subsequent generation, which is then adopted in the next iteration until a maximum number of generations or satisfactory fitness level for the population is reached. The use of GA-based feature selection for ECG analysis has been reported in several studies [70, 71]. In addition to GAs, particle swarm optimization may also provide promising results for feature selection in the future [72].

5 Traditional Machine Learning Algorithms

After the features has been extracted and selected from the ECG signals, classification models can be constructed from these features by traditional machine learning algorithms. In the literature, there are various classifiers that have been utilized for ECG analysis and classification tasks. According to the recent researches, these methods can be mainly divided into artificial neural networks (ANNs), linear discriminant analysis (LDA), *k*-nearest neighbor (KNN), support vector machine (SVM), decision tree (DT), and Bayesian classifiers, which will be discussed in detail below.

5.1 Artificial Neural Networks (ANNs)

Inspired by biological neural networks, ANNs are composed of interconnected artificial neurons with adjustable weights. ANNs typically comprise three components, namely, input, hidden, and output layers, and have been extensively applied in an effort to address both linear and nonlinear classification issues using various network structures. Recently, several ANN architectures have been employed for arrhythmia classification, including probabilistic neural network (PNN) and multilayer perceptron (MLP). PNN is a feed-forward neural network that exhibit shorter training times than some other neural network models such as backpropagation network [73]. Furthermore, PNN are computationally more robust and efficient than traditional MLP [74] and have been exploited in numerous studies over recent years [75, 76]. Hybrid neuro-fuzzy networks have also been proposed to overcome the problems of MLP, such as their generalization performance and training time [77]. Radial basis function neural networks use radial basis functions as the activation functions to form a simple structure that can be trained effectively. Other models, such as recurrent ANNs, generalized regression neural networks, and

block-based neural networks, have also been applied in some ECG classification studies [78, 79]. Besides single models, some researchers have utilized combined neural networks to obtain more generic methods, which not only reduces the overall error in the neural networks but also decreases the incidence of false negatives [80].

5.2 *Linear Discriminant Analysis (LDA)*

LDA was originally proposed by Fisher and can achieve high classification accuracy compared with more complex algorithms [81]. LDA is a statistical approach based on discriminant functions, which are estimated from the training data and attempt to linearly separate the features by changing the weight vector and a bias, where the criteria applied to compute the weight vector is based on a model. The parameters are selected by maximum likelihood to provide the highest possible discrimination between different classes. LDA has been utilized in several recent ECG classification studies [82, 83].

5.3 k-*Nearest Neighbor (KNN)*

KNN is a widely used machine learning algorithm that has been exploited in numerous fields, including ECG classification [84, 85]. According to the labels of the closest training samples, the KNN category can be determined for an unknown class in the feature space. The closest training samples are selected by calculating the similarity of each feature vector using a distance measure, where the most common distance formula is that for Euclidean distance. Subsequently, the unknown samples can be confirmed to belong to the class of most of the closest k samples, which is analogous to a majority voting approach. In KNN, the choice of the k value is crucial as it strongly influences the accuracy of the ECG signal classification. However, the computational cost is relatively high. Therefore, some studies have combined clustering algorithms with neural networks to improve the robustness and reduce the training time [86, 87].

5.4 *Support Vector Machine (SVM)*

In SVM, the classifiers are constructed by using different hyperplanes in an n-dimensional space and then applied to distinguish different classes [88]. The hyperplanes are formed with the aid of points located at the edges of the support vectors. SVM is a popular method for solving binary classification problems owing to its excellent generalization ability, in which the main purpose is to search for a maximum margin between the training set and decision boundary. Support vectors

are the training samples closest to the decision boundary and employed for maximization of the margin. Depending on the type of kernel function, SVM can be utilized for both linear and nonlinear classification. A linear kernel function makes the SVM deal with a linear classification, whereas nonlinear kernel functions, such as Gaussian radial basis, polynomial, and sigmoid, make the SVM handle a nonlinear classification. In recent years, SVM has been utilized in numerous ECG classification studies [89, 90].

5.5 *Decision Trees (DTs)*

DTs are a decision support method that uses a tree-like model to determine and classify different categories utilizing only conditional control statements to execute the program. According to the difference purposes, the output results may be either a possible target class label or a target value, which are referred to as classification and regression trees, respectively. Recently, DTs have been employed for ECG signal classification [91, 92]. In addition to common DT approaches, some modified DT structures have also been reported, such as ensemble classifiers, which have also been employed for ECG signal classification. In the case of ensemble classifiers, multiple decision trees are trained with subsets of the training data, and the output class is a type of majority voting based on the maximum number of votes from all of the individual trees.

5.6 *Bayesian*

Bayesian classifiers are based on Bayesian decision theory, which is a fundamental statistical method. These classifiers have been widely used to solve pattern recognition problems in various studies [93, 94]. The rules of these classifiers can be divided into two types depending on whether the class is known. If the class is known, the values of the other features can be predicted. If the class is not known, Bayes's rule can be employed to predict the class label according to the given features. In Bayesian classifiers, probabilistic models of the features are constructed to predict the label of an unknown sample. Several types of Bayesian classifiers, including Bayesian networks such as naive Bayes, as well as the Bayes maximum-likelihood classifier, have been applied to ECG classification.

In addition to the aforementioned algorithms, other types of classifiers have also been employed for ECG classification, such as the template matching technique [95], Gaussian mixture model [96], and hidden Markov model. Furthermore, unsupervised learning algorithms such as *k*-means and self-organizing maps [97] have also been utilized.

6 Deep Learning Algorithms

Different from traditional algorithms that rely on hand-designed feature extraction methods, the deep learning algorithms can be used to train an end-to-end DNNs, where the feature extraction function is a hidden part of the model, which is optimized during the model training. Therefore, these methods can reduce the consumption of human efforts for feature extraction. Moreover, due to the auto-learned feature extraction functions avoid the subjectivity and limitation of human designed methods, the deep learning methods can usually achieve a better performance than traditional methods, especially when the dataset is big and comprehensive. In this section, we will discuss the deep learning methods that have been proposed for ECG classification.

6.1 Convolutional Neural Networks (CNNs)

CNNs are a type of DNNs that have found broad application in numerous fields, such as computer vision [98] and speech recognition [99]. A CNN is typically composed of a stack of convolutional layers and other types of assistant layers, such as max pooling layers, nonlinear activation layers, and batch normalization layers. There are numerous studies in the literature that have utilized CNNs for ECG classification.

Beat-level arrhythmia detection with the MITDB is a typical problem where CNNs have been widely adopted. According to the AAMI EC57 standard, heartbeats can be divided into five categories, namely, N (beats originating in the sinus mode), S (supraventricular ectopic beats), V (ventricular ectopic beats), F (fusion beats), and Q (unclassifiable beats) [100]. For classification, heartbeats should be first extracted from the ECG recordings. Typically, a heartbeat is cut out with specified lengths before and after the corresponding QRS complex. The extracted heartbeats can be processed directly using a 1D CNN, which usually consists of a series of 1D convolutional layers, pooling layers (between the convolutional layers), and fully connected layers [101]. For the multi-label classification of heartbeats, the output layer is typically a softmax layer whose output is the respective predicted probabilities of each class. As heartbeat segments are relatively short and have generally been downsampled to a short length (typically 100–200 sampling points), the CNNs used for the classification are usually not very deep, perhaps containing only two or three convolutional layers. To make the CNN parameters independent from the input layer dimension, a variant of CNNs, termed adaptive CNNs [102], has been proposed by extending the convolutional layers such that they are capable of both convolution and downsampling. Another innovation of this work was that the input contained not only the heartbeat to be classified but also its neighbors, which allowed the neural network to identify abnormal beats (e.g., premature atrial contractions) by comparing the current beat to its neighbors. For similar purposes, a dual-beat coupling

method was proposed to prepare for the network input. This method converts two pairs of adjacent beats into a dual-beat coupling matrix, which provides information concerning both the beat waveform and beat-to-beat correlation. Next, the dual-beat coupling matrix is input to a 2D CNN with three convolutional layers for classification. Another study proposed a 1D residual convolutional network for heartbeat classification [103]. This network consists of five residual blocks, each containing two convolutional layers, two ReLU activation layers, a residual skip connection, and a pooling layer. The output of the final residual block is then input into two fully connected layers (the first followed by a ReLU layer and the last followed by a softmax layer) to perform the final classification. This network contains 13 layers with trainable weights, which is much deeper than the networks reported in other studies.

In addition to the MITDB, the PTB is another commonly employed dataset that has been used to study beat-level MI identification methods. Numerous studies have utilized CNNs to address this problem. In an early study, a simple CNN consisting of three convolutional layers, three max pooling layers (each following a convolutional layer), and three fully connected layers was proposed for MI detection [104]. A previous study proposed an MI detector by reusing the design and weights of convolutional layers in the residual CNN that had been trained and evaluated on the MITDB. The final two fully connected layers in the network were retrained on the PTB dataset to adapt to this problem. For better feature extraction from 12-lead ECG recordings, a multiple-feature-branch convolutional neural network (MFB-CNN) [105] was proposed to extract the features of each ECG lead using separate CNN subnetworks (i.e., feature branches). The outputs of the feature branches are then summarized and classified by fully connected softmax layers. As an extension of the MFB-CNN method, the multilead-CNN (ML-CNN) [106] leverages 2D convolutional layers with a 1D kernel to enable parameter sharing among leads, which helps to reduce the number of parameters and avoid overfitting. Furthermore, an extension to the pooling layer, termed a lead asymmetric pooling layer, was also introduced to capture the multiscale features of the different leads [106].

Although the detection of ECG abnormalities at the beat level is relatively easy to handle, the limitation is also obvious. The features of some diseases, such as AF, may manifest themselves not only in the heartbeat waveform but also in the pattern of beat-to-beat variability. Therefore, many researchers have attempted to address the problem of arbitrary ECG segment classification, where the length of ECG segments can be very variable, typically ranging from several seconds to several minutes. The classification of ECG segments is more in line with the needs of clinical practice and has thus attracted the attention of an increasing number of researchers. As the input is substantially longer than that of beat-level classifiers, the networks designed to address this problem are usually much deeper. For example, a 34-layer residual-based CNN was proposed to detect up to 12 types of rhythms using single-lead ECGs. These researchers assembled a very large dataset containing 91,232 ECG recordings obtained from 53,549 patients using single-lead ambulatory ECG monitoring devices. In terms of performance, the average F_1 score of the DNN

model was reported to exceed that of an average cardiologist. In addition, when the specificity was set to the cardiologist level, the DNN displayed a higher sensitivity than the average cardiologist for all rhythm classes, which demonstrates the power of DNN for ECG classification. However, as the dataset used in this study is not publicly available, most of the research in this field has been conducted using other publicly accessible datasets, such as the AFDB and CinC/PhysioNet Challenge 2017 (CinC2017DB) dataset. In one study, the authors created two types of CNNs (one with two convolutional layers and the other with three convolutional layers) using the AFDB to detect AF [107]. The inputs for these CNNs were 5-s ECG segments that had been transformed via short-term Fourier transform and stationary wavelet transform, respectively. In addition, a four-layer CNN was proposed to detect AF using five contiguous heartbeats that had been separately subjected to CWT [7].

The CinC2017DB contains 12,186 single-lead ECG recordings for AF detection and has attracted considerable research attention [108], with many researchers creating CNN-based AF detectors using this dataset. To extract features from signals lasting up to tens of seconds, networks with a deeper structure are typically required, and consequently residual convolutional networks have proved popular [109]. For example, one proposed network was based on a previous network design with variable network depth and number of filters at each layer. In another study, the authors proposed a simple residual convolutional network consisting of 16 convolutional layers to address this problem. As another effective CNN, a densely connected convolutional network (DenseNet) was also applied to detect AF from ECGs obtained from this dataset. Another proposed network consisted of three dense blocks using 15-s ECG segments as the input [110]. Some other works have combined CNNs and handcrafted features to improve the generalization ability of the classifiers. For example, in one study the authors extracted the features using a statistical description of the signal average beats and CNN, and trained an ensemble classifier of a neural network and bagged tree for AF detection [111]. In another study, features to characterize the variability of the RR interval and existence of P waves were extracted via custom methods and then combined with CNN-extracted features for the final classification [112].

6.2 Recurrent Neural Networks (RNNs)

RNNs were designed for time series processing and have demonstrated outstanding performance in fields such as automatic speech recognition and natural language processing. As ECG signals are time series in nature, it is quite natural to expect that RNNs could be used for ECG processing. For beat-level ECG classification, a type of RNN termed the long short-term memory (LSTM) network was applied to classify beats from the MITDB [113]. In another study, the beat waveforms were first converted into frequency sub-band sequences using a wavelet-based layer and then fed into a bidirectional LSTM-based neural network for feature learning and

classification [114]. With respect to arbitrary ECG segment classification, an ensemble of RNNs was created to detect AF using recordings from CinC2017DB [115]. The input for this network was the sequence of features extracted from each heartbeat rather than the raw ECG recordings, which significantly reduces the number of steps that the RNNs must process. Similarly, in another study, each ECG recording was split into a sequence of beats, and each beat was then processed using an MLP for feature extraction [116]. The extracted beat feature sequence was then processed using an LSTM-based neural network to predict the sequence rhythm. Furthermore, a handcrafted feature-based classifier was also constructed and coupled with the LSTM-based DNN to afford an ensemble classifier, which achieved the first place in the challenge.

6.3 Combining CNNs and RNNs

As one ECG segment may contain up to thousands of sampling points, it is neither effective nor efficient to process raw ECG signals using only an RNN model. In addition, CNNs are considered effective for extracting features from the raw waveforms. Thus, the combination of CNNs and RNNs is expected to improve the feature learning and classification of ECG signals, and numerous such studies have been reported. For example, a model combining a CNN and LSTM was constructed for the automatic classification of cardiac arrhythmias using data from CinC2017DB [117]. The CNN was used to learn a high-level representation (a sequence of feature vectors) of the ECGs, while the LSTM further summarized the feature sequence learned by the CNN for classification. In another study, the performance of a model combining a CNN and LSTM was compared with that of a pure CNN model via fivefold cross validation, which demonstrated the efficiency of this combination [118]. In a separate study, two independent CNN models were used to extract features from ECG and heart rate signals, and then the features were merged into an RNN for aggregation [119]. In addition, a deep residual convolutional network with recurrent layers was applied to AF detection using 5-s ECG segments [120]. In another study, the authors proposed a novel framework that combined a CNN, bidirectional LSTM, and an attention mechanism for the detection of paroxysmal atrial fibrillation (PAF) [121]. The input for this network was a sequence of 30-s ECG segments, each of which was first converted to a time-frequency representation via wavelet transform and then fed into a CNN subnet for feature extraction. Then, a bidirectional RNN with an attention layer was used to aggregate the extracted feature sequences for AF detection. Finally, in another study, the authors separately extracted ECG features using an RNN, a combination of a CNN and an RNN, and handcrafted methods, and then used an XGBoost classifier to perform the final classification based on the union of all of these features [122].

6.4 Unsupervised Deep Models

Unsupervised deep learning methods have also been applied to ECG classification, especially for pretraining of the model parameters. For example, heartbeat classification was accomplished by gradually training a stack of restricted Boltzmann machine (RBM) layers for feature learning. The trained RBM layers were then combined to form a deep brief network with a softmax layer for classification. In another study, a stacked denoising auto-encoder (SDAE) was proposed to extract features from the time-frequency spectra of heartbeats that were computed using a modified frequency slice wavelet transform. The classifier was then built by concatenating an encoder layer of the SDAE and a softmax layer. A modified denoising auto-encoder was also reported using the original heartbeat as the input, after which the output was compared with the denoised heartbeat for loss computing [123]. The features extracted by the last encoder layer were fed into a feed-forward neural network to classify the heartbeats into 16 categories, and comparable performance with previous methods based on feature engineering was achieved. In another study, a stacked sparse auto-encoder was proposed for ECG feature learning, in which the extracted features were further handled by a softmax regression model for heartbeat classification [124]. A convolutional auto-encoder was also utilized to extract features from heartbeats, and the features were then fed into an LSTM network for classification of the beats into five types of rhythms [125]. With respect to unsupervised learning using arbitrary ECG segments, the authors of one study proposed an encoder-decoder-based genitive model for limited-channel ECG classification [126]. This network was trained to generate the missing channel information, and then the latent representation could be used as features to perform other supervisory tasks. The model was compared with standard RNNs with respect to disease prediction, which demonstrated the effectiveness of the proposed method.

7 Proposed Deep Learning Algorithm

We have proposed a novel neural network by combining a residual CNN and bidirectional LSTM to discriminate nine different cardiac conditions (including one normal condition and eight abnormalities) using the 12-lead ECG recordings from the CPSC2018 dataset [127]. As the length of these recordings is variable (ranging from 6 to 60 s), we converted them to the same length by truncation (for longer recordings) or padding with zeros (for shorter recordings) to enhance the batch processing during model training. The selected target length was 30 s (15,000 sampling points), which is longer than most recordings in the dataset, to obtain a balance between the information integrity of the ECG recordings and the computational complexity of the DNN model. Then, the ECGs were processed using our proposed DNN for feature extraction and classification. As the input sequences are very long, it is challenging to extract features from them. To overcome this

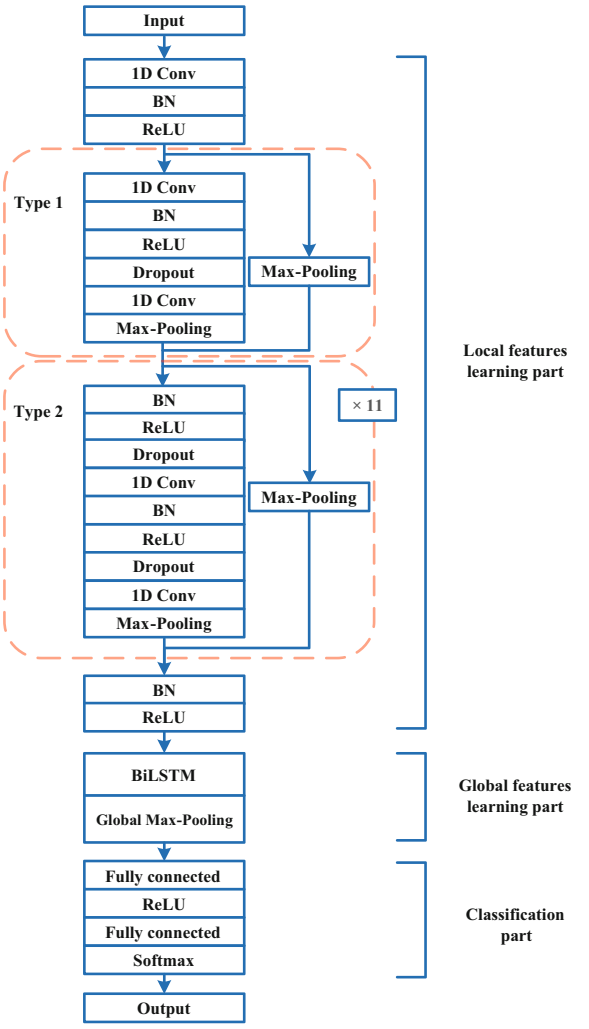

Fig. 2 Structure of the proposed DNN for 12-lead ECG classification [127]

challenge, we constructed our DNN with three parts, namely, a local features learning part, a global features learning part, and a classification part, as depicted in Fig. 2.

The local features learning part is a 1D residual CNN that consists of 11 residual blocks of two types, which was adapted from a previous work. Both types are composed of 1D convolutional (1D Conv) layers, batch normalization (BN) layers, ReLU layers, dropout layers, and max pooling layers. The difference is that type 2 contains three more layers (including a BN layer, a ReLU layer, and a dropout layer) than type 1 at the beginning of the block. Counting the independent convolutional layer at the very beginning, there are a total of 25 convolutional layers and 12 max pooling layers in this part. The initial convolutional layer contains

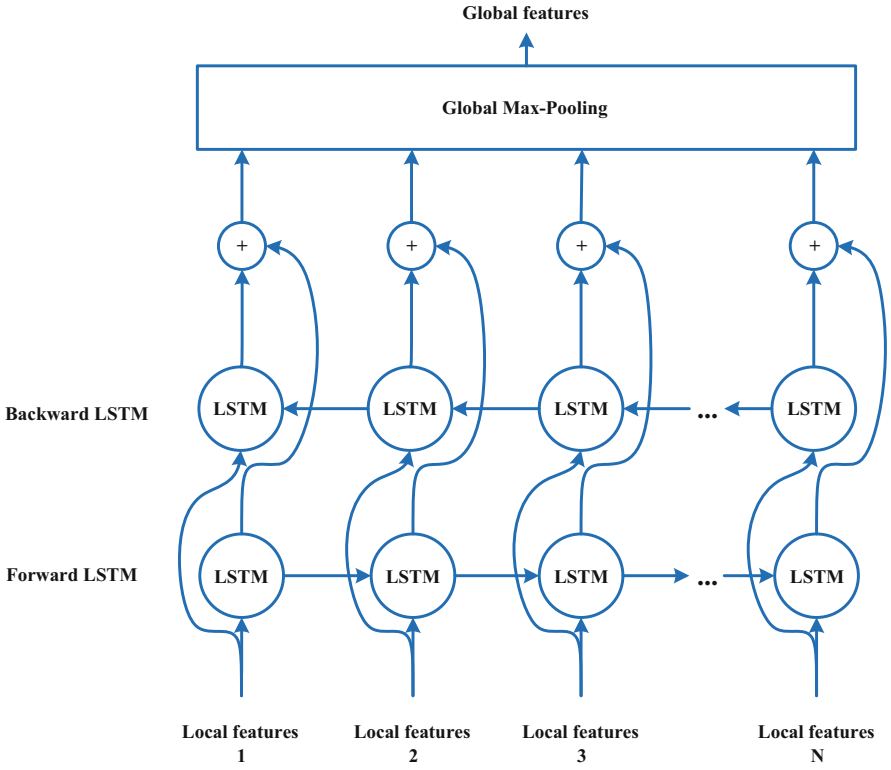

Fig. 3 Structure unfolded along the time sequence of the global features learning part [127]

32 filters with a kernel size of 16. For the deeper layers, the filter number increases by 32 every four blocks, while the kernel size is reduced by half every six blocks. In our design, the length of a feature map remains unchanged when going through a convolutional layer but is compressed by half when going through a max pooling layer that has a pool size of 2. However, only one of every two adjacent max pooling layers has a pool size of 2, whereas the other has a pool size of 1. Therefore, the sequence length is compressed by six times to 1/64 of the original after processing by the local features learning part. The compressed feature maps are then fed into the global features learning part for further processing.

The global features learning part is composed of two layers, namely, a bidirectional LSTM layer and a global max pooling layer, as depicted in Fig. 3. Owing to the advantageous gate designs of this unit, LSTM can remember the properties of a longer temporal up to hundreds of time steps in its internal state. The bidirectional LSTM is composed of two LSTM layers, termed the forward LSTM and the backward LSTM, which process the sequences in opposite directions. Each LSTM layer contains 64 units, which means that the output at each step is also a vector with 64 elements. The outputs of the two LSTM layers are summed into a local-focused global feature vector containing 64 elements, which encapsulates features from the

Table 2 Results of the arrhythmia classification based on 450 test samples [127]

Rank	Team	Normal	AF	I-AVB	LBBB	RBBB	PAC	PVC	STD	STE
1	**He et al.**	**0.748**	**0.920**	**0.882**	**0.889**	**0.883**	**0.787**	**0.851**	**0.780**	**0.780**
2	Cai et al.	0.765	0.927	0.887	0.886	0.880	0.812	0.800	0.784	0.753
3	Chen et al.	0.752	0.930	0.871	0.915	0.839	0.832	0.833	0.800	0.667
4	Mao et al.	0.692	0.940	0.852	1.000	0.899	0.706	0.875	0.762	0.622
5	Yu et al.	0.709	0.907	0.863	0.918	0.838	0.736	0.783	0.714	0.723
6	Zhang et al.	0.757	0.904	0.839	0.887	0.787	0.735	0.755	0.742	0.638

Table 3 The F_1 scores of four sub-abnormal types based on 450 test samples [127]

Rank	Team	F_{af}	F_{block}	F_{pc}	F_{st}	F_1
1	**He et al.**	**0.920**	**0.884**	**0.821**	**0.780**	**0.836**
2	Cai et al.	0.927	0.884	0.806	0.770	0.833
3	Chen et al.	0.930	0.868	0.832	0.738	0.827
4	Mao et al.	0.940	0.898	0.800	0.704	0.816
5	Yu et al.	0.907	0.871	0.758	0.719	0.799
6	Zhang et al.	0.904	0.833	0.745	0.699	0.783

context of the current step in both the forward and backward directions. The outputs of the two LSTM layers are summed in a stepwise manner and then fed into a global max pooling layer, which aggregates the feature sequence into a single feature vector.

In the classification part, the global feature vector is processed by two fully connected (Dense) layers, where the first is followed by a ReLU layer and the second is followed by a softmax layer. The first fully connected layer contains 64 cells, whereas the second contains nine cells (corresponding to nine classes). The output of the softmax layer is thus the predicted probability distribution over the nine classes.

We evaluated our proposed model using the hidden test dataset containing about 3000 ECG recordings. The performance of our proposed model was measured by the F_1 scores of nine classes, the F_1 scores of four sub-normal types, and the overall F_1 score which is the average of the 9 class-wised scores. The four sub-normal types include AF, Block (including I-AVB, LBBB, and RBBB), PC (including PAC and PVC), and ST (including STD and STE). Through the competition process of the CPSC, we got two groups of scores, one (shown in Tables 2 and 3) was computed by using a balance subset (containing 450 recordings, 50 for each class) of the test set, and the other (shown in Table 4) was computed by using the whole test set.

Table 4 The F_1 scores of four sub-abnormal types based on the entire test set [127]

Rank	Team	F_{af}	F_{block}	F_{pc}	F_{st}	F_1
1	Chen et al.	0.933	0.899	0.847	0.779	0.837
2	Cai et al.	0.931	0.912	0.817	0.761	0.830
3	**He et al.**	**0.914**	**0.879**	**0.801**	**0.742**	**0.806**
4	Yu et al.	0.918	0.890	0.789	0.718	0.802
5	Yan et al.	0.924	0.882	0.779	0.709	0.791
6	Chen et al.	0.905	0.902	0.722	0.708	0.783

As shown in Table 2, the classifier achieves a good performance on a small and balanced dataset. The scores for the eight abnormalities are all close to or greater than 0.78, and the highest one (0.920) is the score for AF detection. The lowest score (0.748) is achieved on the identification of normal. The F_1 scores of four sub-normal classes were presented in Table 3. The achieved F_1 scores of AF, Block, PC, and ST are 0.920, 0.884, 0.821, and 0.78, respectively. From the results, we can see that the proposed model has a high-level capability of identifying AF, block, and premature contraction. The score for detection of ST-segment changes is slightly lower, possibly because the ST-segment changes in some cases are too subtle to be detected by a CNN network.

Table 4 presents the F_1 scores of four sub-abnormal types on the entire test set, which is slightly lower than that achieved on the dataset with 450 samples. The possible reason is that we balanced the training set during the training process of our model making the data distribution different from that of the test set. The balance operation is helpful to the feature learning of the network, but it can also cause a bias to the classifier, which needs to be improved in future studies. Nevertheless, the overall score (0.806) is still higher than that of most competitors, which ranked the third place in the CPSC2018 challenge.

8 Discussions

During the history of automatic ECG analysis, lots of algorithms have been proposed to address a series of problems in this field. Though many of these algorithms have been used successfully in some scenarios, many problems in the workflow of ECG processing, e.g., noise removal, feature extraction and selection, remain to be solved. As the development of mobile computing and communication technologies, smaller and cheaper portable ECG monitors are becoming a trend in the industry, and meantime bringing new challenges to the community of ECG analysis due to the lower quality and bigger volume of their generated data.

In recent years, deep learning has exceeded the performance of traditional pattern recognition algorithms in many aspects, making it a very promising technology. Compared to the traditional algorithms, a DNN model can extract features from the raw data by itself rather than relying on handcrafted features. We have demonstrated

its effectiveness to the problems of ECG classification and arrhythmias detection by experiments. Besides, a DNN model is generally supposed to have a better accuracy and robustness when it is trained on a much larger dataset. We have noticed that many entertainments built their private annotated ECG datasets that are much larger than the public available datasets. Open-access large databases, such as CPSC2018, have prompted the researches of deep learning technology in the field of ECG analysis. In the future, larger open-access ECG datasets will benefit the research community greatly, just like what happened in the field of computer vision and natural language processing. Furthermore, as new technologies, e.g., new components and architectures, keep emerging, deep learning is in a stage of rapid development and will continue to promote progress in the field of ECG analysis.

References

1. Alberdi, A., Aztiria, A., Basarab, A.: Towards an automatic early stress recognition system for office environments based on multimodal measurements: a review. J. Biomed. Inform. **59**, 49–75 (2016)
2. He, R., Wang, K., Li, Q., Yuan, Y., Zhao, N., Liu, Y., Zhang, H.: A novel method for the detection of R-peaks in ECG based on k-nearest neighbors and particle swarm optimization. EURASIP J. Adv. Signal Process. **82**, 1–14 (2017)
3. Banerjee, S., Mitra, M.: Application of Cross Wavelet Transform for ECG pattern analysis and classification. IEEE Trans. Instrum. Meas. **63**, 326–333 (2014)
4. De Lannoy, G., François, D., Delbeke, J., Verleysen, M.: Weighted conditional random fields for supervised interpatient heartbeat classification. IEEE Trans. Biomed. Eng. **59**, 241–247 (2012)
5. De Chazal, P., O'Dwyer, M., Reilly, R.B.: Automatic classification of heartbeats using ECG morphology and heartbeat interval features. IEEE Trans Biomed Eng. **51**, 1196–1206 (2004)
6. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 770–778 (2016)
7. He, R., Wang, K., Zhao, N., Liu, Y., Yuan, Y., Li, Q., Zhang, H.: Automatic detection of atrial fibrillation based on continuous wavelet transform and 2d convolutional neural networks. Front. Physiol. **9**, 1206 (2018)
8. Hannun, A.Y., Rajpurkar, P., Haghpanahi, M., Tison, G.H., Bourn, C., Turakhia, M.P., Ng, A. Y.: Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network. Nat. Med. **25**, 65–69 (2019)
9. Llamedo, M., Martinez, J.P.: Heartbeat classification using feature selection driven by database generalization criteria. IEEE Trans. Biomed. Eng. **58**, 616–625 (2011)
10. Lynn, P.: Recursive digital filters for biological signals. Med. Biol. Eng. Comput. **9**, 37–43 (1979)
11. Ferrara, E.R., Widraw, B.: Fetal electrocardiogram enhancement by time-sequenced adaptive filtering. I.E.E.E. Trans. Biomed. Eng. **29**, 458–460 (1982)
12. Yelderman, M., Widrow, B., Cioffi, J.M., Hesler, E., Leddy, J.A.: ECG enhancement by adaptive cancellation of electrosurgical interference. I.E.E.E. Trans. Biomed. Eng. **30**, 392–398 (1983)
13. Thakor, N.V., Zhu, Y.-S.: Applications of adaptive filtering to ECG analysis: noise cancellation and arrhythmia detection. I.E.E.E. Trans. Biomed. Eng. **38**, 785–794 (1991)
14. Xue, Q., Hu, Y.H., Tompkins, W.J.: Neural-network-based adaptive matched filtering for QRS detection. I.E.E.E. Trans. Biomed. Eng. **39**, 317–329 (1992)

15. Phukpattaranont, P.: QRS detection algorithm based on the quadratic filter. Expert Syst. Appl. **42**, 4867–4877 (2015)
16. Elgendi, M.: Fast QRS detection with an optimized knowledge-based method: evaluation on 11 standard ECG databases. PLoS One. **8**, 1–18 (2013)
17. Singh, B.N., Tiwari, A.K.: Optimal selection of wavelet basis function applied to ECG signal denoising. Digital Signal Process. **16**, 275–287 (2006)
18. Chen, S.-W., Chen, H.-C., Chan, H.-L.: A real-time QRS detection method based on moving-averaging incorporating with wavelet denoising. Comput. Methods Programs Biomed. **82**, 187–195 (2006)
19. Zadeh, A.E., Khazaee, A., Ranaee, V.: Classification of the electrocardiogram signals using supervised classifiers and efficient features. Comput. Methods Programs Biomed. **99**, 179–194 (2010)
20. Sameni, R., Shamsollahi, M.B., Jutten, C., Clifford, G.D.: A nonlinear Bayesian filtering framework for ECG denoising. I.E.E.E. Trans. Biomed. Eng. **54**, 2172–2185 (2007)
21. Pan, J., Tompkins, W.J.: A real-time QRS detection algorithm. I.E.E.E. Trans. Biomed. Eng. **32**, 230–236 (1985)
22. Arzeno, N.M., Deng, Z.-D., Poon, C.-S.: Analysis of first-derivative based QRS detection algorithms. I.E.E.E. Trans. Biomed. Eng. **55**, 478–484 (2008)
23. Merah, M., Abdelmalik, T.A., Larbi, B.H.: R-peaks detection based on stationary wavelet transform. Comput. Methods Prog. Biomed. **121**, 149–160 (2015)
24. Berwal, D., Kumar, A., Kumar, Y.: Design of high performance QRS complex detector for wearable healthcare devices using biorthogonal spline wavelet transform. ISA Trans. **81**, 222–230 (2018)
25. Mehta, S.S., Lingayat, N.: A combined entropy-based method for detection of QRS complexes in 12-lead electrocardiogram using SVM. Comput. Biol. Med. **38**, 138–145 (2008)
26. Poli, R., Cagnoni, S., Valli, G.: Genetic design of optimum linear and nonlinear QRS detectors. I.E.E.E. Trans. Biomed. Eng. **42**, 1137–1141 (1995)
27. Kim, H., Yazicioglu, R.F., Merken, P., van Hoof, C., Yoo, H.-J.: ECG signal compression and classification algorithm with quad level vector for ECG holter system. IEEE Trans. Inf. Technol. Biomed. **14**, 93–100 (2010)
28. Moody, G.B., Mark, R.G.: The MIT-BIH arrhythmia database on CD-ROM and software for use with it. In: 1990 Proceedings Computers in Cardiology, pp. 185–188 (1990)
29. American Heart Association: AHA database. http://www.ahadata.com/ (1998)
30. Van Bemmel, J.H., Williams, J.L.: Standardisation and validation of medical decision support systems: the CSE project. Methods Inf. Med. **29**, 261–262 (1990)
31. Liu, F., Liu, C., Zhao, L., Zhang, X., Wu, X., Xu, X., Liu, Y., Ma, C., Wei, S., He, Z., Li, J., Ng, E.Y.: An open access database for evaluating the algorithms of electrocardiogram rhythm and morphology abnormality detection. J. Med. Imaging Health Inform. **8**, 1368–1373 (2018)
32. Leevy, J.L., Khoshgoftaar, T.M., Bauder, R.A., Seliya, N.: A survey on addressing high-class imbalance in big data. J. Big Data. **5**, 1–30 (2018)
33. Guo, H., Li, Y., Shang, J., Gu, M., Huang, Y., Gong, B.: Learning from class imbalanced data: review of methods and applications. Expert Syst. Appl. **73**, 220–239 (2017)
34. Chawla, N., Bowyer, K., Hall, L.O., Philip Kegelmeyer, W.: Smote: synthetic minority over-sampling technique. J. Artif. Intell. Res. **16**, 321–357 (2002)
35. He, H., Bai, Y., Garcia, E.A., Li, S.: Adasyn: adaptive synthetic sampling approach for imbalanced learning. In: 2008 IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence), pp. 1322–1328 (2008)
36. Guo, H., Viktor, H.L.: Learning from imbalanced data sets with boosting and data generation: the DataBoost-IM approach. ACM SIGKDD Explor. Newslett. **6**, 30–39 (2004)
37. Maldonado, S., López, J., Vairetti, C.: An alternative SMOTE oversampling strategy for high-dimensional datasets. Appl. Soft Comput. **76**, 380–389 (2019)
38. Jiang, J., Zhang, H., Pi, D., Dai, C.: A novel multi-module neural network system for imbalanced heartbeats classification. Expert Syst. Appl. X. **1**, 100003 (2019)

39. Sun, Y., Kamel, M.S., Wong, A.K.C., Wang, Y.: Cost-sensitive boosting for classification of imbalanced data. Pattern Recogn. **40**, 3358–3378 (2007)
40. Zhou, Z.H., Liu, X.Y.: Training cost-sensitive neural networks with methods addressing the class imbalance problem. IEEE Trans. Knowl. Data Eng. **18**, 63–77 (2006)
41. Wang, S., Liu, W., Wu, J., Cao, L., Meng, Q., Kennedy, P.J.: Training deep neural networks on imbalanced data sets. In: International Joint Conference on Neural Networks, pp. 4368–4374 (2016)
42. Maldonado, S., López, J.: Imbalanced data classification using second-order cone programming support vector machines. Pattern Recogn. **47**, 2070–2079 (2014)
43. Liu, X.Y., Wu, J., Zhou, Z.H.: Exploratory undersampling for class imbalance learning. IEEE Trans. Syst. Man Cybern. B. **39**, 539–550 (2009)
44. Havaei, M., Davy, A., Wardefarley, D., Biard, A., Courville, A., Bengio, Y., Pal, C., Jodoin, P., Larochelle, H.: Brain tumor segmentation with deep neural networks. Med. Image Anal. **35**, 18–31 (2017)
45. Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y.: Generative adversarial nets. In: Ghahramani, Z., Welling, M., Cortes, C., Lawrence, N.D., Weinberger, K.Q. (eds.) Advances in Neural Information Processing Systems 27, pp. 2672–2680 (2014)
46. Chen, X., Duan, Y., Houthooft, R., Schulman, J., Sutskever, I., Abbeel, P.: Infogan: interpretable representation learning by information maximizing generative adversarial nets. In: Advances in Neural Information Processing Systems, pp. 1–14 (2016)
47. Antoniou, A., Storkey, A., Edwards, H.: Data augmentation generative adversarial networks. arXiv:1711.04340v2 (2018)
48. Frid-Adar, M., Klang, E., Amitai, M., Goldberger, J., Greenspan, H.: Synthetic data augmentation using gan for improved liver lesion classification. In: 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), pp. 289–293 (2018)
49. Bolelli, F., Pollastri, F., Palacios, R.P., Grana, C.: Improving skin lesion segmentation with generative adversarial networks. In: 2018 IEEE 31st International Symposium on Computer-Based Medical Systems (CBMS), pp. 442–443 (2018)
50. Lima, J.L., Macêdo, D., Zanchettin, C.: Heartbeat anomaly detection using adversarial oversampling. arXiv preprint arXiv:1901.09972 (2019)
51. Korürek, M., Doğan, B.: ECG beat classification using particle swarm optimization and radial basis function neural network. Expert Syst. Appl. 37, 7563–7569 (2010)
52. Kumar, R.G., Kumaraswamy, Y.S.: Investigation and classification of ECG beat using input output additional weighted feed forward neural network. In: International Conference on Signal Processing, Image Processing & Pattern Recognition (ICSIPR), pp. 200–205 (2013)
53. Ye, C., Kumar, B.V.K.V., Coimbra, M.T.: Heartbeat classification using morphological and dynamic features of ECG signals. I.E.E.E. Trans. Biomed. Eng. **59**, 2930–2941 (2012)
54. Chen, S.S., Hua, W., Li, Z., Li, J., Gao, X.J.: Heartbeat classification using projected and dynamic features of ECG signal. Biomed. Signal Process. **31**, 165–173 (2017)
55. Sharma, P., Ray, K.C.: Efficient methodology for electrocardiogram beat classification. IET Signal Process. **10**, 825–832 (2016)
56. Song, C.Y., Liu, K.B., Zhang, X., Chen, L.L., Xian, X.C.: An obstructive sleep apnea detection approach using a discriminative hidden Markov model from ECG signals. I.E.E.E. Trans. Biomed. Eng. **63**, 1532–1542 (2016)
57. Afkhami, R.G., Azarnia, G., Tinati, M.A.: Cardiac arrhythmia classification using statistical and mixture modeling features of ECG signals. Pattern Recogn. Lett. **70**, 45–51 (2016)
58. Martis, R., Acharya, R., Ray, A.: Application of higher order cumulants to ECG signals for the cardiac health diagnosis. In: 2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 1697–1700, Boston, MA (2011)
59. Kumar, M., Pachori, R.B., Acharya, U.R.: Characterization of coronary artery disease using flexible analytic wavelet transform applied on ECG signals. Biomed. Signal Process. **31**, 301–308 (2017)

60. Hassan, A.R., Haque, M.A.: An expert system for automated identification of obstructive sleep apnea from single-lead ECG using random under sampling boosting. Neurocomputing. **235**, 122–130 (2017)
61. Chawla, M.: A comparative analysis of principal component and independent component techniques for electrocardiograms. Neural Comput. Appl. **18**, 539–556 (2009)
62. Kallas, M., Francis, C., Kanaan, L., Merheb, D., Honeine, P., Amoud, H.: Multi-class SVM classification combined with kernel PCA feature extraction of ECG signals. In: International Conference on Telecommunications (ICT), pp. 1–5 (2012)
63. Ubeyli, E.D.: Recurrent neural networks employing Lyapunov exponents for analysis of ECG signals. Expert Syst. Appl. **37**, 1192–1199 (2010)
64. Ergin, S., Uysal, A.K., Gunal, E.S., Gunal, S., Gulmezoglu, M.B.: ECG based biometric authentication using ensemble of features. In: 9th Iberian Conference on Information Systems and Technologies (CISTI). IEEE. pp. 1–6 (2014)
65. Vafaie, M.H., Ataei, M., Koofigar, H.R.: Heart diseases prediction based on ECG signals' classification using a genetic-fuzzy system and dynamical model of ECG signals. Biomed. Signal Process. **14**, 291–296 (2014)
66. Gunal, S., Gerek, O.N., Ece, D.G., Edizkan, R.: The search for optimal feature set in power quality event classification. Expert Syst. Appl. **36**, 10266–10273 (2009)
67. Kittler, J.: Feature set search algorithms. In: Pattern Recognition and Signal Processing, vol. 29. Springer, Netherlands (1978)
68. Pudil, P., Novovicova, J., Kittler, J.: Floating search methods in feature-selection. Pattern Recogn. Lett. **15**, 1119–1125 (1994)
69. Goldberg, D.E.: Genetic Algorithms in Search, Optimization, and Machine Learning. Addison-Wesley, Reading (1989)
70. Chui, K.T., Tsang, K.F., Chi, H.R., Ling, B.W.K., Wu, C.K.: An accurate ECG-based transportation safety drowsiness detection scheme. IEEE Trans. Ind. Inf. **12**, 1438–1452 (2016)
71. Li, H.Q., Yuan, D.Y., Ma, X.D., Cui, D.Y., Cao, L.: Genetic algorithm for the optimization of features and neural networks in ECG signals classification. Sci. Rep. **7**, 41011 (2017)
72. Lin, S.-W., Ying, K.-C., Chen, S.-C., Lee, Z.-J.: Particle swarm optimization for parameter determination and feature selection of support vector machines. Expert Syst. Appl. **35**, 1817–1824 (2008)
73. Nimbhorkar, N.B., Alaspurkar, S.J.: Probabilistic neural network in solving various pattern classification problems. Int. J. Comput. Sci. Netw. Secur. **14**, 133–137 (2014)
74. Yu, S.-N., Chen, Y.-H.: Electrocardiogram beat classification based on wavelet transformation and probabilistic neural network. Pattern Recogn. Lett. **28**, 1142–1150 (2007)
75. Martis, R.J., Acharya, U.R., Min, L.C.: ECG beat classification using PCA, LDA ICA and discrete wavelet transform. Biomed. Signal Process. **8**, 437–448 (2013)
76. Wang, J.S., Chiang, W.C., Hsu, Y.L., Yang, Y.T.C.: ECG arrhythmia classification using a probabilistic neural network with a feature reduction method. Neurocomputing. **116**, 38–45 (2013)
77. Özbay, Y., Ceylan, R., Karlik, B.: A fuzzy clustering neural network architecture for classification of ECG arrhythmias. Comput. Biol. Med. **36**, 376–388 (2006)
78. Li, P.F., Wang, Y., He, J.C., Wang, L.H., Tian, Y., Zhou, T.S., Li, T.C., Li, J.S.: High-performance personalized heartbeat classification model for long-term ECG signal. I.E.E.E. Trans. Biomed. Eng. **64**, 78–86 (2017)
79. Jewajinda, Y., Chongstitvatana, P.: A parallel genetic algorithm for adaptive hardware and its application to ECG signal classification. Neural Comput. Appl. **22**, 1609–1626 (2013)
80. Osowski, S., Markiewicz, T., Hoai, L.T.: Recognition and classification system of arrhythmia using ensemble of neural networks. Measurement. **41**, 610–617 (2008)
81. Liu, Q., Pitt, D., Wu, X.: On the prediction of claim duration for income protection insurance policyholders. Ann. Actuar. Sci. **8**, 42–62 (2014)

82. Chen, T.H., Mazomenos, E.B., Maharatna, K., Dasmahapatra, S., Niranjan, M.: Design of a low-power on-body ECG classifier for remote cardiovascular monitoring systems. IEEE J. Emerging Sel. Top. Circuits Syst. **3**, 75–85 (2013)
83. Wang, J.S., Lin, C.W., Yang, Y.T.C.: A k-nearest-neighbor classifier with heartrate variability feature-based transformation algorithm for driving stress recognition. Neurocomputing. **116**, 136–143 (2013)
84. Homaeinezhad, M.R., Atyabi, S.A., Tavakkoli, E., Toosi, H.N., Ghaffari, A., Ebrahimpour, R.: ECG arrhythmia recognition via a neuro-SVM-KNN hybrid classifier with virtual QRS image-based geometrical features. Expert Syst. Appl. **39**, 2047–2058 (2012)
85. Martis, R.J., Acharya, U.R., Prasad, H., Chua, C.K., Lim, C.M., Suri, J.S.: Application of higher order statistics for atrial arrhythmia classification. Biomed. Signal Process. **8**, 888–900 (2013)
86. Lagerholm, M., Peterson, C., Braccini, G., Edenbrandt, L., Sornmo, L.: Clustering ECG complexes using Hermite functions and self-organizing maps. IEEE Trans. Biomed. Eng. **47**, 838–848 (2000)
87. Özbay, Y., Tezel, G.: A new method for classification of ECG arrhythmias using neural network with adaptive activation function. Digital Signal Process. **20**, 1040–1049 (2010)
88. Schölkopf, B., Smola, A.J.: Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond. MIT Press, Cambridge (2002)
89. Elhaj, F.A., Salim, N., Harris, A.R., Swee, T.T., Ahmed, T.: Arrhythmia recognition and classification using combined linear and nonlinear features of ECG signals. Comput. Methods Prog. Biomed. **127**, 52–63 (2016)
90. Raj, S., Ray, K.C.: ECG signal analysis using DCT-based DOST and PSO optimized SVM. IEEE Trans. Instrum. Meas. **66**, 470–478 (2017)
91. Seera, M., Lim, C.P., Liew, W.S., Lim, E., Loo, C.K.: Classification of electrocardiogram and auscultatory blood pressure signals using machine learning models. Expert Syst. Appl. **42**, 3643–3652 (2015)
92. Fayn, J.: A classification tree approach for cardiac ischemia detection using spatiotemporal information from three standard ECG leads. I.E.E.E. Trans. Biomed. Eng. **58**, 95–102 (2011)
93. Martis, R.J., Acharya, U.R., Prasad, H., Chua, C.K., Lim, C.M.: Automated detection of atrial fibrillation using Bayesian paradigm. Knowl.-Based Syst. **54**, 269–275 (2013)
94. Lee, J., McManus, D.D., Bourrell, P., Sornmo, L., Chon, K.H.: Atrial flutter and atrial tachycardia detection using Bayesian approach with high resolution time-frequency spectrum from ECG recordings. Biomed. Signal Process. **8**, 992–999 (2013)
95. Singh, Y.N., Gupta, P.: Correlation-based classification of heartbeats for individual identification. Soft. Comput. **15**, 449–460 (2011)
96. Abo-Zahhad, M., Ahmed, S.M., Abbas, S.N.: Biometric authentication based on PCG and ECG signals: present status and future directions. Signal Image Video Process. **8**, 739–751 (2014)
97. Yang, H., Kan, C., Liu, G., Chen, Y.: Spatiotemporal differentiation of myocardial infarctions. IEEE Trans. Autom. Sci. Eng. **10**, 938–947 (2013)
98. Gidaris, S., Komodakis, N.: Object detection via a multi-region and semantic segmentation-aware CNN model. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 1134–1142 (2015)
99. Abdel-Hamid, O., Mohamed, A. R., Jiang, H., Penn, G.: Applying convolutional neural networks concepts to hybrid NN-HMM model for speech recognition. In: 2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE. pp. 4277–4280 (2012)
100. Association for the Advancement of Medical Instrumentation: Testing and reporting performance results of cardiac rhythm and ST segment measurement algorithms. ANSI/AAMI EC38 (1998)

101. Li, D., Zhang, J., Zhang, Q., Wei, X.: Classification of ECG signals based on 1D convolution neural network. In: 2017 IEEE 19th International Conference on e-Health Networking, Applications and Services (Healthcom). IEEE. pp. 1–6 (2017)
102. Kiranyaz, S., Ince, T., Gabbouj, M.: Real-time patient-specific ECG classification by 1-D convolutional neural networks. IEEE Trans. Biomed. Eng. **63**(3), 664–675 (2015)
103. Kachuee, M., Fazeli, S., Sarrafzadeh, M.: ECG heartbeat classification: a deep transferable representation. In: 2018 IEEE International Conference on Healthcare Informatics (ICHI). IEEE. pp. 443–444 (2018)
104. Acharya, U.R., Fujita, H., Oh, S.L., Hagiwara, Y., Tan, J.H., Adam, M.: Application of deep convolutional neural network for automated detection of myocardial infarction using ECG signals. Inf. Sci. **415**, 190–198 (2017)
105. Liu, W., Zhang, M., Zhang, Y., Liao, Y., Huang, Q., Chang, S., Wang, H., He, J.: Real-time multilead convolutional neural network for myocardial infarction detection. IEEE J. Biomed. Health Inform. **22**, 1434–1444 (2017)
106. Liu, W., Huang, Q., Chang, S., Wang, H., He, J.: Multiple-feature-branch convolutional neural network for myocardial infarction diagnosis using electrocardiogram. Biomed. Signal Process. Control. **45**, 22–32 (2018)
107. Xia, Y., Wulan, N., Wang, K., Zhang, H.: Detecting atrial fibrillation by deep convolutional neural networks. Comput. Biol. Med. **93**, 84–92 (2018)
108. Clifford, G.D., Liu, C., Moody, B., Li-wei, H.L., Silva, I., Li, Q., Mark, R.G.: AF Classification from a short single lead ECG recording: the PhysioNet/Computing in Cardiology Challenge 2017. In: 2017 Computing in Cardiology (CinC). IEEE. pp. 1–4 (2017)
109. Andreotti, F., Carr, O., Pimentel, M.A., Mahdi, A., De Vos, M.: Comparing feature-based classifiers and convolutional neural networks to detect arrhythmia from short segments of ECG. In 2017 Computing in Cardiology (CinC). IEEE. pp. 1–4 (2017)
110. Rubin, J., Parvaneh, S., Rahman, A., Conroy, B., Babaeizadeh, S.: Densely connected convolutional networks and signal quality analysis to detect atrial fibrillation using short single-lead ECG recordings. In: 2017 Computing in Cardiology (CinC). IEEE. pp. 1–4 (2017)
111. Plesinger, F., Nejedly, P., Viscor, I., Halamek, J., Jurak, P.: Automatic detection of atrial fibrillation and other arrhythmias in holter ECG recordings using rhythm features and neural networks. In: 2017 Computing in Cardiology (CinC). IEEE. pp. 1–4 (2017)
112. Ghiasi, S., Abdollahpur, M., Madani, N., Kiani, K., Ghaffari, A.: Atrial fibrillation detection using feature based algorithm and deep convolutional neural network. In: 2017 Computing in Cardiology (CinC). IEEE. pp. 1–4 (2017)
113. Singh, S., Pandey, S.K., Pawar, U., Janghel, R.R.: Classification of ECG arrhythmia using recurrent neural networks. Proc. Comput. Sci. **132**, 1290–1297 (2018)
114. Yildirim, Ö.: A novel wavelet sequence based on deep bidirectional LSTM network model for ECG signal classification. Comput. Biol. Med. **96**, 189–202 (2018)
115. Schwab, P., Scebba, G.C., Zhang, J., Delai, M., Karlen, W.: Beat by beat: classifying cardiac arrhythmias with recurrent neural networks. In: 2017 Computing in Cardiology (CinC). IEEE. pp. 1–4 (2017)
116. Teijeiro, T., García, C.A., Castro, D., Félix, P.: Arrhythmia classification from the abductive interpretation of short single-lead ECG records. In: 2017 Computing in Cardiology (CinC). IEEE. pp. 1–4 (2017)
117. Warrick, P., Homsi, M.N.: Cardiac arrhythmia detection from ECG combining convolutional and long short-term memory networks. In: 2017 Computing in Cardiology (CinC). IEEE. pp. 1–4 (2017)
118. Zihlmann, M., Perekrestenko, D., Tschannen, M.: Convolutional recurrent neural networks for electrocardiogram classification. In: 2017 Computing in Cardiology (CinC). IEEE. pp. 1–4 (2017)
119. Limam, M., Precioso, F.: Atrial fibrillation detection and ECG classification based on convolutional recurrent neural network. In 2017 Computing in Cardiology (CinC). IEEE. pp. 1–4 (2017)

120. Xiong, Z., Nash, M.P., Cheng, E., Fedorov, V.V., Stiles, M.K., Zhao, J.: ECG signal classification for the detection of cardiac arrhythmias using a convolutional recurrent neural network. Physiol. Meas. **39**, 094006 (2018)
121. Shashikumar, S.P., Shah, A.J., Clifford, G.D., Nemati, S.: Detection of paroxysmal atrial fibrillation using attention-based bidirectional recurrent neural networks. In: Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. ACM. pp. 715–723 (2018)
122. Hong, S., Wu, M., Zhou, Y., Wang, Q., Shang, J., Li, H., Xie, J.: ENCASE: an ENsemble ClASsifiEr for ECG classification using expert features and deep neural networks. In: 2017 Computing in Cardiology (CinC). IEEE. pp. 1–4 (2017)
123. Jiang, C., Song, S., Meng, M.Q.H.: Heartbeat classification system based on modified stacked denoising autoencoders and neural networks. In: 2017 IEEE International Conference on Information and Automation (ICIA). IEEE. pp. 511–516 (2017)
124. Yang, J., Bai, Y., Lin, F., Liu, M., Hou, Z., Liu, X.: A novel electrocardiogram arrhythmia classification method based on stacked sparse auto-encoders and softmax regression. Int. J. Mach. Learn. Cybern. **9**, 1733–1740 (2018)
125. Rajan, D., Thiagarajan, J.J.: A generative modeling approach to limited channel ECG classification. In: 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE. pp. 2571–2574 (2018)
126. Yildirim, O., Baloglu, U.B., Tan, R.S., Ciaccio, E.J., Acharya, U.R.: A new approach for arrhythmia classification using deep coded features and LSTM networks. Comput. Methods Prog. Biomed. **176**, 121–133 (2019)
127. He, R., Liu, Y., Wang, K., Zhao, N., Yuan, Y., Li, Q., Zhang, H.: Automatic cardiac arrhythmia classification using combination of deep residual network and bidirectional LSTM. IEEE Access. **7**, 102119–102135 (2019)

ECG Interpretation with Deep Learning

Wenjie Cai and Danqin Hu

Abstract Electrocardiography (ECG), which can trace the electrical activity of the heart noninvasively, is widely used to assess heart health. Accurate interpretation of ECG requires significant amounts of education and training. With the application of deep learning, the accuracy of ECG diagnostic analysis has reached a new high level and even outperforms that of individual cardiologists. And the automated ECG diagnostic model makes it possible for analyzing ECG signals from wearable devices in real time. The common deep learning networks for analyzing ECG are mainly based on convolutional neural networks (CNN), recurrent neural networks (RNN), CNN plus RNN, and some other architectures. This chapter gives a systematical review on the CNN-based, RNN-based, as well as CNN and RNN-based intelligent analysis models for the automated ECG interpretation.

Keywords Convolutional neural network · Recurrent neural network · Electrocardiography interpretation

An electrocardiogram (ECG) is a record of the electrical activity of the heart. It is one of the routine means of detecting cardiovascular disease. Millions of ECGs were generated in hospitals and healthcare centers annually throughout the world. With the rise of wearable devices, more and more ECGs are being produced all the time. Automated analysis and interpretation of these data can help improve the work efficiency of medical physicians.

Computerized interpretation of ECG can date to late 1950s. Despite of the improvement of ECG analysis algorithms, the total accuracy of computer programs was 5.8% lower than average cardiologists [1]. The algorithms were very sensitive to the quality of ECG signals and made most frequent mistakes in arrhythmias, conduction disorders, and electronic pacemaker rhythms [2]. However, the

W. Cai (✉) · D. Hu
School of Medical Instrument and Food Engineering, University of Shanghai for Science and Technology, Shanghai, China
e-mail: wjcai@usst.edu.cn

C. Liu, J. Li (eds.), *Feature Engineering and Computational Intelligence in ECG Monitoring*, https://doi.org/10.1007/978-981-15-3824-7_8

computerized interpretation does help facilitating clinical decision making and saving physician readers' time. Combination of algorithm and physician readers can make the ECG diagnosis more accurate [3]. The improvement of algorithms can further reduce the workload of doctors and increase diagnostic accuracy.

Machine learning, which is a branch of artificial intelligence (AI), provides more intelligent ways to interpret ECGs. Algorithms such as k-nearest neighbor (KNN), support vector machine (SVM), and random forest parse ECG features learn from them and then make intelligent decisions to new samples [4]. Machine learning is very popular because the algorithm automatically recognizes the correlation between data features and classification labels, automatically finds the optimal solution, and shows strong robustness. However, disadvantage exists and is that professional knowledge is needed to extract the ECG features. The type, quality, and quantity of feature extraction are decisive for the final classification results.

Deep learning, which is a subset of machine learning, uses multilayered artificial neural network (ANN) to decrypt the data and has achieved great success in many areas such as computer vision, speech recognition, and natural language processing. It is so successful due to the application of convolutional neural network (CNN) and recurrent neural network (RNN), which are two major ideas of deep learning. With CNN, RNN, or combination of them, deep learning makes state-of-the-art solutions for many kinds of complicated tasks such as medical imaging analysis. The biggest advantage of deep learning is that no manual feature extraction is required during model training, which reduces the need for domain expertise. Actually, the deep learning model can learn the abstract features from data in an incremental manner. And in many cases the features extracted by deep learning model are more representative than that extracted by human experts [5]. The performance of such a model can outperform human experts. Figure 1 shows the difference between traditional machine learning and deep learning. "Deep" means more hidden layers in deep learning neural networks than in ANNs. Preprocessing of deep learning can be the same as traditional machine learning, or it can be different or even none at all. It mainly depends on the model's training strategy. Feature extraction is done automatically when training the deep learning model. In this chapter, we focus on the deep learning application on ECG analysis.

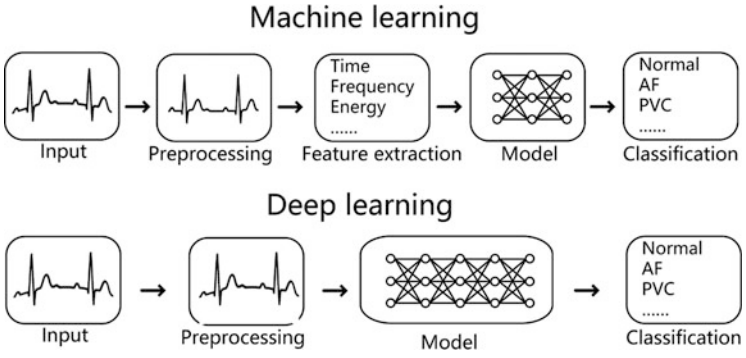

Fig. 1 Comparison between traditional machine learning and deep learning. *AF* atrial fibrillation, *PVC* premature ventricular contraction

1 Preprocessing

The resolution of ECG can be resampled to 120–500 Hz. Sampling rate should remain the same in the training data set and the test data set. An ECG with too low sampling frequency will lose some useful details and will affect the diagnostic accuracy, while an ECG with very high sampling frequency increases computational cost.

Deep learning does not require the data to have the same size. However, to speed up training, the data is calculated in batches, and the dimensions of the data in each batch must be the same. Normally, we arbitrarily choose the model input length. If the raw signal is longer than the chosen length, the ECGs can be divided into segments of that length. Typically, ECG of 5–10 s is enough for a medical practitioner to make a diagnosis conclusion. When the sample is too short for segmentation, it can be padded with zeros at the beginning or ending. It has to be noted that the shorter the sample is segmented, the finer the annotation must be. For most arrhythmias such as atrial fibrillation (AF), the symptom will last for a while. If the entire ECG is not very long, such as only 1 min, it is usually safe to say that the diagnosis conclusion of any small segmentation is just the same as the entire ECGs. However, for premature contractions such as premature ventricular contraction (PVC), abnormal heartbeats may appear multiple times in the entire 1-min ECG or occur just once. For these kinds of ECGs, the entire ECG is required for classification, or it needs more detailed annotation to indicate specifically where the premature beats are.

Performance of deep learning systems depends heavily on the number of training samples. In the real world, the training data is limited and sometimes even very little. Data augmentation creates new training data by applying domain-specific techniques to original data. It is an effective way to improve classification accuracy and reduce overfitting. For ECG, data augmentation can be done in several ways. First, random noise can be added to the signals. Second, sinusoidal signal can be fused with raw ECG to simulate baseline drifting. Third, samples can be renormalized with a random mean and standard deviation [6]. Data augmentation forces the model to learn the real pattern and helps improve model performance. ECG segmentation is also a kind of data augmentation because it can increase training samples significantly. The principle of data augmentation for ECG data is that cardiologists will make the same diagnosis conclusion for the converted data and the original data.

2 CNN Model

CNN is widely used in computer vision due to its powerful ability to extract abstract features from various images. CNN layers take inputs and calculate with filters. Different filters generate different features. The outputs contain certain spatial features of the inputs. When stacked with several CNN layers, the model can identify a variety of complex features. Through these features, the model makes state-of-the-

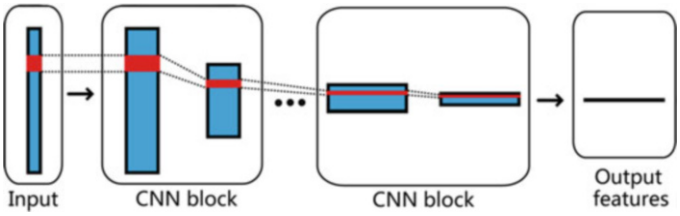

Fig. 2 Block diagram of a CNN model

art image recognition. For ECG signals, whatever single lead, two leads, or 12 leads were used, one-dimensional CNN is usually used to decode the time series signals and preserve the features of adjacent signals' voltage values. The learned features are used to make final classifications (Fig. 2).

Acharya et al. used beat-to-beat strategy with a deep neural network, which contained three layers of 1D-CNN to classify five different categories (non-ectopic, supraventricular ectopic, ventricular ectopic, fusion, and unknown) of heartbeats, and they achieved an accuracy of 94.03% [7] on the MIT-BIH Arrhythmia database (MITDB) [8]. They also tested segment strategy and proposed a model that contained four layers of 1D-CNN to automatically discriminate four different arrhythmias (normal, AF, atrial flutter, and ventricular flutter) on the data from three different open databases. ECG with 2-s and 5-s segments were input into the model, and the model achieved accuracies of 92.50% and 94.90%, respectively [9]. Yıldırım et al. designed a model that contained seven CNN layers to detect 17 classes of arrhythmia with the MITDB. The analysis is based on 10-s ECG fragments and reached overall accuracy of 91.33% with low computational complexity [10]. The more CNN layers are used, the more complex pattern can be extracted.

Hannun et al. developed a deep residual neural network to discriminate ECG into 12 classes. The architecture had 34 layers that contained 16 residual blocks, and each residual block had two convolutional layers. The model was trained and tested on a large ECG data set containing 91,232 single-lead ECGs. The average $F1$ scores for this model reached 0.837 and exceeded that of average cardiologists of 0.780. Additionally, the model also ranked among the best in the 2017 PhysioNet/CinC Challenge which used a medium-sized ECG data set [11, 12]. With the network depth increases, the varnishing gradient problem makes the model hard to train. Residual network can overcome this problem by skipping connection to add the output to a later layer from an earlier layer [13]. It has advantages of keeping the network deep but trainable.

Apart from 1D-CNN, 2D-CNN can also be used to analyze ECG signals. In this case, the signal needs to be converted to 2D version. Ji et al. extracted heartbeats and converted each beat into an image. Then the image was fed to the 2D-CNN model which was composed of the ZF net (a CNN architecture winner of ImageNet competition 2013), region proposal network (RPN) net, and faster regions with CNN (Faster R-CNN) net. The model classified ECG beats into five categories

using the MITDB with an accuracy of 99.21%, which was much higher than that of SVM algorithm [14]. Rahhal et al. applied continuous wavelet transform (CWT) to the ECG signals to generate time–frequency 2D features. Then they fed these features into a CNN model which was pretrained on the ImageNet data set and got state-of-the-art results on the detection of ventricular ectopic beats (VEB) with an accuracy of 99.9% and supraventricular ectopic beats (SVEB) with an accuracy of 99.8% [15]. Similarly, Xia et al. used short-term Fourier transform (STFT) and stationary wavelet transform (SWT) to transform the 1D ECG signals to 2D matrix. STFT and SWT features were then fed into two deep CNN architectures, and the results showed better diagnostic sensitivity, specificity, and accuracy for AF detection compared to other conventional machine learning techniques with an accuracy of 98.63% [16]. 2D-CNN is the core part of compute vision and is believed to be able to obtain better spatial details. A comparison study on classifying ECG into normal and abnormal beats with the MITDB showed that 2D-CNN model could reach an accuracy of 98% and was 2% better than the best 1D-CNN model [17].

3 RNN Model

When doctors read ECG trying to get a diagnosis conclusion, they read it segment by segment while keeping memories of the information of early segments. RNN is such a deep learning architecture. It processes the input sequence one by one while keeping a state containing early information (Fig. 3). It has been successfully applied to sequence data such as speech recognition and natural language processing. Long short-term memory (LSTM) and gated recurrent unit (GRU) are most popular RNN architectures. LSTM maintains three gates, which are an input gate, an output gate, and a forget gate. These three nonlinear gates determined what past information to be remembered [18]. GRU is a kind of simplified version of LSTM. It has two gates: an update gate and a reset gate [19, 20]. LSTM is more powerful but computationally expensive while GRU has the similar effects with cheaper computation.

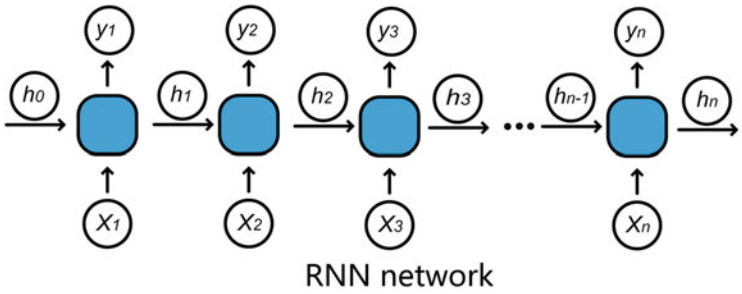

Fig. 3 Block diagram of an RNN model

Faust et al. used a bidirectional LSTM neural network to detect AF with the MIT-BIH Atrial Fibrillation Database (AFDB). They partitioned the RR interval signals with a sliding window of 100 samples. Then the samples were fed into the model and got an accuracy of 98.51% with tenfold cross-validation. The results were promising, but the training process was really slow [21]. Wang et al. proposed a global classification model named global recurrent neural network (GRNN). Inputs of the GRNN include two parts. The first one was the raw, single heartbeat signals, which were input into two stacked LSTM layers to extract morphological features. The other one was temporal features related to RR intervals. These two kinds of features were then put into an RNN layer that served as feature learning layer, and the output was the classification results. The model achieved accuracies of more than 99% to detect supraventricular ectopic and ventricular ectopic beats on several ECG database [22].

Apart from the raw signal input into RNN model, the data can be preprocessed first to get some preliminary local features and then fed into the model. This process will save lot of training time. Yildirim tried to classify five types of heartbeats, which are normal, PVC, paced beat (PB), left bundle branch block (LBBB), and right bundle branch block (RBBB) with the MITDB. He applied discrete wavelet transform (DWT) to ECGs to get meaningful features and then input these features to two stacked, bidirectional LSTM layers. Data preprocessing with DWT significantly improved the model performance compared with no data preprocessing. And the model with bidirectional LSTM layers was superior to the model with unidirectional LSTM layers. Their best model got a very high recognition performance of 99.39% on the test set [23]. Actually, the most common preprocess for RNN model is CNN. We will elaborate it in the next section.

4 CNN + RNN Model

The combination of CNN and RNN is becoming more and more popular. Such a model can not only extract local spatial features by using CNN but also make full use of the sequence processing capability of RNN. More importantly, the computational cost is greatly reduced while ensuring the accuracy of the model. Figure 4 shows a CNN plus GRU model that we used to in the first China Physiological Signal Challenge (CPSC 2018). This model is composed of three stacked CNN blocks, two stacked bidirectional GRU layers, and two fully connected layers. The inputs are 5-s 12-lead ECG segments that have been downsampled to 125 Hz. With this model and some optimization skills, we achieved the $F1$ score of 0.83 for nine classes ECG and won the second place. Similarly, Zihlmann et al. first computed logarithmic spectrograms of ECGs and used stacked CNN layers to extract spectrogram features. Then a LSTM layer was used to calculate the features across time. This model performed better than the model without a LSTM layer and got second place in the competition of the 2017 PhysioNet/CinC Challenge [24]. Xiong et al. developed a 21-layer 1D convolutional recurrent neural network, which contained 16 residual

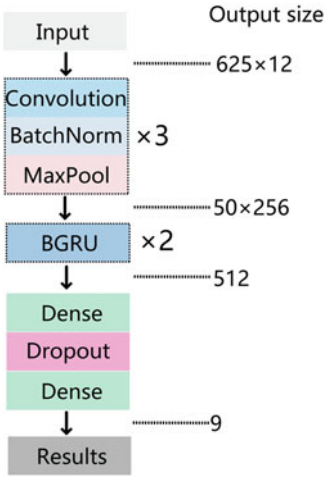

Fig. 4 CNN + RNN network structure used in the CPSC 2018

CNN blocks and 3 RNN layers to classify ECGs in the 2017 PhysioNet/CinC Challenge. They used 5-s segments of ECG signals as inputs. The model had an $F1$ score of 0.82, which is 0.01 less than the top score [25]. These excellent models emerged in the ECG challenges indicate that the combination of CNN and RNN has superior advantages.

The diversity of the models lies in the combination of different numbers of CNN and RNN, and all have extraordinary performance. Tan et al. proposed a model that consisted of two stacked CNN layers and three stacked LSTM layers to automatically detect coronary artery disease (CAD) ECG signals. The CAD ECG data were from Fantasia and St Petersburg Institute of Cardiology Technics 12-leads arrhythmia. The model was able to diagnose CAD ECG signals with an accuracy of 99.85% [6]. Andersen et al. designed a five-layer model for real-time detection of AF. The data were from the MITDB, the AFDB, and the MIT-BIH NSR Database (NSRDB). Their network contained two stacked CNN and one LSTM layer. The input sequence was 30 RR intervals. This model achieved 97.8% for accuracy [26]. Oh et al. developed a network combined with CNN and LSTM to classify ECGs into five classes, which were normal, PAC, PVC, LBBB, and RBBB. The model had three stacked convolutional layers followed by one LSTM layer. The model achieved an accuracy of 98.10% on the MITDB [27].

5 Other Deep Learning Architectures

Deep ANN can also be used for ECG classification. Xu et al. proposed a five-layer deep ANN for heartbeat classification. Their strategy was based on beat-by-beat detection and required heartbeat segmentation and alignment. The model achieved an overall accuracy of 94.7% for detecting non-ectopic, supraventricular ectopic,

ventricular ectopic, fusion, and unknown beats [28]. Sannino et al. proposed a nine-layer deep ANN to automatically recognize abnormal heartbeats from the MITDB without using CNN or RNN. They segmented ECGs into individual heartbeats, and then took 50 uniformly distributed samples from each P wave to T wave and concatenated them with four temporal features as model inputs. The model showed an accuracy of 99.09% for classifying normal and abnormal heartbeats [29].

Deep learning models can be used just as feature extractors. Liu fed deep learning features and 174 other expert features into an XGBoost (eXtreme Gradient Boosting of decision trees) model to classify the ECG data from the CPSC 2018 into nine categories. Deep learning features were from a deep residual neural network that contains 17 layers of convolution. Expert features included time domain features, frequency domain features, and other specific features that had statistics and physiology significance. The model reached a $F1$ score of 0.81 and was ranked among top ten of the leaderboard of the CPSC 2018 [30].

There are some other deep learning methods that can be used to classify ECG signals such as restricted Boltzmann machines and deep belief networks [31]. With the development of deep learning technology, more sophisticated models will be developed. Table 1 summarizes the studies on ECG interpretation with deep learning.

6 Conclusion

Artificial intelligence has made great strides in many fields, and technological innovation has changed people's lives. The application of deep learning in ECG analysis has made remarkable achievement. It performs better than the traditional machine learning and even has higher diagnostic accuracy than cardiologists [12]. Along with the emerging of lots of wearable devices, which can collect ECG signals, fully automated analysis of ECG is becoming very important. It can improve people's health management level and screen out suspicious heart disease. The use of artificial intelligence to aid diagnosis can greatly increase work efficiency and reduce the rate of misdiagnosis [3].

The most common means of deep learning today include CNN, RNN, or a combination of both. CNN is good at capturing spatial features and has low computational cost. RNN is good at processing time series data but had high computational cost. Combining CNN and RNN may be a smart solution under the limitation of computing power. Despite of great achievements of automated ECG interpretation with deep learning, we have several concerns. First, the models proposed in most researches can only identify a few kinds of abnormal ECGs, so their clinical application value is limited. Second, many models are reported to have high ECG interpretation accuracy. However, they were tested in different databases or they discriminated different ECG classes, so it is hard to tell which one is better with the metrics such as accuracy or $F1$ score. The main bottleneck is the lack of standardized large databases. The most used ECG database in literature is the

Table 1 Studies on ECG interpretation with deep learning

Author, Year	Database	Preprocessing	Analysis strategy	Classes	Performance
CNN					
Acharya et al. [7]	MITDB	Heartbeat segmentation and alignment; Total: 109,449 beats	Beat-by-beat	Non-ectopic, Supraventricular ectopic, Ventricular ectopic, Fusion, Unknown	Accuracy: 94.03%
Acharya et al. [9]	MITDB AFDB CUDB	2-s segments or 5 s segments; Total: 21,709 segments (2-s), 8683 segments (5-s)	Segments	Normal, AF, Atrial flutter, Ventricular flutter,	Accuracy (2-s): 92.50%, Accuracy (5-s): 94.90%
Yıldırım et al. [10]	MITDB	10-s segments; Total: 1000 segments	Segments	Normal, PAC, Atrial flutter, AF, Supraventricular tachyarrhythmia, Pre-excitation, PVC, Ventricular bigeminy, Ventricular trigeminy, Ventricular tachycardia, Idioventricular rhythm, Ventricular flutter, Fusion, LBBB, RBBB, Second-degree heart block, Pacemaker rhythm	Accuracy: 91.33%

(continued)

Table 1 (continued)

Author, Year	Database	Preprocessing	Analysis strategy	Classes	Performance
Hannun et al. [12]	Data set from iRhythm Technologies Inc.	Total: 91,232 records	Records	AF and atrial flutter, AVB (atrioventricular block), Bigeminy, EAR (ectopic atrial rhythm), IVR (idioventricular rhythm), Junctional rhythm, Noise, Sinus rhythm, SVT (supraventricular tachycardia), Trigeminy, Ventricular tachycardia, Wenckebach	*F*1 score: 0.837
Ji et al. [14]	MITDB Own data	Heartbeat segmentation and alignment; ECG beats were transformed into images; Total: 48,992 beats	Beat-by-beat	Normal LBBB RBBB PVC Fusion	Accuracy: 99.21%
Rahhal et al. [15]	MITDB INCART SVDB	Heartbeat segmentation and resampling; CWT; Total:100,635 beats (MITDB), 157,797 beats (INCART), 164,551 beats (SVBD)	Beat-by-beat	Non-ectopic, Supraventricular ectopic, Ventricular ectopic, Fusion	Accuracy: V(MITDB): 99.9% S(MITDB): 99.8% V(INCART): 99.23% S(INCART): 99.82% V(SVDB): 99.40% S(SVDB): 98.40%

Xia et al. [16]	AFDB	5-s data segments; STFT or SWT; Total: 162,536 segments	Segments	Non-AF AF	Accuracy (STFT): 98.29% Accuracy (SWT): 98.63%
RNN					
Faust et al. [21]	AFDB	Segments with 100 beats	Segments	Non-AF AF	Accuracy: 98.51%
Wang et al. [22]	MITDB INCARTDB SVDB	Beat segmentation; RR-interval calculation; Total:100,420 beats (MITDB), 175,575 beats (INCARTDB), 184,023 beats (SVDB)	Beat-by-beat	Non-ectopic, Supraventricular ectopic, Ventricular ectopic, Fusion	Accuracy: V (MITDB): 99.7% S (MITDB): 99.7% V (INCART): 99.8% S (INCART): 99.9% V (SVDB): 99.4% S (SVDB): 99.0%
Yildirim [23]	MITDB	Beat segmentation; Wavelet transformation; Total: 7326 beats	Beat-by-beat	Normal LBBB RBBB PVC PB	Accuracy: 99.39%
CNN + RNN					
Zihlmann et al. [24]	PhysioNet/CinC Challenge 2017	Total: 8528 records	Records	Normal, AF, Other, Noisy	*F*1 score: 0.83
Xiong et al. [25]	PhysioNet/CinC Challenge 2017	5-s segments; Total: 8528 ECG records	Records	Normal, AF, Other, Noisy	*F*1 score: 0.82
Tan et al. [6]	Fantasia and INCARTDB	5-s segments; Total: 38,120 segments	Segments	Normal, CAD	Accuracy: 99.85%

(continued)

Table 1 (continued)

Author, Year	Database	Preprocessing	Analysis strategy	Classes	Performance
Andersen et al. [26]	AFDB MITDB NSRDB	Segments with 30 RR intervals	Segments	Non-AF, AF	Accuracy: 97.8%
Oh et al. [27]	MITDB	2.78-s segments; Total:16,499 segments	Segments	Normal, LBBB, RBBB, PAC, PVC	Accuracy: 98.10%
Other deep learning architectures					
Xu et al. [28]	MITDB	Heartbeat segmentation and alignment; Total: 100,645 beats	Beat-by-beat	Non-ectopic, Supraventricular ectopic, Ventricular ectopic, Fusion, Unknown	Accuracy: 94.7%
Sannino and De Pietro [29]	MITDB	Heartbeat segmentation; RR-interval calculation; Total: 4576 beats	Beat-by-beat	Normal, Abnormal	Accuracy: 99.68%
Liu et al. [30]	CPSC 2018	CNN was used to extract features; Expert features were extracted; Total: 9831 records	Records	Normal, AF, AVB, LBBB, RBBB, PAC PVC, ST-segment depression ST-segment elevated	*F*1 score: 0.81

MITDB, which has only 48 records. The CPSC2018 has made a good attempt to provide a moderate-sized clinical 12-lead ECG database [32]. More open, large, standard ECG databases will definitely take the automated ECG interpretation to the next level.

References

1. Rautaharju, P.M.: Eyewitness to history: landmarks in the development of computerized electrocardiography. J. Electrocardiol. **49**, 1–6 (2016)
2. Guglin, M.E., Thatai, D.: Common errors in computer electrocardiogram interpretation. Int. J. Cardiol. **106**, 232–237 (2006)
3. Smulyan, H.: The computerized ECG: friend and foe. Am. J. Med. **132**, 153–160 (2019)
4. Berkaya, S.K., Uysal, A.K., Gunal, E.S., Ergin, S., Gunal, S., Gulmezoglu, M.B.: A survey on ECG analysis. Biomed. Signal Process. Control. **43**, 216–235 (2018)
5. Miotto, R., Wang, F., Wang, S., Jiang, X., Dudley, J.T.: Deep learning for healthcare: review, opportunities and challenges. Brief. Bioinform. **19**, 1236–1246 (2018)
6. Tan, J.H., Hagiwara, Y., Pang, W., Lim, I., Oh, S.L., Adam, M., Tan, R.S., Chen, M., Acharya, U.R.: Application of stacked convolutional and long short-term memory network for accurate identification of CAD ECG signals. Comput. Biol. Med. **94**, 19–26 (2018)
7. Acharya, U.R., Oh, S.L., Hagiwara, Y., Tan, J.H., Adam, M., Gertych, A., Tan, R.S.: A deep convolutional neural network model to classify heartbeats. Comput. Biol. Med. **89**, 389–396 (2017)
8. Goldberger, A.L., Amaral, L.A., Glass, L., Hausdorff, J.M., Ivanov, P.C., Mark, R.G., Mietus, J.E., Moody, G.B., Peng, C.K., Stanley, H.E.: PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. Circulation. **101**, E215 (2000)
9. Acharya, U.R., Fujita, H., Lih, O.S., Hagiwara, Y., Tan, J.H., Adam, M.: Automated detection of arrhythmias using different intervals of tachycardia ECG segments with convolutional neural network. Inf. Sci. **405**, 81–90 (2017)
10. Yıldırım, Ö., Pławiak, P., Tan, R.S., Acharya, U.R.: Arrhythmia detection using deep convolutional neural network with long duration ECG signals. Comput. Biol. Med. **102**, 411–420 (2018)
11. Clifford, G.D., Liu, C.Y., Moody, B., Silva, I., Li, Q., Johnson, A., Mark, R.: AF classification from a short single lead ECG recording: the PhysioNet/computing in cardiology challenge. Comput. Cardiol. **44**, 469 (2017)
12. Hannun, A.Y., Rajpurkar, P., Haghpanahi, M., Tison, G.H., Bourn, C., Turakhia, M.P., Ng, A. Y.: Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network. Nat. Med. **25**, 65–69 (2019)
13. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. arXiv:1512.03385 (2015)
14. Ji, Y., Zhang, S., Xiao, W.: Electrocardiogram classification based on faster regions with convolutional neural network. Sensors. **19**, 2558 (2019)
15. Rahhal, M.M.A., Bazi, Y., Zuair, M.A., Othman, E., Benjdira, B.: Convolutional neural networks for electrocardiogram classification. J. Med. Biol. Eng. **38**, 1014–1025 (2018)
16. Xia, Y., Wulan, N., Wang, K., Zhang, H.: Detecting atrial fibrillation by deep convolutional neural networks. Comput. Biol. Med. **93**, 84–92 (2018)
17. Wu, Y., Yang, F., Liu, Y., Zha, X., Yuan, S.: A comparison of 1-D and 2-D deep convolutional neural networks in ECG classification. arXiv:1810.07088 (2018)
18. Hochreiter, S., Schmidhuber, J.: Long short-term memory. Neural Comput. **9**, 1735–1780 (1997)

19. Cho, K., Van Merrienboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., Bengio, Y.: Learning phrase representations using RNN encoder-decoder for statistical machine translation. arXiv:1406.1078 (2014)
20. Chung, J., Gulcehre, C., Cho, K.H., Bengio, Y.: Empirical evaluation of gated recurrent neural networks on sequence modeling. arXiv:1412.3555 (2014)
21. Faust, O., Shenfield, A., Kareem, M., San, T.R., Fujita, H., Acharya, U.R.: Automated detection of atrial fibrillation using long short-term memory network with RR interval signals. Comput. Biol. Med. **102**, 327–335 (2018)
22. Wang, G., Zhang, C., Liu, Y., Yang, H., Fu, D., Wang, H., Zhang, P.: A global and updatable ECG beat classification system based on recurrent neural networks and active learning. Inf. Sci. **501**, 523–542 (2018)
23. Yildirim, Ö.: A novel wavelet sequence based on deep bidirectional LSTM network model for ECG signal classification. Comput. Biol. Med. **96**, 189–202 (2018)
24. Zihlmann, M., Perekrestenko, D., Tschannen, M.: Convolutional recurrent neural networks for electrocardiogram classification. arXiv:1710.06122 (2017)
25. Xiong, Z., Nash, M.P., Cheng, E., Fedorov, V.V., Stiles, M.K., Zhao, J.: ECG signal classification for the detection of cardiac arrhythmias using a convolutional recurrent neural network. Physiol. Meas. **39**(9), 094006 (2018)
26. Andersen, R.S., Peimankar, A., Puthusserypady, S.: A deep learning approach for real-time detection of atrial fibrillation. Expert Syst. Appl. **115**, 465–473 (2019)
27. Oh, S.L., Ng, E.Y.K., Tan, R.S., Acharya, U.R.: Automated diagnosis of arrhythmia using combination of CNN and LSTM techniques with variable length heart beats. Comput. Biol. Med. **102**, 278–287 (2018)
28. Xu, S.S., Mak, M., Cheung, C.: Towards end-to-end ECG classification with raw signal extraction and deep neural networks. IEEE J. Biomed. Health Inform. **23**, 1574–1584 (2019)
29. Sannino, G., De Pietro, G.: A deep learning approach for ECG-based heartbeat classification for arrhythmia detection. Futur. Gener. Comput. Syst. **86**, 446–455 (2018)
30. Liu, Z., Meng, X., Cui, J., Huang, Z., Wu, J.: Automatic identification of abnormalities in 12-Lead ECGs using expert features and convolutional neural networks. In: 2018 International Conference on Sensor Networks and Signal Processing (SNSP), vol. 1, pp. 163–167 (2018)
31. Mathews, S.M., Kambhamettu, C., Barner, K.E.: A novel application of deep learning for single-lead ECG classification. Comput. Biol. Med. **99**, 53–62 (2018)
32. Liu, F., Liu, C., Zhao, L., Zhang, X., Wu, X., Xu, X., Liu, Y., Ma, C., Wei, S., He, Z.: An open access database for evaluating the algorithms of electrocardiogram rhythm and morphology abnormality detection. J. Med. Imaging Health Inform. **8**, 1368–1373 (2018)

Visualizing ECG Contribution into Convolutional Neural Network Classification

Yaowei Li, Frédéric Precioso, and Chengyu Liu

Abstract Convolutional Neural Network (CNN) demonstrated impressive classification performance on ElectroCardioGram (ECG) analysis and has a great potential to extract salient patterns from signal. Visualizing local contributions of ECG input to a CNN classification model can both help us understand why CNN can reach such high-level performances and serve as a basis for new diagnosis recommendations. This chapter builds a single-layer 1-D CNN model on ECG classification with a 99% accuracy. The trained model is then used to build a 1-D Deconvolved Neural Network (1-D DeconvNet) for visualization. We propose Feature Importance Degree Heatmap (FIDH) to interpret the contribution of each point of ECG input to CNN classification results, and thus to show which part of ECG raises attention of the classifier. We also illustrate the correlation between two visualization methods: first-order Taylor expansion and multilayer 1-D DeconvNet.

Keywords Convolutional neural network · Deconvolved neural network · Deep learning · Electrocardiography interpretation

Y. Li
Universite Cote d'Azur, CNRS, I3S, Nice, France

School of Instrument Science and Engineering, Southeast University, Nanjing, China

F. Precioso (✉)
Universite Cote d'Azur, CNRS, I3S, Nice, France
e-mail: Frederic.Precioso@unice.fr

C. Liu (✉)
The State Key Laboratory of Bioelectronics, School of Instrument Science and Engineering, Southeast University, Nanjing, China
e-mail: chengyu@seu.edu.cn

C. Liu, J. Li (eds.), *Feature Engineering and Computational Intelligence in ECG Monitoring*, https://doi.org/10.1007/978-981-15-3824-7_9

1 Introduction

Deep learning is a branch of artificial intelligence, which outperforms traditional machine learning methods and can even outperform humans on signal analysis tasks, for instance, achieving higher accuracy on heart disease diagnose than cardiologists [1]. Convolutional Neural Network (CNN) was first introduced by Fukushima in 1980 [2] and later improved by LeCun et al. [3]. It is a kind of neural network where hidden layers follow a structure providing invariance or robustness to spatial transformations (i.e., scale changes, translation changes) [3]. CNN has become widely used in many computer vision tasks, achieving state-of-the-art performance, for instance, for object detection [4]. It also makes remarkable solutions on electrocardiogram (ECG) analysis and has a great potential to extract salient patterns of signal [1, 5–10].

Kiranyaz et al. [11] studied the real-time "patient-specific" ECG classification system using 1-D CNN, reaching an accuracy of 97.60% and 99.00% in the detection of supraventricular ectopic beats and ventricular ectopic beats, respectively, on MIT-BIH Arrhythmia database [12]. Acharya et al. [13] tested segment strategy and proposed a model which contained four layers of 1-D CNN to automatically discriminate four different arrhythmias (normal, atrial fibrillation, atrial flutter, and ventricular flutter). The data were collected from three different open databases. ECGs with 2-s and 5-s segments were input into the model achieving accuracies of 92.50% and 94.90%, respectively [14].

However, it is hard to interpret why the deep neural network perform so well. In deep neural network, the input components are merged after the first layer, which cause diluted links between higher layers and input data and lead to a lack of interpretability. This drawback hinders the cardiologist to integrate their expertise knowledge into the network or verify the classification decision. "Why your deep network makes such decision?" is hard to explain to physicians or patients, which considerably limits values and trustworthiness of the decision result.

Several works have been dedicated to understand the internal operations of multilayer network classification and explain what features are captured and learned inside the network. Explanation of neural network behavior on the level of single neurons is given in [15, 16]. These works try to find inputs which maximize the activation of neurons by means of optimization problems which can be solved by gradient ascent. To visualize CNN, Zeiler and Fergus [17] proposed "Deconvolved Network" to visualize deep models by mapping the hidden features in each intermediate layer to input pixel space, showing which patterns from the training set activate the feature map. However, it does not illustrate how the network makes decision and why this visualization would be valid. Simonyan et al. [18] establish an interpretation of the work done in [17] as an approximation to partial derivatives with respect to pixels in the input image. However, it does not match multilayer network.

In this chapter, we aim to propose a simple CNN to classify 1-D ECG signals and interpret which feature in the input signal is discriminative during the CNN

classification. We first train a 1-D Convolutional Neural Network (CNN) to classify four types of ECG beats. To interpret the feature from the trained 1-D CNN, we visualize the network by applying the layer-wise propagation "Deconvolved Network" in 1-D signal space, and thus, in this chapter, we call it "1-D deconvolved neural network (1-D DeconvNet)." The principle of 1-D CNN and 1-D deconvolved neural network is presented in Sect. 2.2. In this section, we interpret the relationship between Taylor expansion and DeconvNet. The architecture of 1-D CNN (see Sect. 2.3.1) and 1-D deconvolved neural network (see Sect. 2.3.2) is described in Sect. 2.3. The dataset used in this chapter is a wearable ECG database collected using conductive textile dry electrodes embedded in the Wearable 12-lead ECG Lenovo SmartVest system (see Sect. 3.1). The classification result of 1-D CNN reached an accuracy of 96%; more details are shown in Sect. 3.2. The visualization of 1-D CNN is presented as Feature Importance Degree Heatmap (FIDH), which is shown in Sect. 3.3.

2 Method

2.1 Principle of 1-D Convolutional Neural Network (1-D CNN)

CNN is a hierarchical network which consists of convolutional layers alternate with pooling layers to automatically extract features, followed with fully connected layers as a classifier. Compared to a 2-D CNN commonly used in vision tasks which requires a large training set and numerous iterations to converge (often hundreds of epochs), 1-D CNN is easier to train with only a few dozens of epochs, making it a good choice for 1-D ECG signal classification.

For a convolution layer, the convolution operation is as follows:

$$Z^{(l)} = F\left(X^{(l)} * K^{(l)}\right) \tag{1}$$

where $X^{(l)}$ and $K^{(l)}$ are feature map and kernel at layer l, respectively, $Z^{(l)}$ is the output of the convolutional layer, $*(\cdot)$ is the filtering (convolution) operation, and $F(\cdot)$ is the activate function.

In this study, we choose ReLU as activation, defined by the following:

$$F(Y_i) = \max(0, Y_i) \tag{2}$$

where Y is a 1-D vector and Y_i is the ith element of Y.

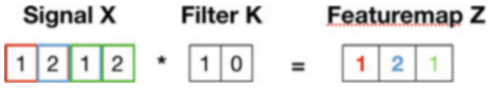

Fig. 1 1-D convolution operation

Fig. 2 1-D max-pooling operation

The convolutional layer is then followed with the pooling layer. We choose max-pooling in this chapter:

$$X^{(l+1)} = \text{maxpooling}\left(Z^{(l)}\right) \tag{3}$$

$X^{(l+1)}$ is the output of max-pooling in layer l, which is also the 1-D input vector of layer $l + 1$. $Z^{(l)}$ is the input of max-pooling in layer l, which is also the 1-D output vector of convolution operation.

In the max-pooling operation, each element in the output vector $X^{(l+1)}$ is selected from the input vector $Z^{(l)}$:

$$X_i^{(l+1)} = \max_{j \in \Omega(i)} \left(Z_j^{(l)}\right) \tag{4}$$

$X_i^{(l+1)}$ is the ith element of $X^{(l+1)}$. $Z_j^{(l)}$ is the jth element of the 1-D vector $Z^{(l)}$, which is the output of the convolution layer.

Thus, in a multilayer CNN, the input vector $X^{(0)}$ is processed by a combination of convolutions, max-pooling, and ReLU:

$$\begin{aligned} X^{(n)} &= F\left(\text{maxpooling}\left(X^{(n-1)} * K^{(n-1)}\right)\right) \\ &= F\left(\text{maxpooling}\left(F\left(\textit{maxpooling}\left(X^{(n-2)} * K^{(n-2)}\right)\right) * K^{(n-1)}\right)\right) \\ &= \ldots \\ &= F\left(\ldots F\left(\text{maxpooling}\left(X^{(0)} * K^{(0)}\right)\right)\right) \end{aligned} \tag{5}$$

The output of layer n $X^{(n)}$ is then followed by a classification operation, such as fully connected layers and softmax layer. In this study, a flatten layer, a dense layer (also called fully connected layer), and a softmax layer lead to the classification decision.

The whole architecture is shown in Fig. 5. A flatten layer concatenates all channels from the previous layer. The output of this architecture is a classification score $f(x)$ obtained to predict which class the input vector belongs to.

2.2 *Visualization of 1-D CNN: 1-D Deconvolved Neural Network*

2.2.1 Visualization by First-Order Taylor Expansion

Although we build high-accuracy network in Sect. 2.1, we still cannot explicit what makes the CNN performing so well. The overall idea of CNN visualization is to understand the contribution of each single component from an input vector x to the prediction $f(x)$ made by a classifier f. The objective is, thus, to measure how much each input vector component contributes to a positive or negative classification results.

Simonyan et al. [18] address the visualization of deep CNN by using first-order Taylor expansion. They used the score of a given class $f_c(x)$. In such a setup, there is a mapping $f_c : R^v \rightarrow R^1$ such that $f_c(x) > 0$ denotes the presence of the learned structure. We are interested in finding out the contribution of each input component $x(d)$ of an input signal x to a specific prediction $f_c(x)$.

Given a signal x_0, we can approximate $f_c(x)$ with a linear function in the neighborhood of x_0 by computing the first-order Taylor expansion:

$$f_c(x) = f_c(x_0) + w \cdot (x - x_0) \tag{6}$$

where x_0 is given and w is the first-order derivative of $f(x_0)$:

$$w = \frac{\partial f_c(x_0)}{\partial x_0} \tag{7}$$

Thus, the influence of each position in a given signal x_0 can be ranked by Eq. (7). Another interpretation using the class score derivative is that the magnitude of the derivative indicates which position needs to be changed the least to affect the class score the most. Baehrens et al. [19] also visualize classifier predictions by using partial derivatives at the prediction point x but in the context of Bayesian classification.

In the forward propagation of CNN, the input of each layer $X^{(l+1)}$ is calculated by the following:

$$X^{(l+1)} = G\left(X^{(l)} * K^{(l)}\right) \tag{8}$$

where $X^{(l)}$ is the convolutional kernel of layer l and $G(\cdot)$ is a nonlinear function. In this chapter, it consists of ReLU and max-pooling.

The gradient of $f(X^{(0)})$:

$$\begin{aligned}\frac{\partial f_c}{\partial X^{(0)}} &= \frac{\partial f_c}{\partial X^{(1)}} \cdot \frac{\partial X^{(1)}}{\partial X^{(0)}} \\ &= \frac{\partial f_c}{\partial X^{(1)}} \cdot \frac{\partial G}{\partial X^{(0)} * K^{(0)}} * K^{(0)^{\mathrm{T}}} \\ &= \frac{\partial f_c}{\partial X^{(2)}} \cdot \frac{\partial G}{\partial X^{(1)} * K^{(1)}} * K^{(1)^{\mathrm{T}}} \cdot \frac{\partial G}{\partial X^{(1)}} * K^{(0)^{\mathrm{T}}} \\ &= \frac{\partial f_c}{\partial X^{(l)}} \cdot \frac{\partial \mathrm{relu}}{\partial \mathrm{maxpool}} \cdot \frac{\partial \mathrm{maxpool}}{\partial X^{(l)}} * K^{(l-1)^{\mathrm{T}}} \cdot \ldots \cdot \frac{\partial \mathrm{relu}}{\partial \mathrm{maxpool}} \cdot \frac{\partial \mathrm{maxpool}}{\partial X^{(1)}} * K^{(0)^{\mathrm{T}}}\end{aligned} \tag{9}$$

The gradient of feature importance $f_l(X^0)$ in layer l:

$$\begin{aligned}\frac{\partial f_l(X^0)}{\partial X^0} &= \frac{\partial f_l}{\partial X^1} \cdot \frac{\partial G}{\partial X^1 * K^1} * K^{(1)^{\mathrm{T}}} \\ &= \frac{\partial f_l}{\partial X^{(l)}} \cdot \frac{\partial \mathrm{relu}}{\partial \mathrm{maxpool}} \cdot \frac{\partial \mathrm{maxpool}}{\partial X^{(l)}} * K^{(l-1)^{\mathrm{T}}} \cdot \ldots \cdot \frac{\partial \mathrm{relu}}{\partial \mathrm{maxpool}} \cdot \frac{\partial \mathrm{maxpool}}{\partial X^1} * K^{(0)^{\mathrm{T}}}\end{aligned} \tag{10}$$

where $K^{(l)\mathrm{T}}$ is the flip kernel (transpose of $K^{(l)}$) and $X^{(0)}$ is the input signal of a CNN.

This whole operation is similar to a backward propagation. In backward propagation, since the max-pooling operation is not differentiable, the "derivative" of max-pooling is obtained by recording the locations of the maxima within each pooling region in a set of switch variables. In Eq. (10), we just represent it as $\frac{\partial \mathrm{maxpool}}{\partial X^{(n)}}$.

The operation is shown in Fig. 3.

2.2.2 Visualization by DeconvNet

In DeconvNet, to reproject any feature map onto the input space, the feature map convolved from CNN is repeatedly processed with three operators: UnReLU, Unpooling, and Unconv.

1. Unpooling: Unpooling approximates the inverse pooling operation by recording the locations of the maxima within each pooling region in a set of switch variables. This operation corresponds to the max-pooling in backward propagation (Fig. 4).

   Thus, the unpooling operation can be expressed as follows:

$$\mathrm{unpooling}(X) = \frac{\partial \mathrm{maxpool}}{\partial X}$$

where X is the input of unpooling operation; in real case, it is always a feature map in one layer of CNN.

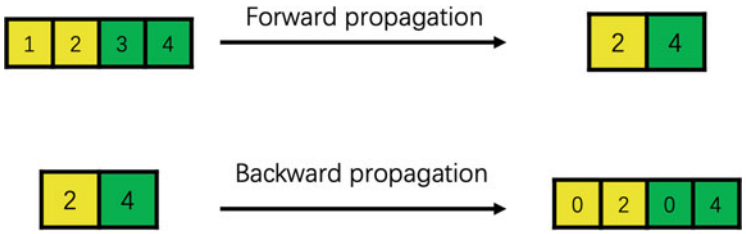

Fig. 3 Max-pooling backward propagation

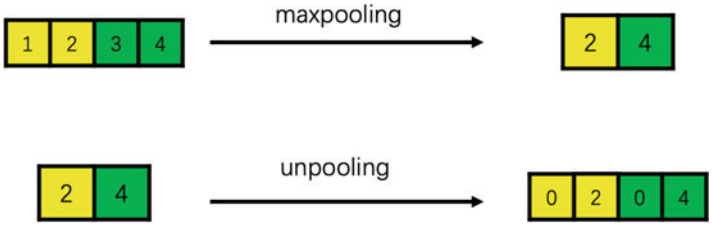

Fig. 4 Unpooling operation, which is corresponding to the max-pooling backward propagation

1. UnReLU: The ReLU function used in CNN rectifies the feature maps, thus ensuring the feature maps are always positive. To obtain valid feature reconstructions at each layer (which also should be positive), we pass the reconstructed signal through a ReLU function.
   Thus, the UnReLU operation is expressed as follows:

$$\text{UnReLU}(X) = \text{ReLU}(X)$$

2. Unconv: To invert a convolution operation in CNN, the DeconvNet applies transposed versions of the initial filters to the rectified maps. In practice, this means flipping each filter vertically and horizontally:

$$\text{conv}(X) = X * K$$

$$\text{unconv}(X) = X * K^{\text{T}}$$

   where $*$ is convolution operation, K is the filter during convolution, and K^{T} is the transpose filter of K.

In multilayer network, given an input signal $X^{(0)}$, the reconstruction R^l of the feature importance within layer l to the input space is expressed as in Eq. (11). This could visualize the feature importance of layer l in the input space. To better present the equation, we represent ReLU($\star$) as $\star$: ReLU.

$$R^l = X^{(l)} \vdots \text{ReLU} * \frac{\partial \text{maxpool}}{\partial \text{X}^{(\text{l})}} * \text{K}^{(\text{l}-1)^{\text{T}}} * \ldots \vdots \text{ReLU} * \frac{\partial \text{maxpool}}{\partial \text{X}^{(1)}} * \text{K}^{(0)^{\text{T}}} \quad (11)$$

We can show that the equation is similar to Eq. (10):

$$R^l \cong \frac{\partial f_l}{\partial X^{(0)}}$$

The contribution of each component to the classification result is expressed as Eq. (12):

$$R^n = f_{\text{c}}\left(X^{(0)}\right) \vdots \text{relu} \cdot \ldots \vdots \text{relu} \cdot \frac{\partial \text{maxpool}}{\partial \text{X}^{(1)}} * \text{K}^{(0)^{\text{T}}} \quad (12)$$

We can say that in DeconvNet, the expression of each input component contribution is very similar to the one in Eq. (10), the gradient of $f_{\text{c}}(X^{(0)})$.

$$R^n \cong \frac{\partial f_{\text{c}}}{\partial X^{(0)}} \quad (13)$$

Different from R^l which visualizes the layer feature importance and highlights which input part is important for CNN feature extraction, R^n focuses on the decision or classification result. Thus, R^n implies the contribution of each component of input signal $X^{(0)}$ to the classification result $f_{\text{c}}(X^{(0)})$, while R^l implies the feature importance of layer l backward to the input space.

2.3 Model

2.3.1 1-D CNN Model

Figure 5 provides details about the 1-D CNN architecture. This model only includes one convolutional layer. A 5000 × 1 signal is presented as the input. It is convolved with four different feature maps (blue block), each of size 35 by 1, using a stride of 1. The resulting feature maps are then: (1) passed through a rectified linear unit function (not shown) and (2) pooled (max within 2 × 1 regions, using stride 2). These four feature maps are then passed to a flatten layer (green block), which is able to transform multichannel features (in this case 4) into a 1-channel 10,000 × 1 vector that can be fed into a fully connected neural network classifier. This last network consists of two dense layers, taking features from the former layers, and results in the final 4 × 1 vector of classification scores.

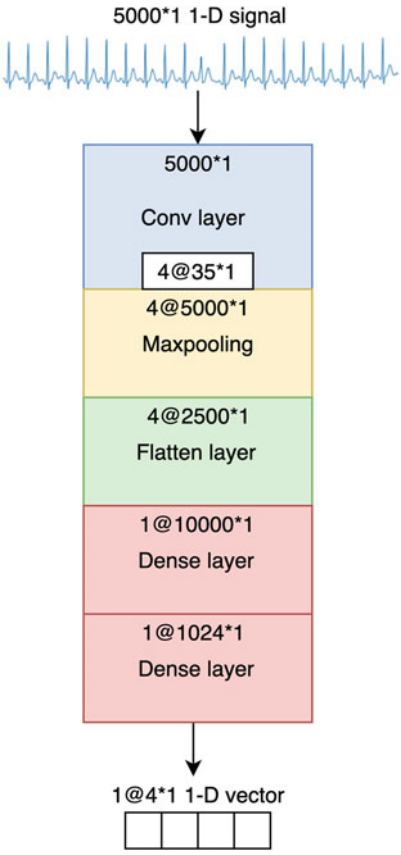

Fig. 5 1-D CNN architecture includes one convolutional layer and two dense layers (i.e., fully connected layers). A flatten layer is added before the first dense layer

2.3.2 1-D DeconvNet Model

Once the 1-D CNN training exposed in Sect. 2.3.1 is completed, a DeconvNet can be built to visualize the contribution of each input component to the classification results. Figure 6 illustrates the flowchart of 1-D CNN model and DeconvNet. A 1×5000 1-D signal passes through the trained 1-D CNN model (left). To visualize this signal, a DeconvNet (right) is attached to the CNN layer, providing a path to project the feature maps of CNN layer back onto the input space. The feature maps of the CNN layer are then successively processed by "UnReLU," "Unpooling," and "Unconv" to reconstruct the activity stored in the feature maps into the input space, and thus get a 1×5000 1-D signal as output of the DeconvNet.

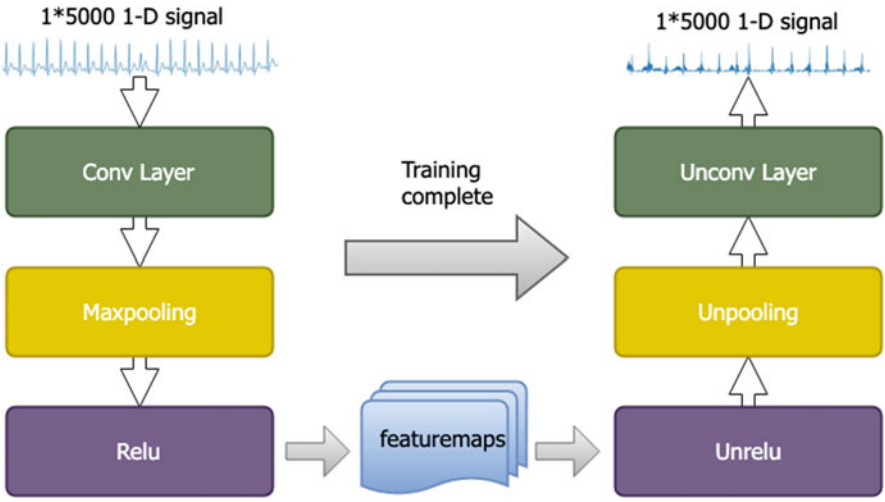

Fig. 6 1-D CNN (left) and DeconvNet (right) flowchart: feature maps are generated from the 1-D CNN and then used as the input of the DeconvNet. The input of CNN and the output of DeconvNet are both $1*5000$ vector

3 Experiment

3.1 Dataset

A wearable ECG database was collected using conductive textile dry electrodes embedded in the Wearable 12-lead ECG Lenovo SmartVest system. A self-charged ECG module was embedded in the back of SmartVest, which could start-up signal recording and implement hardware filtering, denoising, and amplifying (frequency band = 0.05–125 Hz, gain = 400). The ECG was sampled at 500 Hz and was stored locally in the remember card in ECG module, being transmitted to a connected smartphone via Bluetooth. Twenty volunteers aged 26–65 with a history of premature contractions participated in this study. The time length of each recording was 30 min. Wearable ECGs were also segmented into 10-s segments without overlapping. Signal quality was manually checked. Each R-peak location and its type (N, PAC or PVC) of the wearable ECG database were also annotated by two independent cardiologists and arbitrated by a third. The total number of beats was 3642, including 906 normal, 900 noise, 928 PAC, and 967 PVC beats.

3.2 Training Details of 1-D CNN Classification

We now describe the training details of the CNN model shown in Sect. 2.3. The network architecture is shown in Fig. 5. The implementation of the CNN model is

based on tensorflow. The model is trained on the dataset described in Sect. 3.1 (3642 ECG signals, four classes, including 906 normal, 900 noise, 928 PAC, and 967 PVC beats).

The dataset is split into five folds randomly. One fold was selected for validation set, and the other four folds are the training set. Stochastic gradient descent with a mini-batch size of 500 has been used to update the parameters, starting with a learning rate of 0.01. We anneal the learning rate throughout training manually when the loss on the validation converges. All weights are initialized to 0.01 and biases are set to 0. We stopped training after 200 epochs; however, it always converges before 20 epochs given filter size of 34 and max-pooling size of 2.

During training, dropout is used at each convolutional and dense layer, with a learning rate of 0.5. The filter size is set 35×1 for convolution and 2×1 for max-pooling. The overall accuracy is 96% on the validation set.

Figure 7 illustrates the relationship between the convergence speed and the max-pooling size. Larger max-pooling size will lead to a longer convergence time with the same accuracy. Figure 8 illustrates the relationship between the model stability and the filter size. A larger filter size will not change the convergence speed but will reduce the model stability.

3.3 Feature Visualization

3.3.1 Feature Importance Degree Heatmap (FIDH)

In this study, we first train a 1-D CNN for classification of ECG beat types on the wearable ECG database mentioned in Sect. 3.1. The classified ECG includes four different types of beats, which are shown in Fig. 3: (a) PAC (premature atrial contraction), (b) PVC (premature ventricular contraction), (c) normal beat, and (d) noise beat. All recordings are 10-s signal sampled at 500 Hz.

Once the training of the 1-D CNN model described in Sect. 2.3.1 is complete, we start to visualize the features of these four types of beats. Using the 1-D DeconvNet (cf. Sect. 2.3.2), a 5000×1 vector is then obtained as the model output. Each point illustrates the contribution of the corresponding input value in the original signal to the classification result. To better visualize the signals in Fig. 9 and to interpret the correlation between the contribution and the original signal, we plot the Feature Importance Degree Heatmap (FIDH), adding the degree of feature importance as colors at each point of the original signal.

Figure 10 shows FIDH, visualizing the degree of contributions at each point in Fig. 9. This means that samples of beats shown in Fig. 10 are the same as the samples in Fig. 9 selected from the dataset. Here, we only intercept 3.5- to 7-s recording for a clearer visualization, which contains the moments that PAC and PVC appear. Thus, the same four types of beats are included: PAC, PVC, normal beat, and noise beat. The degree of contribution to classification is represented as colors in Fig. 10. The red color means “the hidden weight” at this point is positive, and the blue color

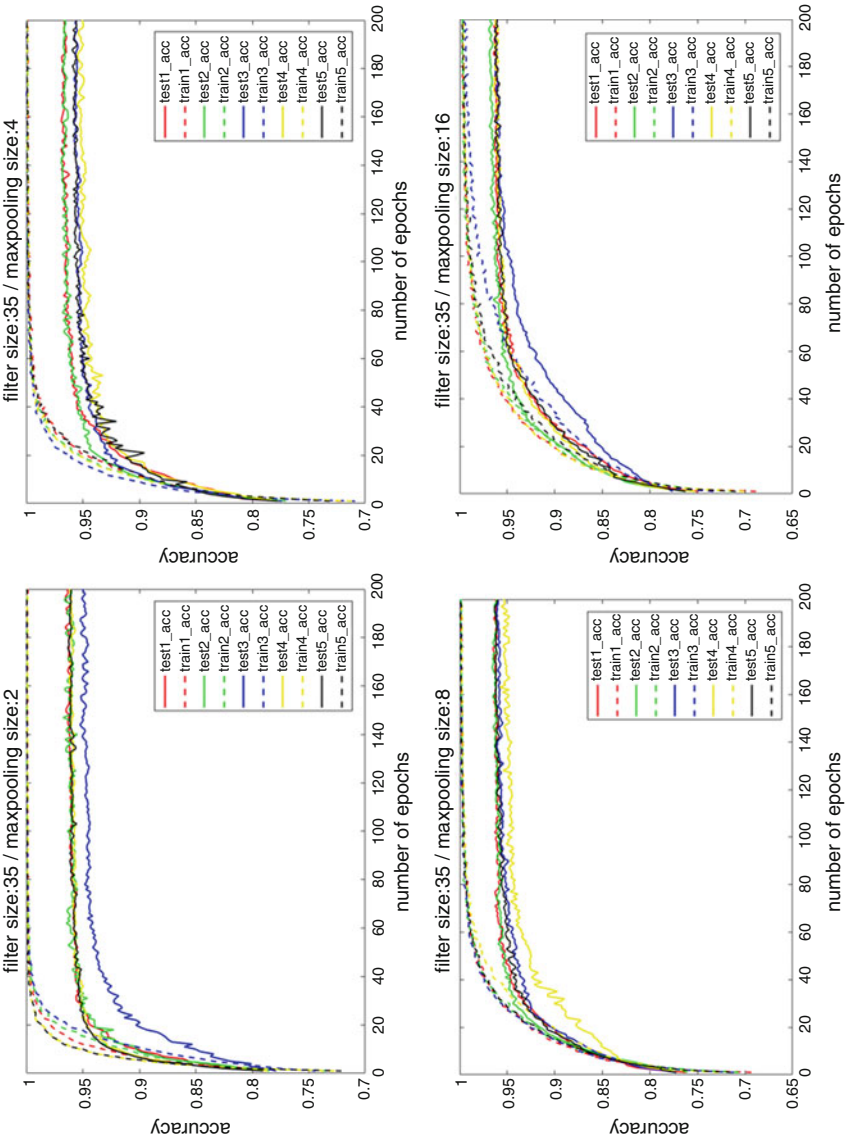

Fig. 7 Classification performance of 1-D CNN with different max-pooling size and unchanged filter size (35). The dotted line is the fivefold accuracy on training set, while the solid line represents the fivefold accuracy on test set

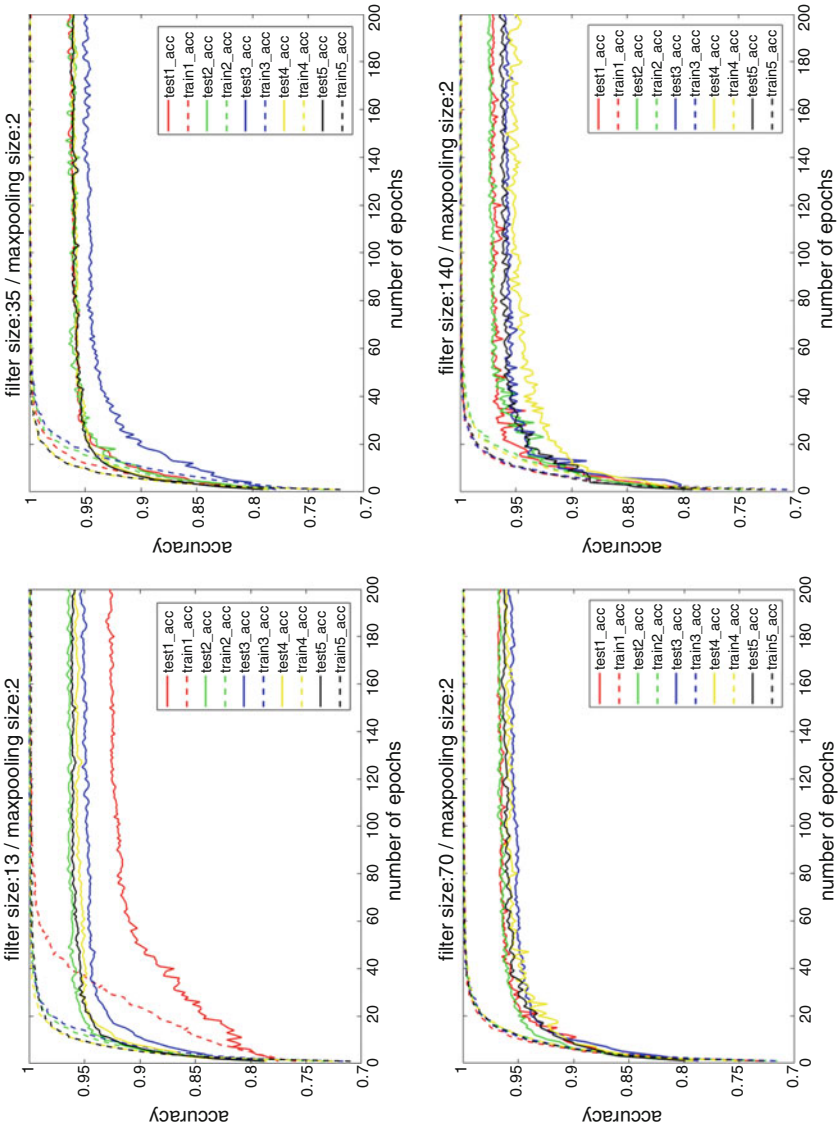

Fig. 8 Classification performance of 1-D CNN with different filter size and unchanged max-pooling size (35). The dotted line is the fivefold accuracy on training set, while the solid line represents the fivefold accuracy on test set

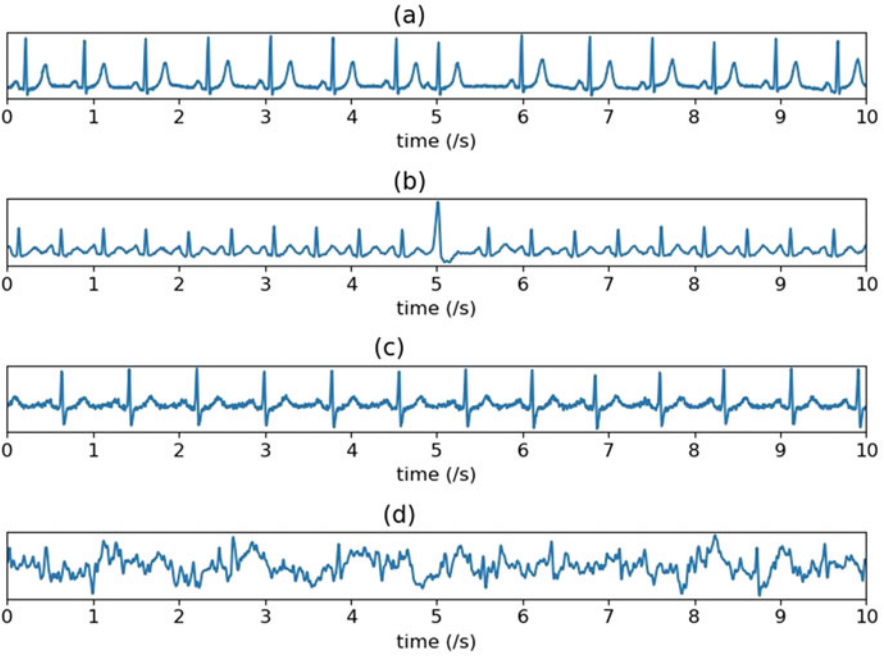

Fig. 9 ECG beats examples of PAC (**a**), PVC (**b**), normal (**c**), and noise (**d**)

means "the hidden weight" at this point is negative. The more intense the color is, the higher the absolute value of importance. Positive (red) or negative (blue) does not mean it will contribute to a positive or negative result. For example, in PVC case, the QRS wave at 5 s, which is colored in deeper red compared to other instant, is the most discriminant part during CNN feature extraction.

In Fig. 10, it is illustrated that only several parts of signal are discriminant for the CNN classifier:

1. In PAC, PVC, and normal beats, all R peaks contribute to a positive classification result. In PAC and normal case, the T waves are also slightly discriminant.
2. However, in PVC case, T waves are hardly highlighted. The R peak at 5 s becomes the most discriminant part, while other R peaks are uniformly less discriminant.
3. In noise case, all parts at high value are positively noticed. The lowest parts of every wave are almost blue (negatively noticed).

To further investigate the correlation between QRS waves, T waves, amplitude, and discrimination, we have selected three other PAC beat samples from the same patient but from different leads. Figure 11 illustrates the DCDH of the three PAC cases. Thus, for each recording, the PAC appears at the same time. Evidently, QRS waves are still discriminative in these three PAC cases with uniform degree. It is visible that T waves are comparably higher in case 1 than in case 2 and in case

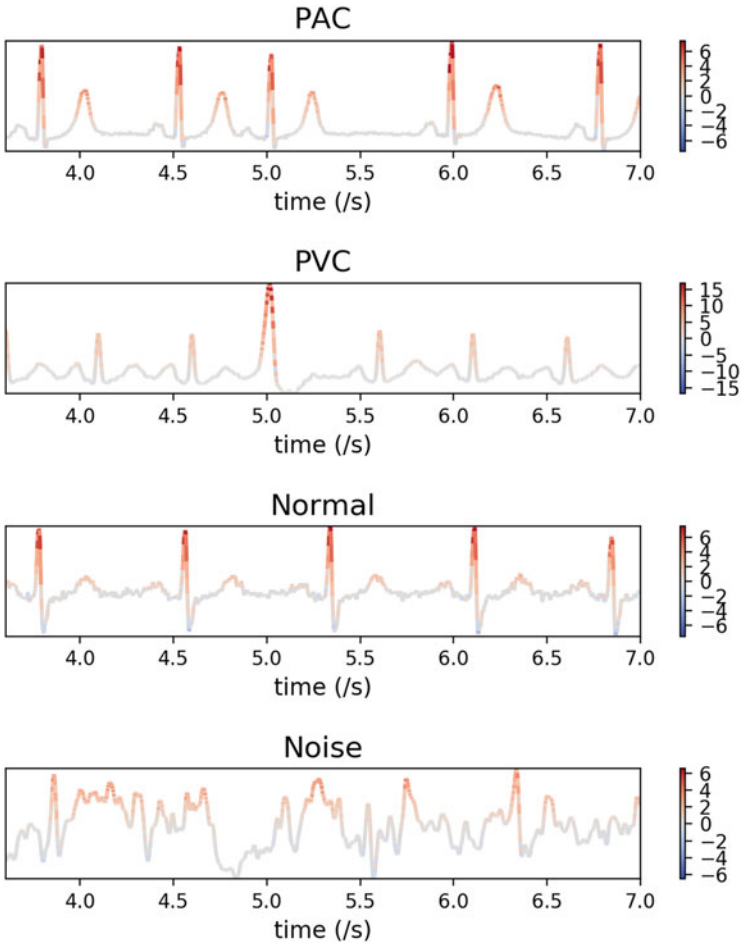

Fig. 10 Feature Importance Degree Heatmap (FIDH) of four types of signal. The samples in Fig. 9 are the input of 1-D CNN, and the generated feature maps are the input of DeconvNet. The color of each curve represents the output of DeconvNet. The red color means "the hidden weight" at this point is positive, and the blue color means "the hidden weight" at this point is negative. Color intensity corresponds to absolute value of importance

3. Accordingly, T waves in case 1 are significantly more discriminant than case 2 and case 3.

Similarly, Fig. 12 includes three cases of PVC selected from the same patient but from different leads.

1. In case 1, the QRS wave at 5 s is significantly higher than other QRS complexes, and thus, the PVC is easily noticeable.
2. In case 2 similarly as in case 1, the QRS wave at 5 s is discriminative. However, all amplitudes of QRS waves are negative values.

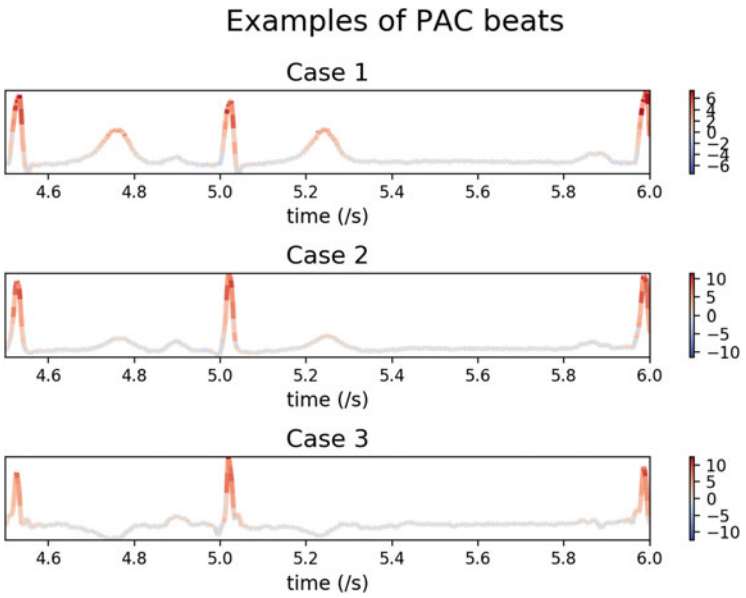

Fig. 11 Feature Importance Degree Heatmap (FIDH) of PAC beats. Three PAC beat samples are selected from the same patient but from different leads. For each recording, PAC appears at the same time (in this case, it is right after 5 s). The red color means "the hidden weight" at this point is positive, and the blue color means "the hidden weight" at this point is negative. The deeper the color is the higher absolute value of the contribution

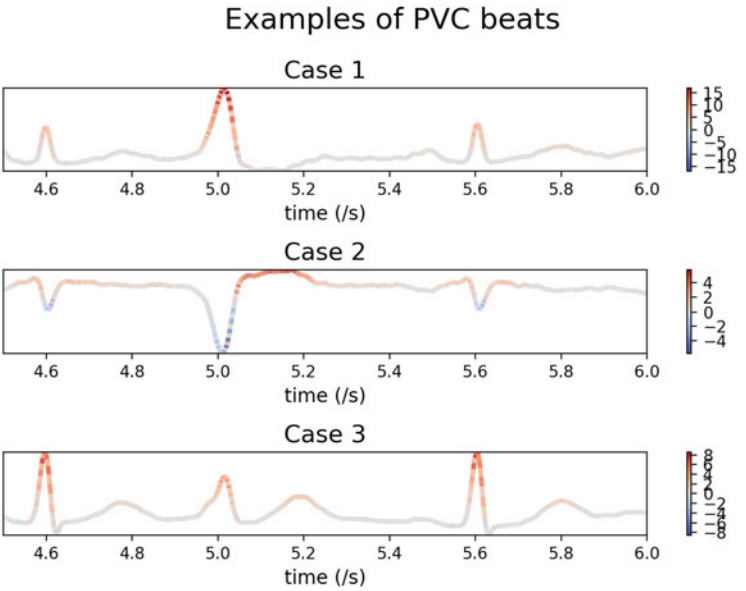

Fig. 12 Feature Importance Degree Heatmap (FIDH) of PVC beats. For each recording, PAC appears at the same time (in this case 5 s)

3. However, it is significantly different in case 3. Though the QRS wave at 5 s is still different from other QRS waves, the absolute amplitude of this wave is the lowest of all QRS waves. It has to be noticed that this case is correctly classified by the model.

4 Conclusion

In this chapter, we have investigated how interesting are the information, the deconvolution process first proposed by Zeiler and Fergus [17] can provide. In order to simplify the analysis, we restrict our study to a CNN of one convolutional layer and confirm that known standard ECG features can be retrieved. This study shows that main ECG features (such as QRS complexes) can be highlighted by CNN deconvolution process based on the absolute amplitude of ECG signal. The next step will be to extend this study to deeper networks such as the one described in the work by Hannun [20].

References

1. Rautaharju, P.M.: Eyewitness to history: landmarks in the development of computerized electrocardiography. J. Electrocardiol. **49**(1), 1–6 (2016)
2. Fukushima, K.: Neocognitron: a self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. Biol. Cybern. **36**(4), 193–202 (1980)
3. LeCun, Y., Bengio, Y., Hinton, G.: Deep learning. Nature. **521**(7553), 436 (2015)
4. Krizhevsky, A., Sutskever, I., Hinton, G.E.: Imagenet classification with deep convolutional neural networks. Adv. Neural Inf. Proces. Syst. **2012**, 1097–1105 (2012)
5. Miotto, R., Wang, F., Wang, S., Jiang, X., Dudley, J.T.: Deep learning for healthcare: review, opportunities and challenges. Brief. Bioinform. **19**(6), 1236–1246 (2017)
6. Oh, S.L., Ng, E.Y.K., Tan, R.S., Acharya, U.R.: Automated diagnosis of arrhythmia using combination of CNN and LSTM techniques with variable length heart beats. Comput. Biol. Med. **102**, 278–287 (2018)
7. Yıldırım, O., Pławiak, P., Tan, R.S., Acharya, U.R.: Arrhythmia detection using deep convolutional neural network with long duration ECG signals. Comput. Biol. Med. **102**, 411–420 (2018)
8. Rahhal, M.M.A., Bazi, Y., Zuair, M.A., Othman, E., Benjdira, B.: Convolutional neural networks for electrocardiogram classification. J. Med. Biol. Eng. **38**, 1014–1025 (2018)
9. Andersen, R.S., Peimankar, A., Puthusserypady, S.: A deep learning approach for real-time detection of atrial fibrillation. Expert Syst. Appl. **115**, 465–473 (2019)
10. Sannino, G., De Pietro, G.: A deep learning approach for ECG-based heartbeat classification for arrhythmia detection. Futur. Gener. Comput. Syst. **86**, 446–455 (2018)
11. Kiranyaz, S., Ince, T., Gabbouj, M.: Real-time patient-specific ECG classification by 1-D convolutional neural networks. IEEE Trans. Biomed. Eng. **63**(3), 664–675 (2015)
12. Goldberger, A.L., Amaral, L.A.N., Glass, L., Hausdorff, J.M., Ivanov, P.C., Mark, R.G., Mietus, J.E., Moody, G.B., Peng, C.-K., Stanley, H.E.: PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. Circulation. **101**(23), e215–e220 (2000)

13. Acharya, U.R., Oh, S.L., Hagiwara, Y., Tan, J.H., Adam, M., Gertych, A., Tan, R.S.: A deep convolutional neural network model to classify heartbeats. Comput. Biol. Med. **2017**(89), 389–396 (2017)
14. Acharya, U.R., Fujita, H., Oh, S.L., Hagiwara, Y., Tan, J.H., Adam, M.: Automated detection of arrhythmias using different intervals of tachycardia ECG segments with convolutional neural network. Inf. Sci. **2017**(405), 81–90 (2017)
15. Erhan, D., Bengio, Y., Courville, A., Vincent, P.: Visualizing higher-layer features of a deep network. University of Montreal. **1341**(3), 1 (2009)
16. Le, Q.V.: Building high-level features using large scale unsupervised learning. In: IEEE International Conference on Acoustics, Speech and Signal Processing, pp. 8595–8598 (2013)
17. Zeiler, M.D., Fergus, R.: Visualizing and understanding convolutional networks. In: European Conference on Computer Vision, pp. 818–833. Springer, Cham (2014)
18. Simonyan, K., Vedaldi, A., Zisserman, A.: Deep inside convolutional networks: visualising image classification models and saliency maps. arXiv preprint arXiv:1312.6034 (2013)
19. Baehrens, D., Schroeter, T., Harmeling, S., Kawanabe, M., Hansen, K., Müller, K.-R.: How to explain individual classification decisions. J. Mach. Learn. Res. **11**, 1803–1831 (2010)
20. Hannun, A.Y., Rajpurkar, P., Haghpanahi, M., Tison, G.H., Bourn, C., Turakhia, M.P., Ng, A. Y.: Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network. Nat. Med. **25**(1), 65–69 (2019)

Part V
Practical Applications

Atrial Fibrillation Detection in Dynamic Signals

Caiyun Ma, Shoushui Wei, and Chengyu Liu

Abstract Atrial fibrillation (AF) is the most common and sustained heart rhythm disorder, increasing the risk of stroke and death, and its incidence is destined to increase as the population ages. Current diagnostic methods are primarily through symptom or other indirect medical assessment methods. The fast-developing wearable technologies significantly promote the progress in ambulatory electrocardiogram (ECG) monitoring. This is a challenge to develop the devices that can detect AF in wearable electronic devices, with accessibility, sensitivity, ease of use, low-cost efficiency, and high computing power. Here, we first give a brief introduction to physiological concepts for development of detection algorithms. Then, we describe several kinds of AF features in dynamic signals. These features are important part of the automatic detection of AF, and a thorough understanding of these concepts can help researchers gain better insight into AF detection. Finally, seven AF features were extracted from the RR interval time series and were input into a SVM model to train AF/non-AF classifier. The results on the wearable ECGs verified that the proposed model could provide good identification for AF events.

Keywords Atrial fibrillation (AF) · Electrocardiogram (ECG) · Wearable · Support vector machines (SVMs)

C. Ma · S. Wei (✉)
The School of Control Science and Engineering, Shandong University, Jinan, China
e-mail: sswei@sdu.edu.cn

C. Liu (✉)
The State Key Laboratory of Bioelectronics, School of Instrument Science and Engineering, Southeast University, Nanjing, China
e-mail: chengyu@seu.edu.cn

C. Liu, J. Li (eds.), *Feature Engineering and Computational Intelligence in ECG Monitoring*, https://doi.org/10.1007/978-981-15-3824-7_10

1 Introduction

Atrial fibrillation (AF), a rapid and irregular fibrillation from the atrium, is a very important type of arrhythmia. According to statistics, the disease of AF affects approximately 1.5–2% of the world's total population, and this figure is likely to increase in the next 50 years [1, 2]. The prevalence of AF increases with age, from <0.5% at 40–50 years to 5–15% at 80 years [3]. In addition, AF can lead to stroke, heart failure, and sudden death, with a high morbidity and mortality [4, 5]. Therefore, the early detection and auxiliary diagnosis of AF have important clinical and social significance for improving patients' treatment strategies and the quality of treatment, reducing the incidence of critical illness and mortality.

The surface electrocardiogram (ECG) signal contained the high potential diagnostic information, and its characteristics directly reflect the nature of pathophysiologic events occurring in both the cardiac chambers. In addition, it is painless to record a surface ECG for the patient. Long-term surface ECG recordings can be performed with minimal risk compared to other invasive diagnostic techniques [6]. Therefore, ECG is a powerful tool to reveal initiation, maintenance, and termination of AF.

Recently, the fast-developing wearable and Internet of things (IoT) technologies significantly promote the progress in ambulatory electrocardiogram (ECG) monitoring, which is an essentially useful tool for the early detection of AF. However, when performed on the relatively noisy wearable ECGs, poor generalization capabilities are inevitable due to the individual waveform variability and external noises. Wearable electronic devices for ECG monitoring are usually highly sensitive to motion artifacts and susceptible to noise interference [7]. This is a challenge to develop AF detection that can be robust in noisy ECGs.

Our work describes physiological concepts behind the development of an AF detection algorithm and some of the most recent signal processing techniques to reveal atrial fibrillation. AF detectors can commonly follow ECG analysis to reveal the arrhythmia. The chapter is dedicated to an overview of some different algorithms to detect AF.

2 Physiological Concepts for Development of Detection Algorithms

During AF, ECG has obvious features: P-wave disappears, f-waves (a series of continuous and irregular atrial excitation waves) appear [8], and RR interval is absolutely irregular [9]. Figure 1 shows the AF episode and normal rhythm episode. In literature, AF detectors can be separated into two major classes: methods based on P-wave features and RR interval features.

Many scholars analyzed the morphology of P-wave to achieve AF detection. Andrikopoulos et al. [10] presented that increased variance of P-wave duration on

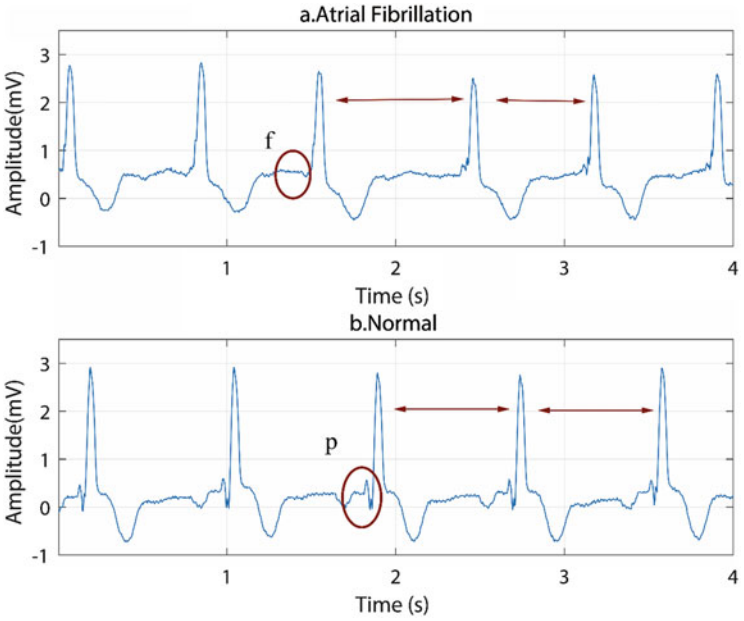

Fig. 1 (**a**, **b**) Atrial fibrillation episode and normal rhythm episode

the electrocardiogram distinguishes patients with idiopathic paroxysmal atrial fibrillation (PAF). Prerfellner et al. [11] used P-wave evidence as a method for improving algorithm to detect AF. Alcaraz et al. [12] proposed a new approach to predict termination of AF: wavelet sample entropy. Alcaraz et al. [13] applied wavelet entropy to AF prediction. Garcia et al. [14] presented application of the relative wavelet energy to heart rate independent detection of AF. However, P-waves or f-waves in atrial activity have small amplitude, which are extremely susceptible to noise interference. This situation can be worse in the dynamic ECG monitoring, where many complicated interferences from the daily activities occur, and is difficult to obtain stable, high-quality signals in real-time long-term recordings.

There are many AF detection algorithms based on RR interval features, including variability analysis, complexity estimation, statistical method, and entropy estimation. The R peak is the most prominent feature of an ECG and the least confounded by muscle noise. Indeed, methods based on RR irregularity should be preferred for external devices. These methods need a high accurate QRS detector, since extra and missed beats would affect algorithms' performance [15]. For the wearable ECG detection, poor generalization capabilities are inevitable due to the individual waveform variability and external noises. This is a challenge to develop AF detection that can be robust in the wearable ECG. This chapter introduces some RR interval features for atrial fibrillation detection algorithm and provides guidance for the identification of wearable ECG signals.

It is worth noting that AF detector combining of several of the above single features with machine learning algorithms could enhance its performance. Couceiro et al. [16] combined the P-wave disappearance, irregular heart rhythm, atrial activity, and other ECG characteristics and developed a neural network model on the MIT-BIH AF database, achieving a sensitivity of 93.8% and a specificity of 96.1%. Babaeizadeh et al. [17] presented a AF detector using decision tree classifier and RR interval, PR interval variability, and a P-wave morphology similarity measure, reporting a sensitivity of 92% and a positive predictivity of 97%. The 2017 PhysioNet/CinC Challenge [18] aims to classify normal sinus rhythm, AF, an alternative rhythm, or noisy ECGs. Many contestants have developed AF detectors based on RR interval features and analysis of the absence of P-waves or f-waves present in TQ interval. Shreyasi Datta et al. [19] used morphological features, frequency features, heart rate variability (HRV) features, statistical features, and other abnormality features with a multilayer cascaded binary classification approach in the PhysioNet/Computing in Cardiology (CinC) Challenge 2017 and won shared 1st places.

3 AF Features in Dynamic Signals: Description and Comparison

This section summarizes the state of the art in AF detection based on RR interval features, that is because pulse beats of ventricles are less likely to be influenced by baseline wandering and noise. In addition, since we cannot predict when the paroxysm of AF will come about, it will be useful to make a real-time portable monitoring electrocardiograph. Some AF features based on RR interval are introduced, and the underlying principle to reveal atrial fibrillation is briefly discussed.

3.1 Lorenz Plot

The characteristics of the Lorentz scatter plot can calculate atrial fibrillation [20]. The RR interval of atrial fibrillation signal is irregular, and the distribution of scatter plots is significantly different from that of normal people. Figure 2 shows Lorenz plot of ΔRR intervals for normal signal and AF signal from the recording of 04936 from the MIT-BIH AF database. x axis is ΔRR interval, and y axis is ΔRR $(i-1)$ interval.

$$\Delta \mathrm{RR} = \mathrm{RR}(i) - \mathrm{RR}(i-1)$$

$$\Delta \mathrm{RR}(i-1) = \mathrm{RR}(i-1) - \mathrm{RR}(i-2)$$

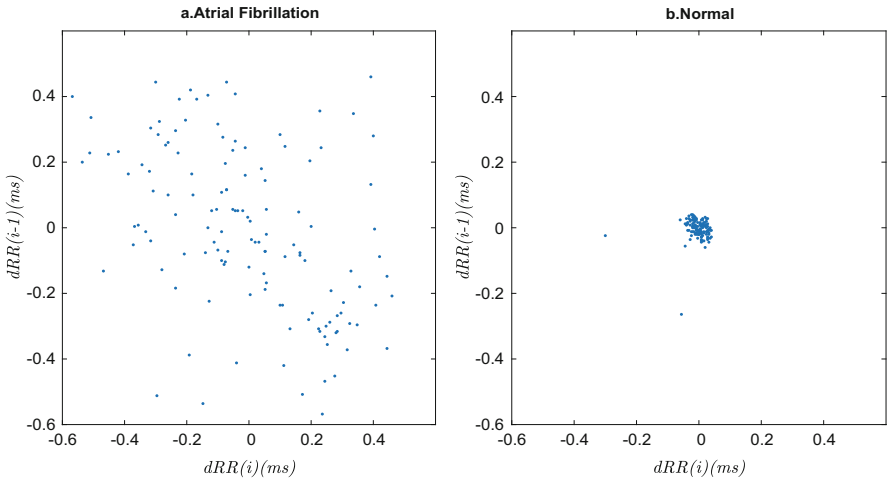

Fig. 2 (**a**, **b**) Lorentz plot of ΔRR interval for AF signal and normal signal

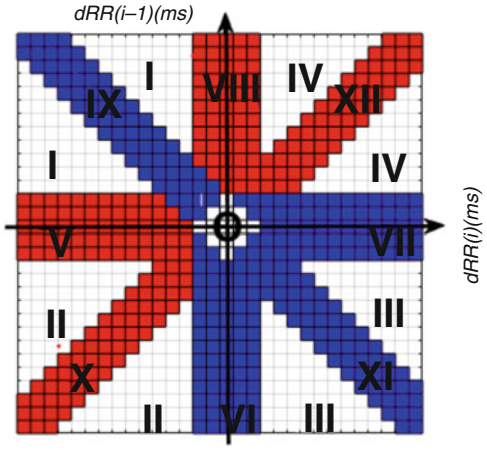

Fig. 3 The two-dimensional histogram, numeric representation of a Lorenz plot

Since the continuous RR interval difference of the normal signal is small, it can be seen that the points of the normal signal are gathered around the starting point, and the atrial fibrillation signal is sparsely distributed due to the irregular RR interval from Fig. 2.

If we consider the digital representation of the Lorentz plot, such as a two-dimensional histogram, it is divided into 13 discrete segments, shown in Fig. 3. For each signal's rhythm, its points have a higher probability of positioning in a particular subdivision. For a normal signal, the point is almost concentrated near the origin.

Therefore, the origin of the histogram, represented by "O," will contain a large number of points, while the boxes in other sections are almost empty. Conversely,

for atrial fibrillation signals, the points will spread across the histogram, filling a larger number of intervals from "I" to "VII."

AFEvidence, which accounts for points positioning within the plot, is a test indicator, used to quantify the possibility of atrial fibrillation. IrregularityEvidence is used to measure the sparsity of a high-value distribution of atrial fibrillation signals. BinCount_n (BC_n) is the number of nonempty bins contained in the nth segment of Fig. 3. PACEv is used to measure the low value of the normal signal, which is corresponding to the number of all the dots filled in the "O" segment in Fig. 3.

$$\text{IrregularityEvidence} = \sum_{1}^{12} \text{BinCount}_n \tag{1}$$

$$\text{AFEvidence} = \text{IrrEv} - \text{OriginCount} - 2 * \text{PACEv} \tag{2}$$

$$\text{PACEv} = \sum_{n=1}^{4} (\text{PC}_n - \text{BC}_n) + \sum_{n=5,6,10} (\text{PC}_n - \text{BC}_n) + \sum_{n=7,8,12} (\text{PC}_n - \text{BC}_n) \tag{3}$$

where PC_n is the number of points contained in the nth segment of Fig. 3.

The method was evaluated on the MIT-BIH AF database, showing sensitivity of 97.5% and positive predictive value of 99%. In the large number of atrial tachycardia (AT) presence, detector performance is worsened. This condition needs to design a supplemental detector to distinguish AT with regular ventricular response from AF. Through clinical testing, this algorithm is eligible for a further implementation on phone by clinical truth. The optimal window size is 2 min at least, and minimum percentage of AF in AF ECG episode is currently 60%.

3.2 Poincare Plot

Poincare plot [21] from non-AF data showed some pattern, while the plot from AF data showed irregular shape. Figure 4 shows Poincare plot of RR interval for AF signal and normal signal. In case of non-AF data, the definite pattern in the plot manifested itself with some limited number of clusters or closely packed one cluster. In case of AF data, the number of clusters in the plot was too many. Making a Poincare plot using the inter-beat intervals, the author extracted three-feature measures characterizing AF and non-AF: the number of clusters, mean stepping increment of inter-beat intervals, and dispersion of the points around a diagonal line in the plot. The author divided distribution of the number of clusters into two and calculated mean value of the lower part by k-means clustering method and classified data whose number of clusters is more than one and less than this mean value as non-AF data. In the other case, the author tried to discriminate AF from non-AF using

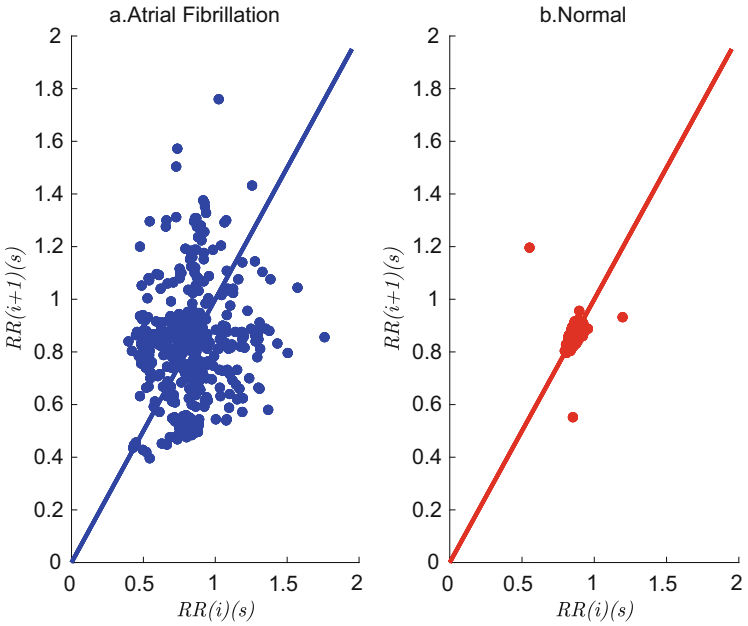

Fig. 4 (**a**, **b**) Poincare plot of RR interval for AF signal and normal signal

support vector machine with the other feature measures: the mean stepping increment and dispersion of the points in the Poincare plot.

The study [21] evaluated the accuracy using leave-one-out cross-validation on Computers in Cardiology Challenge 2001 and 2004. Mean sensitivity and mean specificity were 91.4% and 92.9%, respectively. It could be installed in a portable heart monitoring system. This AF detector was designed as automated algorithm, which did not require any human intervention and any specific threshold and could be installed in a portable AF monitoring system.

3.3 RR Interval Variance

This AF detector is simply based on RR interval variance and is designed to provide an automatic, robust detection of AF [22]. Evaluated by the authors on the MIT-BIH AF database, it showed sensitivity of 87.30% and specificity of 90.31% with an optimal window size of 120 s. This AF detector has been conceived for ambulatory monitoring situations, where arbitrary lead placements, muscle artifact, and potentially changing morphology of the signal can represent a challenge for an AF detector.

3.4 The Median of the Variation in the Absolute Standard Deviation from Mean of Heart Rate in Three Adjacent Segments of the RR Interval Series (MAD)

MAD, defined as the median of the variation in the absolute standard deviation from mean of heart rate in three adjacent segments of the RR interval series, reveals the irregular nature of AF [23]. It was developed for long-term monitoring of AF on a portable monitoring device. In the comparative study [24], this method showed the highest sensitivity and the smallest window length (10 s), which can help in detecting additional AF cases, such as paroxysmal events, whose onset is often unexpected and of short duration and is developed to be easy-to-implement, simple, and to have low-memory requirements.

3.5 Density Histogram of Delta RR Intervals

Huang et al. proposed a novel method for detection of the transition between AF and sinus rhythm based on RR intervals [25]. First, we obtained the delta RR interval distribution difference curve from the density histogram of delta RR intervals and then detected its peaks, which represented the AF events. Once an AF event was detected, four successive steps were used to classify its type, and thus to determine the boundary of AF: (1) histogram analysis, (2) standard deviation analysis, (3) numbering aberrant rhythms recognition, and (4) Kolmogorov–Smirnov (K-S) test.

A dataset of 24-h Holter ECG recordings ($n = 433$) and two MIT-BIH databases (MIT-BIH AF database and MIT-BIH normal sinus rhythm (NSR) database) were used for development and evaluation. Using the receiver operating characteristic curves for determining the threshold of the K-S test, the authors have achieved the highest performance of sensitivity and specificity (Sp) (96.1% and 98.1%, respectively) for the MIT-BIH AF database, compared with other previously published algorithms. The Sp was 97.9% for the MIT-BIH NSR database. The algorithm has been integrated into a Holter system for the automatic detection of AF, and it is also suitable for applying to the continuous AF monitoring situations.

3.6 Coefficient of Sample Entropy (CosEn)

The coefficient of sample entropy (CosEn) is used to distinguish AF and atrial flutter (AFL) from sinus rhythm and other arrhythmias, which is an optimized combination [26]. It includes sample entropy (SampEn) [27] and is able to encode the irregular nature of short RR interval segments during AF and mean heart beat interval (RR), which adds further independent information to the discrimination. Refer Appendix 6.1 for specific calculation process.

This method used a very small window length (only a 12-beat segment). It was validated on MIT-BIH database and achieved sensitivity of 91% and specificity of 94%. But, it had poor performance on Holter monitoring recordings from the University of Virginia (UVa), which is because of recordings with frequent, complex ventricular ectopy or electronic pacemakers which challenge AF identification.

3.7 Normalized Fuzzy Entropy (NFEn)

On the basis of SampEn, Liu et al. developed the FuzzyMEn method by using fuzzy function instead of 0–1 judgment rule [28]. Then, combined with CosEn and FuzzyMEn, normalized fuzzy entropy (NFEn), a novel entropy measure suitable for AF detection based on a short-term RR time series, was proposed again [29]. NFEn uses a fuzzy function to determine vector similarity, replaces a probability estimate with a density estimate for entropy approximation, utilizes a flexible distance threshold parameter, and adjusts for heart rate by subtracting the natural log of mean RR intervals.

NFEn was tested on the MIT-BIH AF, NSR, and arrhythmia databases, demonstrating that NFEn is an accurate measure for detecting AF. For classifying AF and non-AF rhythms, NFEn achieved the highest area under receiver operating characteristic curve (AUC) values of 92.72%, 95.27%, and 96.76% for 12-beat, 30-beat, and 60-beat window lengths, respectively.

3.8 Entropy_AF

Zhao et al. [30] proposed the Entropy_AF method to enhance the performance of entropy-based AF detectors. This algorithm combines the distance normalization function and the entropy-based AF detection concept and uses the flexible threshold parameters. Refer Appendix 6.2 for specific calculation process.

On the MIT-BIH AF database, Entropy_AF achieved the highest area under receiver operating characteristic curve (AUC) values of 98.15% when using a 30-beat time window, which was higher than CosEn with AUC of 91.86%. For classifying AF and non-AF rhythms in the clinical wearable AF database, Entropy_AF also generated the largest values of Youden index (77.94%), sensitivity (92.77%), specificity (85.17%), accuracy (87.10%), positive predictivity (68.09%), and negative predictivity (97.18%). Entropy_AF generated highest classification accuracy when using a 12-beat time window and the better discrimination ability for identifying AF when using Entropy_AF method, indicating that it would be useful for the practical clinical wearable AF scanning.

4 Atrial Fibrillation Detection in Dynamic Signals Based on RR Interval Characteristics

Detectors designed are based on the ventricular response analysis. Starting from the ECG, the RR interval series is derived. Seven AF features were extracted from the RR interval time series. AF features were input classifier to return AF diagnosis. The steps for AF detection are shown in Fig. 5.

4.1 Database

The database is wearable ECG data. The wearable ECG database was collected using a wearable ECG device developed by Southeast University and Lenovo, as shown in Fig. 6 [31]. The patients were recruited from the First Affiliated Hospital of Nanjing Medical University and had been diagnosed as AF by ECG Holter. The study protocol was approved by the Ethics Committee of the First Affiliated Hospital of Nanjing Medical University, and the patient has signed the informed consent form. ECGs were sampled as 400 Hz. In our study, we selected ten normal and ten patients (randomly select 1000 ECG for each person) respectively to tenfold cross-validation. In addition, we selected eight persons (24-h ECG) as testing sets. Table 1 shows details of the dataset.

4.2 QRS Detection

For the RR interval feature extraction, the most important thing is to identify the position of the R peak. In recent decades, QRS detection technology has developed

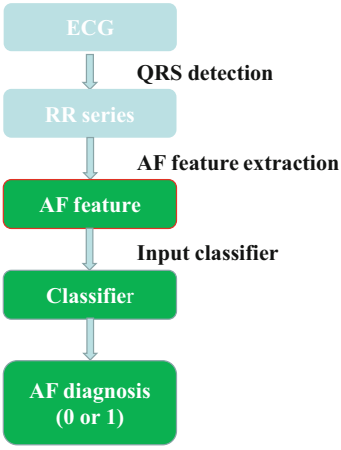

Fig. 5 Steps for AF detection

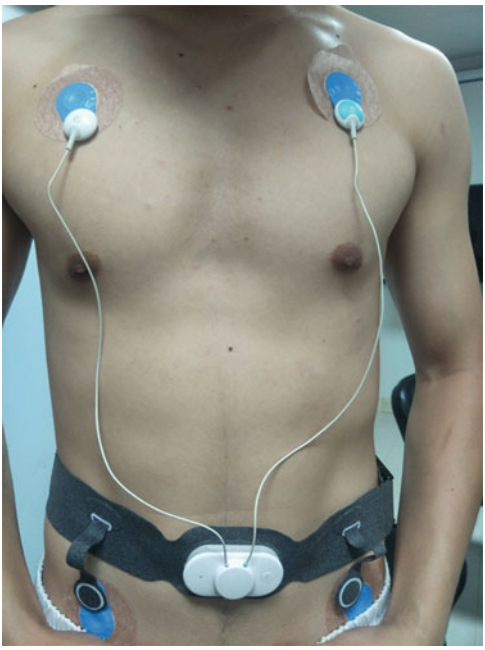

Fig. 6 The wearable ECG device

Table 1 Details of data

Data	60-s ECG episode	30-s ECG episode	10-s ECG episode
Training set	20,000	20,000	20,000
Testing set1	16,313	17,175	23,062
Testing set2	16,688	19,872	15,443
Testing set3	15,738	16,158	16,483
Testing set4	15,744	17,044	27,352
Testing set5	15,582	16,389	17,042
Testing set6	15,366	15,768	16,188
Testing set7	15,785	19,292	15,245
Testing set8	16,016	15,222	16,703

relatively maturely. In this chapter, a fast QRS detection algorithm proposed by Paoletti et al. [15] was used to locate QRS complex waves. This algorithm has a certain anti-noise property and is suitable for the R-peak positioning of dynamic ECG signals [32]. The R-peak location of the 10-s dynamic signal labeled by this QRS detector is shown in Fig. 7.

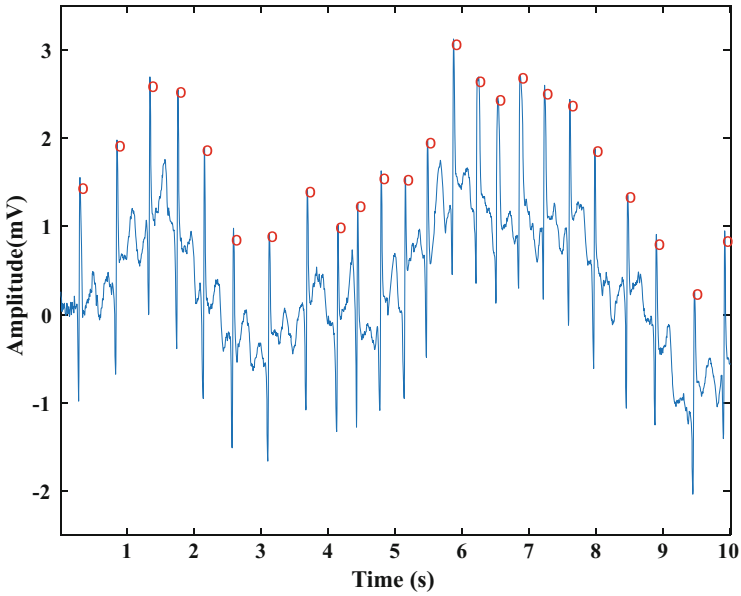

Fig. 7 The R-peak location of the 10-s dynamic signal

4.3 AF Features

The previous section reviews the technology in atrial fibrillation detection based on RR interval features, with particular emphasis on AF screening applications that can be implemented in dynamic signals. We chose seven AF features: Entropy_AF, sample entropy (SampEn), coefficient of sample entropy (CosEn), mean RR intervals of episode (mRR), minimum of heart rate of episode (minHR), maximum of heart rate of episode (maxHR), and median heart rate of episode (medHR) for further evaluation and comparison.

4.4 Support Vector Machine

LIBSVM with the Gaussian kernel function was used as the classifier in Matlab R2017b. Grid search method [33] was used for parameter optimization.

4.5 Evaluation Methods

To get reliable and stable model, we used tenfold cross-validation on 20 people's data on the wearable ECG. In order to verify the generalization ability and

practicability of the model, we used eight people's data as the test data. For model evaluation, three widely used metrics, i.e., accuracy (Acc), sensitivity (Se), specificity (Sp), are used as evaluation indicators. According to the attribute of the label (positive or negative), the result can generate four basic indexes: true positive (TP), false positive (FP), true negative (TN), and false negative (FN). In this case, Acc is the radio of the number of correct predicted labels and total number of the labels, thus Acc = (TP + TN)/(TP + TN + FP + FN). Se is the true positive rate and is probability of incorrectly diagnosing into positive among all positive patients, Se = TP/(TP + FN). Sp is proportion of incorrectly diagnosing into negative among all negative patients, Sp = TN/(TN + FP).

4.6 Results

Table 2 shows the classification results from tenfold cross-validation on the wearable ECG database. The mean and standard deviation (SD) of the experimental results were selected to be evaluated. The result showed an Acc of 95.20% for 10-s episode (Se 93.98% and Sp 97.62%), an Acc of 97.47% for 30-s episode (Se 97.25% and Sp 98.07%), and an Acc of 98.41% for 60-s episode (Se 98.76% and Sp 98.38%).

The combination of the second fold and the seventh fold is the atrial fibrillation signal and the ventricular premature. For 10-s episode, Se is relatively low, that is, the classifier classifies a part of ventricular premature signal into atrial fibrillation signal. But, for 30-s episode and 60-s episode, tenfold cross-validation results are better.

Twenty people's data on the wearable ECG were as training sets, and eight individuals were tested. Table 3 shows the results on the eight wearable ECG data. Testing set2 is a patient with no atrial fibrillation and more atrial and ventricular premature. The test accuracy of 10-s episode and 30-s episode is very low (60.17% and 79.51%, respectively), and the Acc of 60-s episode is 84.56%. Testing set6 is a patient with no atrial fibrillation and atrial and ventricular premature. The testing accuracy of 10-s episode is lower (78.79%), but the Acc is 94.12% of 30-s episode and 97.50% of 60-s episode. For the other testing sets, Acc of 10-s ECG episode is over 90%, Acc of 30-s ECG episode is over 96%, and Acc of 60-s ECG episode is over 98%.

In our purposed algorithm, Acc of 60-s window is the best, but Acc of 10-s window is low on patients with atrial and ventricular premature. In the following model training, the volume and diversity of training sets should be increased. Additional classification rules can be developed for the ECG signals of atrial and ventricular premature.

Table 2 Classification results from tenfold cross-validation on the wearable ECG database

	60-s ECG episode			30-s ECG episode			10-s ECG episode		
Fold	Acc (%)	Se (%)	Sp (%)	Acc (%)	Se (%)	Sp (%)	Acc (%)	Se (%)	Sp (%)
1	100.00	100.00	100.00	99.65	100.00	99.30	98.25	100.00	96.62
2	98.60	97.28	100.00	94.15	89.88	99.44	83.50	75.38	99.26
3	99.55	100.00	99.11	99.20	100.00	98.43	97.40	99.07	95.84
4	96.90	94.16	100.00	95.35	91.64	99.78	93.70	89.58	98.77
5	99.30	100.00	98.62	98.20	99.90	96.62	97.00	99.37	94.85
6	99.95	99.90	100.00	99.60	99.70	99.50	97.90	98.68	97.14
7	98.55	97.27	99.89	96.15	93.25	99.46	90.00	83.90	98.78
8	99.50	99.01	100.00	99.05	98.23	99.90	97.95	96.24	99.79
9	91.95	100.00	86.60	93.60	100.00	88.64	97.80	99.69	96.05
10	99.80	100.00	99.61	99.75	99.90	99.60	98.50	97.92	99.09
Mean	98.41	98.76	98.38	97.47	97.25	98.07	95.20	93.98	97.62
SD	2.454	1.961	4.167	2.421	4.021	3.450	4.900	8.390	1.724

Table 3 The results on the eight wearable ECG data

	60-s ECG episode	30-s ECG episode	10-s ECG episode
Testing sets	Acc (%)	Acc (%)	Acc (%)
Testing set1	100.00	99.97	99.83
Testing set2	84.56	79.51	60.17
Testing set3	99.77	99.16	91.92
Testing set4	98.68	96.01	84.54
Testing set5	98.07	97.14	91.80
Testing set6	97.50	94.12	78.79
Testing set7	99.60	98.24	90.61
Testing set8	99.91	98.96	92.19

5 Conclusions

In this section, eight types of short-term features that can be used for dynamic AF signal detection are described. Particular emphasis is placed on AF screening applications that can be implemented in dynamic signals. Finally, seven AF features were selected to train AF/non-AF model, and the detection effect of the algorithm on the different window lengths of wearable ECG was evaluated. The testing results on wearable ECG show that this is a feasible method for AF detection in dynamic signals.

6. Appendix

6.1. Coefficient of Sample Entropy (CosEn)

A data record consists of a series of N consecutive RR intervals: $x_1, x_2, \ldots, x_n$. For a length $m < n$ and starting point i, the template $x_m(i)$ is the vector containing m consecutive RR intervals $x_i, x_{i+1}, \ldots, x_{i+m-1}$. For a matching tolerance $r > 0$, an instance where all the components of $x_m(i)$ are within a distance r of any other $x_m(j)$ in the record is called a match (or template match). For example, the template x_1 matches x_2, if both $|x_1 - x_3| < r$ and $|x_2 - x_4| < r$. Let B_i denote the number of matches of length m with template $x_m(i)$ and A_i denote the number of matches of length $m + 1$ with template $x_{m+1}(i)$. Let $A = \Sigma_i A_i$, $B = \Sigma_i B_i$ denote, respectively, the total number of matches of length $m + 1$ and m. The sample entropy which refers to the negative natural logarithm of the conditional probability that two short templates with matching length m will continue to match at the next point within any tolerance r.

$$\text{SampEn} = -\ln(A/B) - \ln(B) - \ln(A) \tag{4}$$

where A is the total number of matches with length $m + 1$ and B is the total number of matches of length m. The initial value of the matching tolerance r is 30 ms, but it is allowed to increase until the minimum number of molecules (A) has been completed.

CosEn was proposed by Lake et al., which is a modified version of the sample entropy of a short-time window. It takes into account the length of the entire tolerance window ($2r$) along with the current heart rate, which is defined as follows:

$$\mathrm{CosEn} = \mathrm{SampEn} - \ln(2r) - \ln\left(\overline{\mathrm{RR}}\right) \tag{5}$$

where $\overline{\mathrm{RR}}$ is the mean of the RR interval.

6.2. Entropy_AF

For an RR time series $x(i)$ $(1 \leq i \leq N)$, form the vector sequences X_i^m $(1 \leq i \leq N - m)$:

$$X_i^m = \{x(i), x(i+1), \cdots, x(i+m-1)\} \tag{6}$$

where the vector X_i^m represents m consecutive $x(i)$.

The distance between vector sequences X_i^m and X_j^m was defined as follows:

$$\begin{aligned} \mathrm{d}X_{i,j}^m &= \mathrm{d}\left[X_i^m, X_j^m\right] \\ &= \frac{\max\limits_{0\leq k\leq m-1}|x(i+k) - x(j+k)| - \min\limits_{0\leq k\leq m-1}|x(i+k) - x(j+k)|}{\max\limits_{0\leq k\leq m-1}|x(i+k) - x(j+k)| + \min\limits_{0\leq k\leq m-1}|x(i+k) - x(j+k)| + \varepsilon} \end{aligned} \tag{7}$$

where ε is a small positive number to avoid the possible denominator of 0.

Then, calculate the similarity degree $\mathrm{DX}_{i,j}^m(n, r)$ between the vectors X_i^m and X_j^m by a fuzzy function $\mathrm{uX}\left(\mathrm{d}X_{i,j}^m, n, r\right)$ defined as follows:

$$\mathrm{DX}_{i,j}^m(n, r) = \mathrm{uX}\left(\mathrm{d}X_{i,j}^m, n, r\right) = \exp\left(-\frac{\left(\mathrm{d}X_{i,j}^m\right)^n}{r}\right) \tag{8}$$

where n is the similarity weight and r is the flexible tolerance threshold.

Then, define the functions $\mathrm{BX}^m(n,r)$ as follows:

$$\mathrm{BX}^m(n,r) = \frac{1}{N-m}\sum_{i=1}^{N-m}\left(\frac{1}{N-m}\sum_{j=1}^{N-m}\mathrm{DX}_{i,j}^m(n,r)\right) \tag{9}$$

$\mathrm{BX}^m(n,r)$ measures the mean similarity degrees for the vectors at dimension m. Similarly, we define the functions of mean similarity degrees $\mathrm{AX}^{m+1}(n,r)$ for dimension $m + 1$:

$$\mathrm{AX}^{m+1}(n,r) = \frac{1}{N-m}\sum_{i=1}^{N-m}\left(\frac{1}{N-m}\sum_{j=1}^{N-m}\mathrm{DX}_{i,j}^{m+1}(n,r)\right) \tag{10}$$

Then, we use a density-based estimation, rather than probability-based estimation, to generate a quadratic fuzzy entropy using the volume of each matching region, i.e., $(2r)^m$:

$$\begin{aligned}\mathrm{Entropy}_{\mathrm{AF}} &= -\ln\left(\frac{\mathrm{AX}^{m+1}(n,r)/(2r)^{m+1}}{\mathrm{BX}^m(n,r)/(2r)^m}\right)\\ &= -\ln\left(\frac{\mathrm{AX}^{m+1}(n,r)}{\mathrm{BX}^m(n,r)}\right) + \ln(2r)\end{aligned} \tag{11}$$

Subtract the natural log of mean RR interval as follows:

$$\mathrm{Entropy}_{\mathrm{AF}} = -\ln\left(\frac{\mathrm{AX}^{m+1}(n,r)}{\mathrm{BX}^m(n,r)}\right) + \ln(2r) - \ln(\mathrm{RR}_{\mathrm{mean}}) \tag{12}$$

where $\mathrm{RR}_{\mathrm{mean}}$ is the mean of RR intervals in the current RR episode. $\mathrm{RR}_{\mathrm{mean}}$ is expressed in unit of seconds.

As shown in Eq. (12), directly subtracting the item of $\ln(\mathrm{RR}_{\mathrm{mean}})$ is arbitrary. Last, we use a weight to optimize the effect of mean RR interval on the final entropy output of EntropyAF as follows:

$$\mathrm{Entropy}_{\mathrm{AF}} = -\ln\left(\frac{\mathrm{AX}^{m+1}(n,r)}{\mathrm{BX}^m(n,r)}\right) + \ln(2r) - w \times \ln(\mathrm{RR}_{\mathrm{mean}}) \tag{13}$$

where w is a weight for optimization.

References

1. Young, M.: Atrial fibrillation. Crit. Care Nurs. Clin. North Am. **31**(1), 77–90 (2019). https://doi.org/10.1016/j.cnc.2018.11.005

2. Chen, Y., Wang, X., Jung, Y., Abedi, V., et al.: Classification of short single lead electrocardiograms (ECGs) for atrial fibrillation detection using piecewise linear spline and XGBoost. Physiol. Meas. **39**(10), 104006 (2018). https://doi.org/10.1088/1361-6579/aadf0f
3. Lloyd-Jones, D.M., Wang, T.J., Leip, E.P., Larson, M.G., Levy, D., et al.: Lifetime risk for development of atrial fibrillation: the Framingham Heart Study. Circulation. **110**(9), 1042–1046 (2004). https://doi.org/10.1161/01.CIR.0000140263.20897.42
4. Wolf, P.A., Abbott, R.D., Kannel, W.B.: Atrial fibrillation as an independent risk factor for stroke: the Framingham Study. Stroke. **22**(8), 983–988 (1991). https://doi.org/10.1161/01.STR.22.8.983
5. Naccarelli, G.V., Varker, H., Lin, J., Schulman, K.L.: Increasing prevalence of atrial fibrillation and flutter in the United States. Am. J. Cardiol. **104**(11), 1534–1539 (2009). https://doi.org/10.1016/j.amjcard.2009.07.022
6. Fotiadis, D., Likas, A., Michalis, L., Papaloukas, C.: Electrocardiogram (ECG): Automated Diagnosis. Wiley Encyclopedia of Biomedical Engineering, p. 2006. John Wiley & Sons, Hoboken, NJ (2006). https://doi.org/10.1002/9780471740360.ebs0382
7. Cheng, S., Tamil, L.S., Levine, B.: A mobile health system to identify the onset of paroxysmal atrial fibrillation. In: 2015 International Conference on Healthcare Informatics, Dallas, TX, USA. October 21–23, pp. 189–192 (2015)
8. Rangel, M.O., O'Neal, W.T., Soliman, E.Z.: Usefulness of the electrocardiographic P-wave axis as a predictor of atrial fibrillation. Am. J. Cardiol. **117**(1), 100–104 (2016). https://doi.org/10.1016/j.amjcard.2015.10.013
9. Maji, U., Mitra, M., Pal, S.: Differentiating normal sinus rhythm and atrial fibrillation in ECG signal: a phase rectified signal averaging based approach. In: Control, Instrumentation, Energy and Communication (CIEC), 2014 International Conference on, Calcutta, India, January 31–Feb 02, pp. 176–180 (2014)
10. Andrikopoulos, G.K., Dilaveris, P.E., Richter, D.J., Gialafos, E.J., Gialafos, J.E.: Increased variance of P wave duration on the electrocardiogram distinguishes patients with idiopathic paroxysmal atrial fibrillation. Pacing Clin. Electrophysiol. **23**(7), 1127–1132 (2000). https://doi.org/10.1111/j.1540-8159.2000.tb00913.x
11. Pürerfellner, H., Pokushalov, E., Sarkar, S., et al.: P-wave evidence as a method for improving algorithm to detect atrial fibrillation in insertable cardiac monitors. Heart Rhythm. **11**(9), 1575–1583 (2014). https://doi.org/10.1016/j.hrthm.2014.06.006
12. Alcaraz, R., Vayá, C., Cervigón, R., Sánchez, C., Rieta, J.J.: Wavelet sample entropy: a new approach to predict termination of atrial fibrillation. Comput. Cardiol. **33**, 597–600 (2006)
13. Alcaraz, R., Rieta, J.J.: Application of Wavelet Entropy to predict atrial fibrillation progression from the surface ECG. Comput. Math. Methods Med. **2012**, 245213 (2012). https://doi.org/10.1155/2012/245213
14. García, M., Ródenas, J., Alcaraz, R., Rieta, J.J.: Application of the relative wavelet energy to heart rate independent detection of atrial fibrillation. Comput. Methods Prog. Biomed. **131**, 157–168 (2016). https://doi.org/10.1016/j.cmpb.2016.04.009
15. Paoletti, M., Marchesi, C.: Discovering dangerous patterns in long-term ambulatory ECG recordings using a fast QRS detection algorithm and explorative data analysis. Comput. Methods Prog. Biomed. **82**(1), 20–30 (2006). https://doi.org/10.1016/j.cmpb.2006.01.005
16. Couceiro, R., Carvalho, P., Henriques, J., Antunes, M., Harris, M., Habetha, J.: Detection of atrial fibrillation using model-based ECG analysis. In: 2008 19th International Conference on Pattern Recognition, Tampa, FL, USA, December 8–11, pp. 1–5 (2008). https://doi.org/10.1109/ICPR.2008.4761755
17. Babaeizadeh, S., Gregg, R.E., Helfenbein, E.D., Lindauer, J.M., Zhou, S.H.: Improvements in atrial fibrillation detection for real-time monitoring. J. Electrocardiol. **42**(6), 522–526 (2009). https://doi.org/10.1016/j.jelectrocard.2009.06.006
18. Clifford, G.D., Liu, C., Moody, B., et al.: AF classification from a short single lead ECG recording: the PhysioNet/computing in cardiology challenge 2017. Comput. Cardiol. **44**, 1–4 (2017). https://doi.org/10.22489/CinC.2017.065-469

19. Datta, S., Mukherjee, A., Banerjee, R., et al.: Identifying normal, AF and other abnormal ECG rhythms using a cascaded binary classifier. Comput. Cardiol. **44**, 1–4 (2017). https://doi.org/10.22489/CinC.2017.173-154
20. Sarkar, S., Ritscher, D., Mehra, R.: A detector for a chronic implantable atrial tachyarrhythmia monitor. IEEE Trans. Biomed. Eng. **55**(3), 1219–1224 (2008). https://doi.org/10.1109/TBME.2007.903707
21. Park, J., Lee, S., Jeon, M.: Atrial fibrillation detection by heart rate variability in Poincare plot. Biomed. Eng. Online. **8**(1), 38 (2009). https://doi.org/10.1186/1475-925X-8-38
22. Logan, B., Healey, J.: Robust detection of atrial fibrillation for a long term telemonitoring system. Comput. Cardiol. **32**, 619–622 (2005)
23. Linker, D.T.: Long-term monitoring for detection of atrial fibrillation, US Patent 7630756 B2, 2009
24. Larburu, N., Lopetegi, T., Romero, I.: Comparative study of algorithms for atrial fibrillation detection. Comput. Cardiol. **38**, 265 (2011). https://doi.org/10.1016/j.proenv.2011.12.238
25. Huang, C., Ye, S., Chen, H., Li, D., Tu, Y.: A novel method for detection of the transition between atrial fibrillation and sinus rhythm. IEEE Trans. Biomed. Eng. **58**(4), 1113–1119 (2011). https://doi.org/10.1109/TBME.2010.2096506
26. Lake, D.E., Moorman, J.R.: Accurate estimation of entropy in very short physiological time series: the problem of atrial fibrillation detection in implanted ventricular devices. Am. J. Physiol. Heart Circ. Physiol. **300**(1), H319–H325 (2011). https://doi.org/10.1152/ajpheart.00561.2010
27. Richman, J.S., Moorman, J.R.: Physiological time-series analysis using approximate entropy and sample entropy. Am. J. Phys. Heart Circ. Phys. **278**(6), H2039–H2049 (2000). https://doi.org/10.1152/ajpheart.2000.278.6.H2039
28. Liu, C., Zhao, L.: Using Fuzzy Measure Entropy to improve the stability of traditional entropy measures. Comput. Cardiol. **38**, 681 (2011)
29. Liu, C., Oster, J., Reinertsen, E., Li, Q., Zhao, L., et al.: A comparison of entropy approaches for AF discrimination. Physiol. Meas. **39**(7), 074002 (2018). https://doi.org/10.1088/1361-6579/aacc48
30. Zhao, L., Liu, C., Wei, S., Shen, Q., Zhou, F., Li, J.: A new Entropy-based atrial fibrillation detection method for scanning wearable ECG recordings. Entropy. **20**(12), 904 (2018). https://doi.org/10.3390/e20120904
31. Liu, C., et al.: Signal quality assessment and lightweight QRS detection for wearable ECG smartvest system. IEEE Internet Things J. **6**(2), 1363–1374 (2019). https://doi.org/10.1109/JIOT.2018.2844090
32. Liu, F., Liu, C., Jiang, X., Zhao, L., Wei, S.: A decision-making fusion method for accurately locating QRS complexes from the multiple QRS detectors. World Cong. Med. Phys. Biomed. Eng. **68**(Part 2), 351–355 (2019). https://doi.org/10.1007/978-981-10-9038-7_66
33. Hsu, C., Chang, C., Lin, C.: A Practical Guide to Support Vector Classification. National Taiwan University, Taipei (2008). https://doi.org/10.1111/j.1365-3016.1995.tb00168.x

Applications of Heart Rate Variability in Sleep Apnea

Xiaotong Dong, Shoushui Wei, Hongru Jiang, and Chengyu Liu

Abstract Sleep apnea detection based on Electrocardiograph (ECG) is of great significance for the simplification of the detection process and clinical application. Heart rate variability (HRV) is a marker of autonomic nervous system (ANS) activity that can reflect arrhythmias, which often occur during sleep apnea. In this paper, we take sleep apnea-hypopnea syndrome (SAHS) as an example to introduce the time-domain, frequency-domain and nonlinear analysis methods of HRV in detail, as well as each step of ECG signal processing, and study the performance and differences of HRV between normal sleep signals and sleep apnea signals. Through the Mann–Whitney *U* test, we found that HRV has significant differences between normal sleep signals and sleep apnea signals, indicating that it is suitable as a preliminary screening tool for detecting sleep apnea.

Keywords Sleep apnea · Heart rate variability · Electrocardiograph

1 Introduction

1.1 Introduction to Sleep Apnea-Hypopnea Syndrome

SAHS refers to recurrent apnea and/or hypoventilation, hypercapnia and sleep interruption caused by various reasons during sleep, resulting in a series of pathophysiological changes in the human body. According to different causes, it is often divided into obstructive sleep apnea syndrome (OSAS), central sleep apnea syndrome (CSAS), and mixed sleep apnea syndrome (MSAS), of which OSAS is the

X. Dong · S. Wei (✉) · H. Jiang
School of Control Science and Engineering, Shandong University, Jinan, China
e-mail: sswei@sdu.edu.cn

C. Liu (✉)
The State Key Laboratory of Bioelectronics, School of Instrument Science and Engineering, Southeast University, Nanjing, China
e-mail: chengyu@seu.edu.cn

C. Liu, J. Li (eds.), *Feature Engineering and Computational Intelligence in ECG Monitoring*, https://doi.org/10.1007/978-981-15-3824-7_11

most common type [1]. Sleep apnea can cause daytime sleepiness, irritability, mental and intellectual decline, seriously affecting the quality of life, and more serious it may increase the risk of heart disease, high blood pressure and even death [2]. In recent years, studies have shown that sleep apnea has a certain correlation with the occurrence of cancer [3]. According to the World Health Organization epidemiological survey, about one-third of the world's population suffers from sleep disorders, and it has been increasing in recent years.

The apnea hypopnea index (AHI), which is the sum of the average number of apnea and hypopnea per hour, is the main indicator of the severity of sleep apnea syndrome. The specific definition of sleep apnea is as follows: during sleep, the oronasal respiratory airflow disappears or significantly weakens (down 90% from the baseline) for at least 10 s. If the oronasal airflow during sleep is 30% lower than the baseline level for at least 10 s, accompanied by a decrease in blood oxygen saturation (SaO2) of 4%, or if the oronasal airflow decreased by 50% for 10 s with SaO2 decreased by 3%, it is considered as hypoventilation. During the 7 h of sleep, sleep apnea or hypoventilation occurs repeatedly more than 30 times or AHI $\geq$ 5 is considered as SAHS. Among them, 5 $\leq$ AHI $<$ 15, mild, 15 $\leq$ AHI $<$ 30, moderate, and AHI $\geq$ 30, severe [4].

The detection of SAHS is mainly based on polysomnography (PSG). Multiple physiological signals such as electroencephalogram (EEG), electrooculogram (EOG), electromyogram (EMG), ECG, SaO2, respiration and so on are recorded by PSG, and diagnosis is made based on a comprehensive analysis of all of these signals throughout the night. This is the gold standard for SAHS diagnosis [5]. However, due to the high price and troublesome operation of the PSG, many of patients with sleep apnea are not diagnosed in time, and the test process will seriously affect the patient's sleep quality, leading to incorrect test results [6]. With the development of wearable devices, SAHS detection based on single-channel signal is very necessary.

1.2 Introduction to Heart Rate Variability (HRV)

HRV is an analysis of changes in the heartbeat cycle [7]. The change in heart rate is the result of the action of ANS. HRV contains information on the regulation of the cardiovascular system by neurohumoral factors, so that it is a valuable indicator for predicting sudden cardiac death and arrhythmic events [8]. HRV analysis includes time domain analysis, frequency domain analysis and nonlinear analysis and has a good performance in predicting hypertension, myocardial infarction, and arrhythmia.

HRV has important applications and significance in the diagnosis of many diseases. Currently, HRV is used clinically as a predictor of risk after acute myocardial infarction [9] and as an early warning marker of diabetic neuropathy [10]. Yan et al. [11] compared the differences in HRV between patients with congestive heart failure and normal subjects and quantified the distinguishing ability

of HRV features. They considered that HRV features can distinguish heart failure. Bodin et al. [12] studied the impact of smoking on ANS and found that smokers had lower HF than nonsmokers. Pal et al. [13] found that spectral analysis of HRV may predict the future development of essential hypertension. Blom et al. [14] found that adolescent female psychiatric patients with anxiety disorders and/or major depressive disorder show reduced HRV compared with healthy controls.

HRV also has many applications in the field of sleep, such as the judgment of the severity of sleep apnea [15], the classification of sleep stages [16, 17], and the detection of sleep apnea events [18]. Aytemir et al. [19] studied 80 patients with OSAS and 55 healthy people of age-matched age. They found that in OSAS, the sympathetic nervous system and parasympathetic nervous system were affected, but the changes in nocturnal autonomic balance are more beneficial to the sympathetic nervous system. Roche et al. [20] studied 91 subjects, of which 39 were patients. The study believes that time-domain HRV analysis can be used as an accurate and inexpensive screening tool for clinically suspected OSAS patients.

1.3 Research Steps in This Paper

In this paper, we comprehensively analyze the performance and differences of HRV between normal sleep signals and sleep apnea signals. This study includes ten patients with SAHS, which are taken from the PSG monitor of the Provincial Hospital of Shandong Province. ECGI signal with sampling frequency of 200 Hz is used and the signal duration is 1 h. The signals are marked per minute to indicate whether sleep apnea exists during this time. If so, it is marked as 1; if not, it is marked as 0. The details of the data are shown in Table 1. The research process of this paper is as follows.

Table 1 The detailed information of the data

No.	Gender	Age	BMI	Sleep efficiency (%)	Maximum apnea time (s)	Longest hypoventilation time (s)	Minimum SaO2	AHI
120	Male	50	23.39	84.2	60.5	57.5	74	62.6
123	Male	40	21.88	73.6	60.5	51.5	91	14.3
131	Male	43	26.29	78.6	13	47	85	7.7
134	Male	50	18.69	59.4	68.5	58	86	27.8
267	Male	51	30.42	92.8	54	91	82	10.9
271	Male	51	25.83	80.2	104	116.5	47	54.2
285	Male	39	27.89	86.8	30	23	80	31.5
291	Male	49	24.91	66	45.5	52.5	84	20.7
332	Male	40	25.25	89.3	34.5	35.5	92	7.9
355	Male	58	28.37	94.9	50	44.5	71	52.9

1. Signal preprocessing: ECG signals contain a lot of noise, so the signals are first processed for noise reduction. Then, according to the literature [21], sleep apnea is detected best when the signal length is 1 min, so the signals are segmented into 1-min segments, a total of 600 segments.
2. Feature extraction: Based on ECG signals, five time-domain features, four frequency-domain features, six nonlinear features are extracted.
3. Feature analysis: First, significance test is used to analyze which features are significantly different between the sleep apnea signals and the normal sleep signals, then the features with significant differences are further analyzed.

2 Signal Preprocessing

ECG signal is a very weak physiological low-frequency electrical signal. The amplitude is generally only 0.05–4 mV, and the frequency range is generally 0.1–35 Hz, mainly concentrated at 5–20 Hz. ECG signals are recorded by electrodes mounted on the skin surface of the human body and are extremely susceptible to internal and external noise, which affects the shape and recognition of the waveform, as shown in Fig. 1a. Next, the main noises of the ECG signal, including baseline drift, power frequency interference, and myoelectric noise will be filtered.

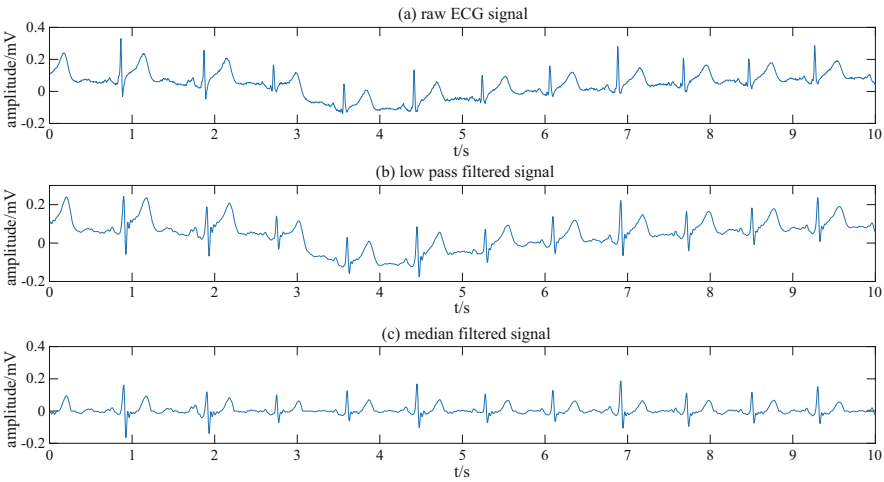

Fig. 1 (**a–c**) Signal preprocessing

2.1 Low-Pass Filtering

The myoelectric noise generally comes from the contraction and tremor of muscles, and its frequency is generally greater than 30 Hz, which appears as an irregular and rapidly changing waveform. Power frequency interference is mainly generated on the power supply equipment and is an inevitable interference of the ECG signal. It shows a "burr" on the ECG signal, which is often filtered by a 50 Hz notch [22]. In order to simplify the processing, we used a 30 Hz Butterworth low-pass filter to filter out both myoelectric and power frequency interference, because this study only needs to identify the position of the QRS wave and the frequency of the QRS wave is about 15 Hz. Figure 1b shows the signal after low-pass filtering. The signal becomes smoother than before.

2.2 Median Filtering

The baseline drift is generally caused by human respiration and electrode movement, and its frequency is generally less than 1 Hz. There are many methods to eliminate it, such as median filtering, high-pass filtering, wavelet transform, and morphological filtering, etc. Many studies have shown that compared with other methods, the wavelet transform method has the best filtering effect, but it has a large amount of calculation and is not suitable for real-time processing [23, 24].

Median filtering is a nonlinear smoothing technique with a small amount of calculation. It has a good suppression effect on both severe and weak baseline drift, and it also has a good protection effect on the ST segment. The main principle is to first extract the baseline using median filtering, and then subtract the baseline from the original signal. The size of the window of the median filter is the main reason that affects the filtering effect. According to previous research experience, the filtering effect is best when the size of the window is about 30% of the sampling frequency. Figure 1c shows the signal after median filtering. The baseline drift is effectively filtered and the signal waveform is well protected.

2.3 Signal Quality Assessment

Although signal preprocessing can improve the signal quality to some extent, some signals are still unusable because of poor quality, as shown in Fig. 2a, b. Therefore, it is necessary to evaluate the quality of the signal and remove signals of poor quality. In 2011, PhysioNet/Computing in Cardiology launched a competition on the evaluation of ECG signal quality and many useful methods were proposed in the competition, such as methods based on signal waveform characteristics [25], methods based on noise type [26], methods based on machine learning [27], etc.

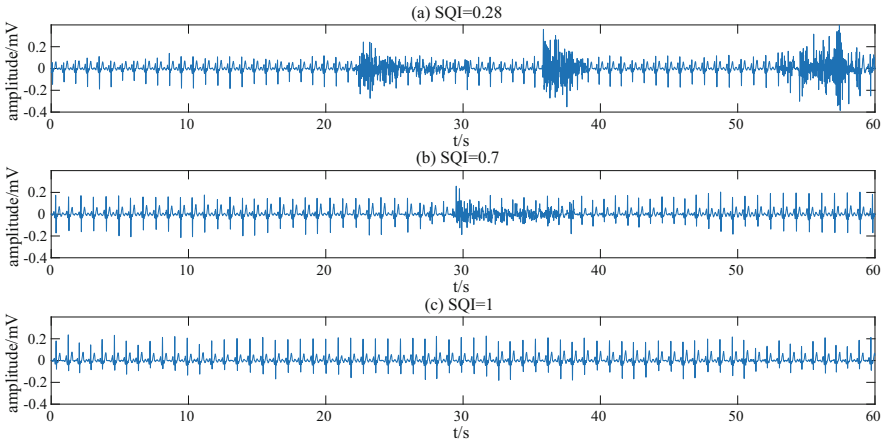

Fig. 2 (**a**–**c**) ECG signals with different SQI

Because the purpose of this paper is to identify the position of the QRS wave, the quality assessment method used in this paper is as follows.

Two QRS wave detection algorithms with different sensitivity to noise are used to detect the location of QRS, and the F1 measure is used as an index to measure the consistency of the two methods. The F1 measure is calculated in the form of a sliding window, where the length of the window is 10 s and the step size is 1 s [28]. If $F1 < 0.9$, the signal quality in this window is poor. Signal quality index (SQI) = number of windows with good signal quality/total number of windows. If $SQI \leq 0.8$, the signal will be discarded. Figure 2 shows signals of different qualities, and this method is very effective. After signal quality assessment, there are 452 1-min signals in total, of which 190 are normal sleep signals and 262 are sleep apnea signals.

3 Feature Description of HRV

In this section, we will explain the computational and physiological significance of each feature.

3.1 Time-Domain Features

The following sections introduce the calculation methods of common time-domain features and their physiological and statistical significance. In the following formula, N represents the length of the RR interval sequence, and μ represents the mean of the RR interval sequence.

1. Standard deviation (SDNN): SDNN reflects all the cyclic components that cause changes in the RR interval sequence.

$$\mathrm{SDNN} = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(\mathrm{RR}_i - \mu)^2} \tag{1}$$

2. Root-mean-squared value (RMSSD): RMSSD is the root mean square of the difference between adjacent RR intervals, and assesses high-frequency changes in heart rate.

$$\mathrm{RMSSD} = \sqrt{\frac{1}{N-1}E_{i=2}^{N}(\mathrm{RR}_i - \mathrm{RR}_{i-1})^2} \tag{2}$$

3. Skewness (Skew): A measure of the asymmetry of the probability distribution of a real-valued random variable. Skew = 0, normal distribution; Skew < 0, left skewness, where there is less data to the left of the mean of RR interval sequence than to the right; Skew > 0, Right skew, opposite to left skew.

$$\mathrm{Skew}(\mathrm{RR}) = E\left[\left(\frac{\mathrm{RR}-\mu}{\sigma}\right)^3\right], \tag{3}$$

where σ represents standard deviation of RR interval sequence.

4. Kurtosis (Kurt): A measure of "heaviness of the tails" of a probability distribution, defined as the fourth cumulant divided by the square of the variance of the probability distribution. Kurt = 3, normal distribution; Kurt < 3, insufficient kurtosis; Kurt > 3, excessive kurtosis.

$$\mathrm{Kurt}(\mathrm{RR}) = \frac{1}{N \times D^2}\sum_{i=1}^{N}(\mathrm{RR}_i - \mu)^4 - 3, \tag{4}$$

where D is the variance of the RR interval sequence.

5. Interquartile range (Iqr): Iqr is the difference between the first and third quartiles, reflecting the degree of dispersion of the middle (50%) of the data. The smaller the value is, the more concentrated the data in the middle is. Especially, Interquartile range is not affected by the extreme value.

3.2 *Frequency-Domain Features*

Compared with the time-domain analysis method, the frequency-domain analysis method can better distinguish the effect of sympathetic and parasympathetic nerves on HRV. Moreover, it is more suitable for short-term HRV analysis [29].

1. Very low frequency power (VLF): power in the range of 0–0.04 Hz.
2. Low frequency power (LF): power in the range of 0.04–0.15 Hz, mainly reflecting the activity of sympathetic nerves.
3. High frequency power (HF): power in the range of 0.15–0.4 Hz, mainly reflecting the activity of the vagus nerves.
4. LF/HF: high and low frequency power ratio, reflecting the balance between sympathetic and vagal regulation.

3.3 Nonlinear Features

1. Poincaré plot is a scatter plot of the current RR interval plotted against the preceding RR interval. It takes R_i as x-axis and R_{i+1} as y-axis, where i is the index of the RR interval sequence, as shown in Fig. 3. It can be found that the distribution of these points is approximately elliptical and the center of the ellipse is located at the coordinate point determined by (μ, μ), μ represents the mean of the RR interval sequence. SD1 is the variance of the RR interval sequence in the $y = -x + 2 \times \mu$ direction, and has been correlated with high frequency power. SD2 is the variance of the RR interval sequence in the $y = x$ direction, which has

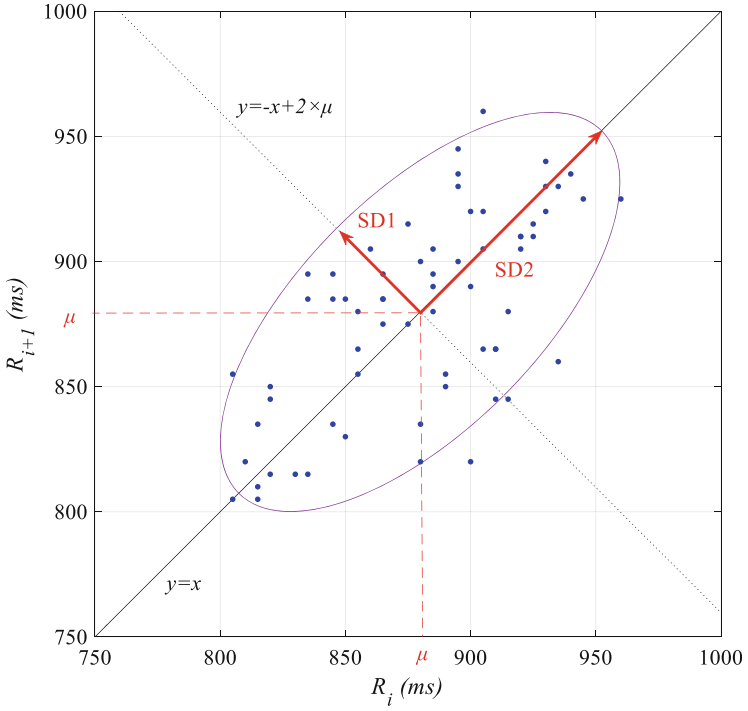

Fig. 3 Poincare diagram of the RR interval

been correlated with both low and high frequency power [7]. SD1 depicts short-term changes, which are mainly caused by respiratory sinus arrhythmia (RSA), while SD2 depicts long-term changes. SD1/SD2 reflects the balance between long-term and short-term heart rate variability [30].

2. Fuzzy measure entropy (FMEn) and approximate entropy (ApEn)
ApEn [31] is a measure of the complexity of an unstable time series. The idea is to detect the probability of a new subsequence in a time series. It is suitable for statistical analysis of relatively short and noisy time series. These statistical analyses are consistent with the general clinical need to distinguish between healthy and abnormal subjects. ApEn is calculated as follows: The time series of N data points are divided into m subsegments according to the order of the data points, and a total of $(N - m + 1)$ subsequence fragments can be obtained. Marking the subsequence fragments with $X(i)$, where $1 < i < N - M + 1$. Then calculate the distance $dX_m(i,j)$ between the current ith subsegment sequence and other subsegments $X(j)$, where $1 < j < N - M + 1$ and $j \neq i$. When $dX_m(i,j) < r$ (r represents the threshold, 0.2 times of the standard deviation of the sequence was used in here), it is considered that $X(i)$ and $X(j)$ are similar. Calculate the proportion of other sequences similar to the current ith subsegment sequence:

$$C_i^m(r) = \frac{\text{num}(dX_m(i,j) < r)}{N - m + 1} \tag{5}$$

The above analysis is performed on all subsegments to obtain the average similarity rate of the subsequence sequences at the m-data point scale:

$$\Phi^m(r) = \frac{\sum_{i=1}^{N-m+1} \log\left(C_i^m(r)\right)}{N - m + 1} \tag{6}$$

Similarly, constructing the $m + 1$ sequence, repeat the above steps, calculate $\Phi^{m+1}(r)$, and ApEn is as follows:

$$\text{ApEn} = \Phi^m(r) - \Phi^{m+1}(r) \tag{7}$$

FMEn [32] uses the membership of the fuzzy function instead of the Heaviside function used in ApEn and Sample entropy (SampEn) as the vector similarity criterion. At the same time, FMEn uses fuzzy local measure entropy and fuzzy global measure entropy to reflect the implicit overall complexity in physiological signals, making up for the limitation that fuzzy entropy only focuses on local complexity.

3. Detrended Fluctuation Analysis (DFA)
DFA [33] can eliminate the influence of trend and analyze the long-range correlation of time series. Firstly, integrate the time series x (length N) to obtain $y(k)$, and divide $y(k)$ into length N. For the isometric interval, the first-order linear fit is performed on the data of each interval by the least square's method, and the

y-coordinate of the straight-line segment is represented by $y_n(k)$. Next, the trend of the integration time series $y(k)$ is eliminated by subtracting the local trend $y_n(k)$ in each interval. The root mean square fluctuation of $y(k)$ is calculated as follows:

$$F(n) = \sqrt{\frac{1}{N}\sum_{k=1}^{N}[y(k) - y_n(k)]^2} \tag{8}$$

The above calculations are repeated on all time scales to describe the relationship between $F(n)$, average fluctuations and interval length n. The linear relationship on the *log–log* plot indicates the existence of a power-law scale. In this case, the fluctuation can be expressed by the slope of the line between log $F(n)$ and log n, expressed by α. When $0.5 < \alpha < 1$, it indicates that the time series has a long-range correlation, showing a trend of increasing trend, that is, an increasing (decreasing) trend in a certain period, and an increasing (decreasing) trend in the next time period. The closer the α is to 1, the stronger the correlation is.

4 Significance Test

In statistics, the significance test is a kind of statistical hypothesis test, which is used to detect whether there is a difference between the experimental group and the control group and whether the difference is significant in a scientific experiment. Significance tests include parametric and nonparametric tests. The commonly used methods of parametric tests include T test, analysis of variance (ANOVA), etc., which require data to follow a normal distribution and homogeneity of variance. When the data does not meet these conditions, the parametric test may give the wrong answer. At this time, a rank-based nonparametric test should be used. Common methods for nonparametric tests include Mann–Whitney U test, Kruskal–Wallis test, etc. [34]. Because some features do not meet the requirements of homogeneity of variance, the method used in this study is the Mann–Whitney U test.

The Mann–Whitney U test [35] is a nonparametric test method that uses sample ranks instead of sample values to calculate whether the difference between two groups is significant. The main idea is to first combine the two sets of data and arrange them in ascending order to obtain the rank of each sample. If the sample values are equal, the rank is defined as the average of the sum of the ranks. Let n_1 and n_2 be the volume of the two groups of samples, respectively, and T_1 and T_2 be the sum of rank of the two groups of samples, respectively. The Mann–Whitney U test statistic was calculated according to the following formula:

$$U_1 = n_1 n_2 + n_1(n_1 - 1)/2 - T_1, \quad U_2 = n_1 n_2 + n_2(n_2 - 1)/2 - T_2$$

The smaller of U_1 and U_2 is selected as the final statistic U. When both sample sizes are small, U value can be directly compared with the critical value table to determine whether there is a significant difference. When two sample sizes are large, their sampling distribution is close to normal distribution, and Z test can be used to detect them.

$$Z = \frac{U - \frac{(n_1 \times n_2)}{2}}{\sqrt{\frac{n_1 \times n_2 (n_1 + n_2 + 1)}{12}}} \tag{9}$$

When $|Z| > 1.96$, there are significant differences between the two groups ($p < 0.5$).

5 Results and Discussion

Table 2 shows the results of the Mann–Whitney U test. "$*$" indicates significant differences between the two groups. Among them, the time-domain features: Skew, Iqr, SDNN, the frequency-domain features: VLF, LF, LF/HF, and nonlinear features: ApEn, FMEn, SD2, SD1/SD2 show significant differences between normal sleep signals and sleep apnea signals. Then, features with significant differences between the two groups will be analyzed in detail.

Table 2 The results of the Mann–Whitney U test

Feature	Average rank of normal sleep group (190)	Average rank of sleep apnea group (262)	Z value	p value
SDNN	186.65	255.4	5.523	<0.001$*$
RMSSD	213.34	236.04	1.824	0.068
Skew	249.42	209.88	−3.176	0.001$*$
Kurt	230.24	223.79	−0.519	0.604
Iqr	182.59	258.34	6.088	<0.001$*$
VLF	180.68	259.73	6.35	<0.001$*$
LF	198.64	246.71	3.862	<0.001$*$
HF	218.28	232.46	1.139	0.255
LF/HF	209.76	238.64	2.32	0.02$*$
ApEn	240.82	216.12	−1.984	0.047$*$
FMEn	254.27	206.36	−3.849	<0.001$*$
SD1	213.41	236	1.815	0.07
SD2	183.99	257.32	5.891	<0.001$*$
SD1/SD2	250.93	208.78	−3.386	0.001$*$
LDA	214.05	235.53	1.726	0.084

5.1 Time-Domain Analysis of HRV

Figure 4 shows that the mean difference between the normal sleep and sleep apnea signals for each feature. The mean value of SDNN of the sleep apnea signals is larger than that of the normal sleep signals. Taking a patient's continuous sleep as an example, it can be seen from Fig. 5 that the SDNN of the sleep apnea signal is mostly above the average line, especially when the sleep apnea occurs continuously. The results are consistent with the results of using the 10-min signal length [36], while the results of using the 24-h signal length [37] are contrary to the results in this paper. However, the results of the frequency-domain features obtained in these papers are the same as those obtained in this paper. And from that, we conclude that the time-domain features are more susceptible to signal length, while the frequency-domain features are more stable for different signal lengths. It can be seen from Fig. 4 that the mean of Iqr of the sleep apnea signals is much larger than that of the normal sleep signals, indicating that the RR interval distribution of the sleep apnea signals is more dispersed. The Skew of the sleep apnea signals is basically distributed in a space smaller than 0, while that of the normal sleep signals is distributed in a space larger than 0. The difference between the two groups is obvious, as shown in Fig. 6. From this, we can infer that when sleep apnea occurs, the RR intervals are significantly prolonged.

5.2 Frequency-Domain Analysis of HRV

In Fig. 4, the mean of VLF, LF, and LF/HF of the sleep apnea signals is higher than that of the normal sleep signals, and HF is not significantly different between the two groups. Studies have shown that HF is regulated by the vagus nerve, and LF is regulated by the sympathetic nerves [38]. It can be concluded that when sleep apnea occurs, the sympathetic nerves are activated, resulting in an imbalance of the sympathetic vagal regulation [39]. The difference of VLF between the two groups is obvious. VLF represents the effects of vasodilatation and contraction, renal angiotensin system, and thermal regulation (body temperature) on heart rate, suggesting that when apnea occurs, hypoxemia and hypercapnia will cause changes in autonomic nervous tone and release of neurotransmitters. And the combined effect of increased intrathoracic pressure and systemic pressure will increase left ventricular afterload, activating the sympathetic nervous system through hypoxia and awakening from sleep, causing multiple disorders of the cardiovascular system including elevated blood pressure, arrhythmias.

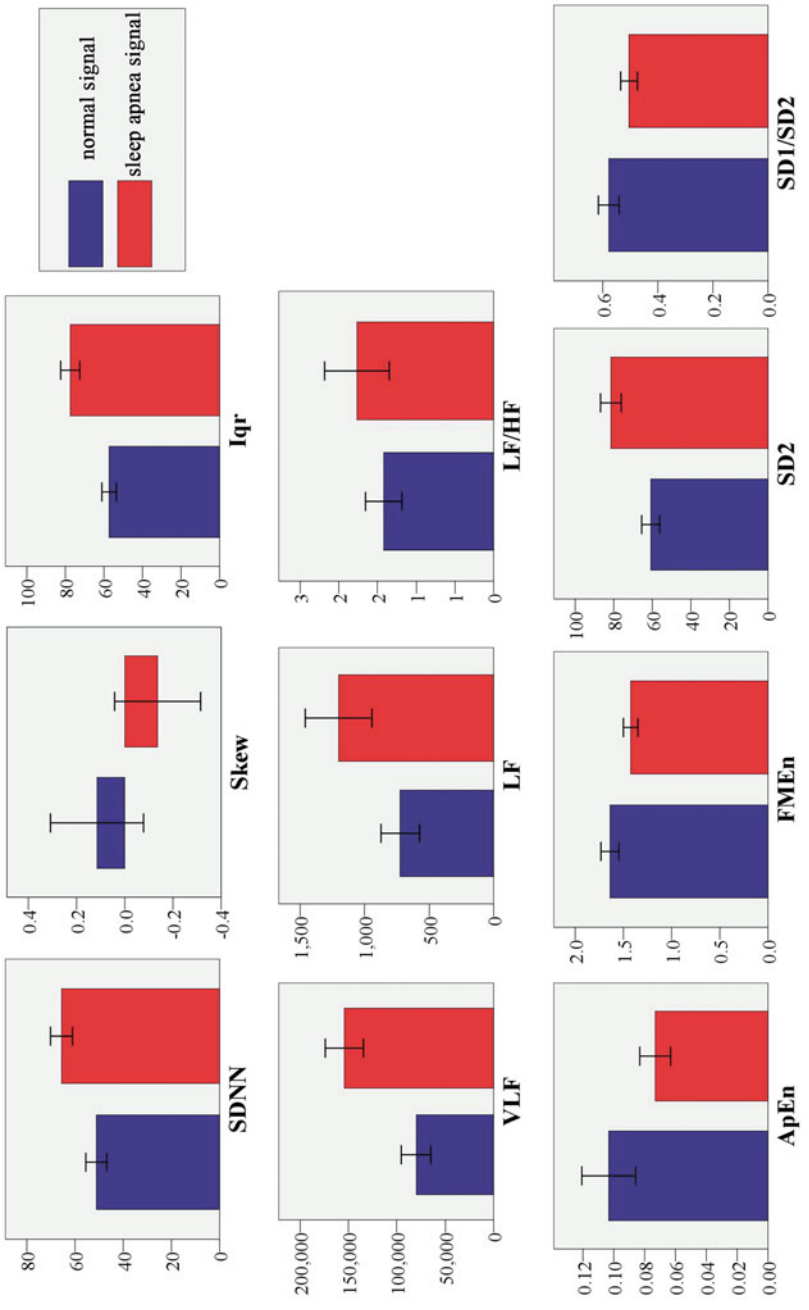

Fig. 4 Mean and 95% confidence interval of each feature

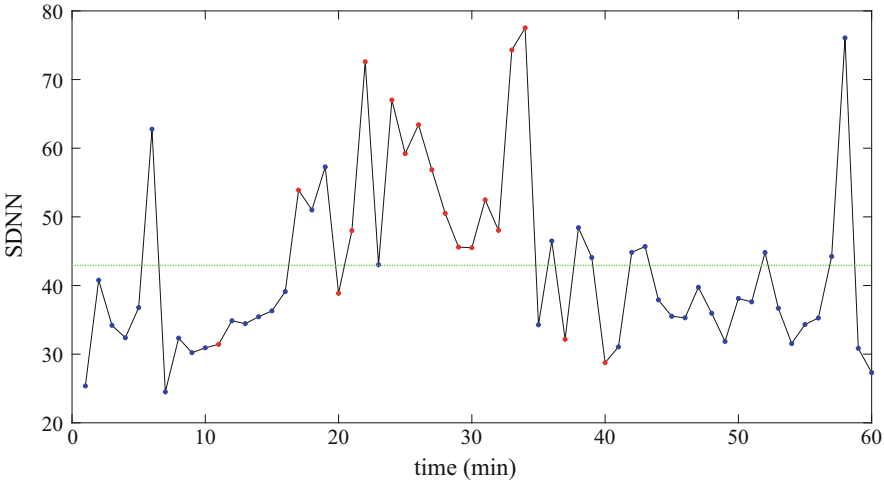

Fig. 5 Changes in SDNN during continuous sleep in patient No. 291. Among them, the blue dot represents the normal sleep, the red dot represents the sleep apnea, and the green line represents the mean value of SDNN

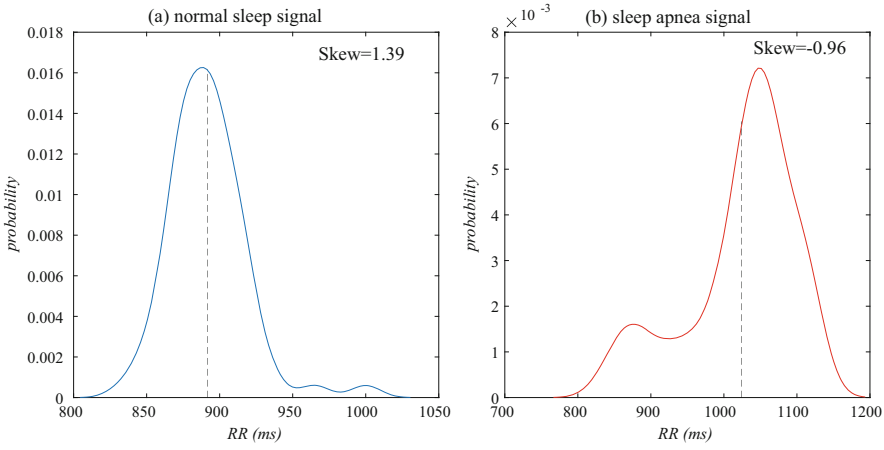

Fig. 6 (**a**, **b**) Difference in Skew between normal sleep signal and sleep apnea signal in patient No. 291

5.3 Nonlinear Analysis of HRV

Figure 4 shows that the mean of ApEn and FMEn of sleep apnea signals are smaller than that of normal sleep signals, indicating that the complexity is lower because the sympathetic excitation increases the certainty of the signal, and thus the entropy

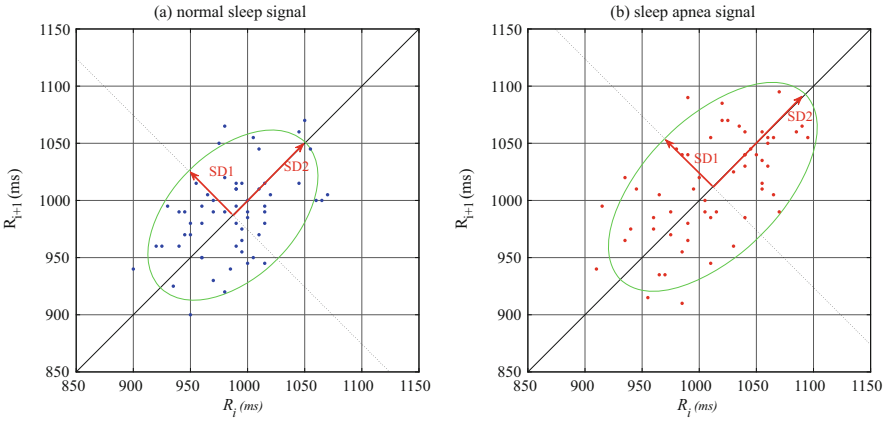

Fig. 7 (**a**, **b**) Difference in Poincare diagram between normal sleep signal and sleep apnea signal in patient No. 291

value decreases [40]. And the difference displayed by ApEn is greater than FMEn. The SD2 of the sleep apnea signals is larger, and the SD1/SD2 is smaller than the normal signals, which can be clearly seen in Fig. 7, which means that the balance of sleep apnea signals is reduced. We already know that SD1, HF, and RMSSD all indicate high-frequency changes in heart rate, and these features have no significant difference between normal sleep signals and sleep apnea signals. Therefore, we infer that the low-frequency information can better reflect the occurrence of sleep apnea.

6 Conclusions and Future Prospects

This paper describes the detailed process of analyzing sleep apnea events using HRV. The results show that HRV can, to some extent, distinguish between normal sleep signals and sleep apnea signals. This is very helpful for detecting the severity of sleep apnea. The shortcoming of this paper is that we did not consider the impact of different sleep stages on HRV, which may cause differences in results. In future experiments, we will do a more detailed study of the difference between normal sleep signals and sleep apnea signals at each sleep stage.

References

1. Javaheri, S., et al.: Sleep apnea: types, mechanisms, and clinical cardiovascular consequences. J. Am. Coll. Cardiol. **69**(7), 841–858 (2017)
2. Hernández, C., et al.: [Innovations in the epidemiology, natural history, diagnosis and treatment of sleep apnea-hypopnea syndrome]. Arch. Bronconeumol. **45**(Suppl 1), 3–10 (2009)

3. Francisco, C.R., et al.: Association between obstructive sleep apnea and cancer incidence in a large multicenter Spanish cohort. Am. J. Respir. Crit. Care Med. **187**(1), 99–105 (2013)
4. 中华医学会呼吸病学分会睡眠呼吸障碍学组: 阻塞性睡眠呼吸暂停低通气综合征诊治指南 (2011 年修订版). 中华结核和呼吸杂志. **25**(1), 162–165
5. Kapur, V.K., et al.: Clinical practice guideline for diagnostic testing for adult obstructive sleep apnea: an American academy of sleep medicine clinical practice guideline. J. Clin. Sleep Med. **13**(3), 479 (2017)
6. Min, Y., Soichiro, M., Kazuo, I.: Evaluation of type 3 portable monitoring in unattended home setting for suspected sleep apnea: factors that may affect its accuracy. Otolaryngol. Head Neck Surg. **134**(2), 204–209 (2006)
7. Malik, M.: Heart rate variability: standards of measurement, physiological interpretation, and clinical use. Ann. Noninvasive Electrocardiol. **93**(5), 1043–1065 (1996)
8. Malik, M., Camm, A.: Heart rate variability. Clin. Cardiol. **13**(8), 570–576 (1990)
9. Bigger, J.T., et al.: RR Variability in healthy, middle-aged persons compared with patients with chronic coronary heart disease or recent acute myocardial infarction. Circulation. **91**(7), 1936–1943 (1995)
10. Pagani, M., et al.: Spectral analysis of heart rate variability in the assessment of autonomic diabetic neuropathy. J. Auton. Nerv. Syst. **23**(2), 143–153 (1988)
11. Yan, W., et al.: Comparison of time-domain, frequency-domain and non-linear analysis for distinguishing congestive heart failure patients from normal sinus rhythm subjects. Biomed. Signal Process. Contr. **42**(4), 30–36 (2018)
12. Bodin, F., et al.: The association of cigarette smoking with high-frequency heart rate variability: an ecological momentary assessment study. Psychosom. Med. **79**(9), 1 (2017)
13. Pal, G.K., et al.: Spectral analysis of heart rate variability (HRV) may predict the future development of essential hypertension. Med. Hypotheses. **72**(2), 183–185 (2008)
14. Blom, E.H., et al.: Heart rate variability (HRV) in adolescent females with anxiety disorders and major depressive disorder. Acta Paediatr. **99**(4), 604–611 (2010)
15. Min, S.K., et al.: Comparison of heart rate variability (HRV) and nasal pressure in obstructive sleep apnea (OSA) patients during sleep apnea. Measurement. **45**(5), 993–1000 (2012)
16. Yilmaz, B., et al.: Sleep stage and obstructive apneaic epoch classification using single-lead ECG. Biomed. Eng. Online. **9**(1), 39 (2010)
17. Penzel, T., et al.: Comparison of detrended fluctuation analysis and spectral analysis for heart rate variability in sleep and sleep apnea. IEEE Trans. Biomed. Eng. **50**(10), 1143–1151 (2003)
18. Moody, G.B., et al.: Stimulating rapid research advances via focused competition: the Computers in Cardiology Challenge 2000. Comput. Cardiol. **27**, 207 (2000)
19. Aytemir, K., et al.: Increased myocardial vulnerability and autonomic nervous system imbalance in obstructive sleep apnea syndrome. Respir. Med. **101**(6), 1277–1282 (2007)
20. Santarelli, G.A.: Screening of obstructive sleep apnea syndrome by heart rate variability analysis: Roche F, Gaspoz JM, Court-Fortune I, et al. Circulation 100:1411, 1999. J. Oral Maxillofac. Surg. **58**(6), 696 (2000)
21. Chazal, P.D., Penzel, T., Heneghan, C.: Automated detection of obstructive sleep apnoea at different time scales using the electrocardiogram. Physiol. Meas. **25**(4), 967 (2004)
22. Hejjel, L.: Suppression of power-line interference by analog notch filtering in the ECG signal for heart rate variability analysis: to do or not to do? Med. Sci. Monit. **10**(1), MT6 (2004)
23. Afsar, F.A., Riaz, M.S., Arif, M.: A Comparison of Baseline Removal Algorithms for Electrocardiogram (ECG) Based Automated Diagnosis of Coronory Heart Disease. IEEE, Washington, DC (2009)
24. Gustavo, L., et al.: Comparison of baseline wander removal techniques considering the preservation of ST changes in the ischemic ECG: a simulation study. Comput. Math. Methods Med. **2017**, 9295029 (2017)
25. Langley, P., et al.: An algorithm for assessment of quality of ECGs acquired via mobile telephones. Comput. Cardiol. **38**, 281 (2011)

26. Liu, C., et al.: Real-time signal quality assessment for ECGs collected using mobile phones. Comput. Cardiol. **38**, 357 (2011)
27. Kalkstein, N., et al.: Using machine learning to detect problems in ECG data collection. Comput. Cardiol. **38**, 437 (2011)
28. Li, Q., Mark, R.G., Clifford, G.D.: Robust heart rate estimation from multiple asynchronous noisy sources using signal quality indices and a Kalman Filter. Physiol. Meas. **29**(1), 15–32 (2008)
29. Lombardi, F.: Clinical implications of present physiological understanding of HRV components. Card. Electrophysiol. Rev. **6**(3), 245–249 (2002)
30. Brennan, M., Palaniswami, M., Kamen, P.: Do existing measures of Poincaré plot geometry reflect nonlinear features of heart rate variability? IEEE Trans. Biomed. Eng. **48**(11), 1342–1347 (2001)
31. Pincus, S.: Approximate entropy (ApEn) as a complexity measure. Chaos. **5**(1), 110–117 (1995)
32. Liu, C., Zhao, L.: Using Fuzzy Measure Entropy to improve the stability of traditional entropy measures. Comput. Cardiol. **38**, 681 (2011)
33. Peng, C., et al.: Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series. Chaos. **5**(1), 82–87 (1995)
34. Chan, Y.: Biostatistics 102: quantitative data - parametric & non-parametric tests. Singap. Med. J. **44**(8), 391–396 (2003)
35. Bagdonavičius, V., Kruopis, J., Nikulin, M.S.: Non-parametric Tests for Complete Data. John Wiley & Sons, Inc., Hoboken, NJ (2011)
36. Szollosi, I., et al.: Sleep apnea in heart failure increases heart rate variability and sympathetic dominance. Sleep. **30**(11), 1509–1514 (2007)
37. Aydin, M., et al.: Cardiac autonomic activity in obstructive sleep apnea: time-dependent and spectral analysis of heart rate variability using 24-hour Holter electrocardiograms. Tex. Heart Inst. J. **31**(2), 132–136 (2004)
38. Montano, N., et al.: Heart rate variability explored in the frequency domain: a tool to investigate the link between heart and behavior. Neurosci. Biobehav. Rev. **33**(2), 71–80 (2009)
39. Vanninen, E., et al.: Cardiac sympathovagal balance during sleep apnea episodes. Clin. Physiol. **16**(3), 209–216 (1996)
40. Seker, R., et al.: Validity test for a set of nonlinear measures for short data length with reference to short-term heart rate variability signal. J. Syst. Integr. **10**(1), 41–53 (2000)

False Alarm Rejection for ICU ECG Monitoring

Jian Dai, Zehui Sun, and Xianliang He

Abstract Alarm fatigue, a pain point in clinic, includes ECG false alarm and ECG meaningless alarm. For one thing, ECG false alarm is mainly caused by clinical care, patients' daily activity and attachment connection failure. Two viable solutions including multi-lead and multi-parameter are provided. For another, ECG meaningless alarm contains arrhythmia class and heart rate over-limit class. Clinical alarm management system improvement and intelligent alarm are proposed in order to reduce ECG meaningless alarm. In a word, this article indicates many effective solutions to relief alarm fatigue.

Keywords Alarm fatigue · False alarm · Meaningless alarm · Intelligent alarm

1 Introduction

The monitoring equipment has been widely used in the clinical setting in order to constantly monitor the vital signs of the patient. When there is a change in the vital sign indicating potential risk in a patient's physiological condition, the medical staff will be alerted by an alarm. However, due to environmental noise, patient movement, inappropriate setting of alarm, and so on, the patient monitoring device may generate many false or unnecessary alarms. Too many false alarms will increase the medical staff cognitive load. As a result, caregivers become desensitized and may simply ignore the alarms—a phenomenon called alarm fatigue [1]. Alarm fatigue can result in impaired recognition of critical event. According to the American Emergency Medical Research Institute (ECRI) in 2013, the alarm fatigue problem ranked first in the top ten medical technology hazards [2]. Moreover, it was not the first time that alarm fatigue appeared on the list. Since 2007, ECRI has listed it in the "top ten medical hazards" for more than once.

J. Dai · Z. Sun · X. He (✉)
Shenzhen Mindray Bio-medical electronics CO., LTD., Shenzhen, China
e-mail: Hexianliang@mindray.com

C. Liu, J. Li (eds.), *Feature Engineering and Computational Intelligence in ECG Monitoring*, https://doi.org/10.1007/978-981-15-3824-7_12

Alarm fatigue is a serious threat to patient safety. In 2010, the US media reported a medical accident in which the patient did not receive timely treatment due to excessive alarms. The report raised great public concern over the problem of how the clinical alarm system may endanger patient safety [3]. According to a report released by Joint the Commission of the United States, 98 adverse cases related to medical device alarms occurred from 2009 to 2012, including 80 cases of patient death, 13 cases of permanent loss of certain physical function, and five cases of prolonged hospital stay [4]. In the NPSG (national patient safety goals) report in 2016, released also by the Joint Commission, reducing the hazards of clinical alarm fatigue is recognized as an important consensus to improve patient safety [5].

2 An Overview of the Alarms in ICU

Depending on whether the alarm reflects a patient's physiological or therapeutical change [6], the clinical alarms can be divided into correct and false alarms; the correct alarms can be further marked as clinically relevant or clinically irrelevant based on whether clinical intervention is needed. Alarm fatigue problems are mainly caused by false alarms and clinically irrelevant alarms. According to Bonafid et al., 24% of the clinical alarms were false alarms, and among the correct alarms, 87% of them were clinically irrelevant [7, 8].

ECG reflects cardiac function, and is one of the most important vital signs in clinical monitoring. Arrhythmia alarms and heart rate alarms generated during ECG monitoring are the major causes of alarm fatigue. Drew et al. showed that half of the monitor alarms over his study of 31-day period are arrhythmia alarms, among which up to 89% were annotated as false alarm [1]. Wu Jun et al. found, in their 7-day study in the intensive care unit (ICU) monitor alarms, the rate of false alarm and the rate of clinical irrelevant alarm reached 65.4% and 32.1%, respectively. In general, 99.9% of the arrhythmia alarms were false alarms [9]. Improving the accuracy of ECG monitoring is therefore the most important way to solve the problem of alarm fatigue.

3 False Alarm Reduction for ECG Monitoring

In the clinical environment, routine care, patient movement, accessory failure, electrode dropping, etc. can all introduce artifacts in ECG signal, resulting false alarms. The complexity of the clinical practice and the diversity of ECG characteristics of different patients together pose a great challenge to ECG false alarm reduction. It is a key issue faced by all patient monitor manufacturers.

3.1 Causes of ECG False Alarms

3.1.1 Interference from Clinical Care

In the ICU, to prevent pressure ulcers, nurses often help the patients change their positions regularly in the bed. The movements may compress or pull the electrode, introducing artifacts in ECG recordings. Figure 1 shows a four-lead ECG signal. The ventricular-tachycardia-like signal segments in the top three recordings are caused by back tapping and body scratching. For this situation, if only ECG II and V lead were analyzed, a false ventricular arrhythmia alarm would be generated. Referring to III lead, however, can help rule out the false alarm.

3.1.2 Interference from Patient Movement

In the subintensive care unit, patients are encouraged to perform their activities of daily living for early postoperative recovery. Those daily activities often cause false alarms in ECG monitoring. For example, patients after cardiothoracic surgery are encouraged to do early functional exercise. The friction between electrode and clothing may cause motion artifacts in ECG signal. As shown in Fig. 2, in the lead II, III, and V, motion artifacts of mid-high frequency are so severe that the QRS wave can hardly be distinguished. Meanwhile the lead I ECG signal is not interfered. The artifacts can be removed by joint analysis of all the leads.

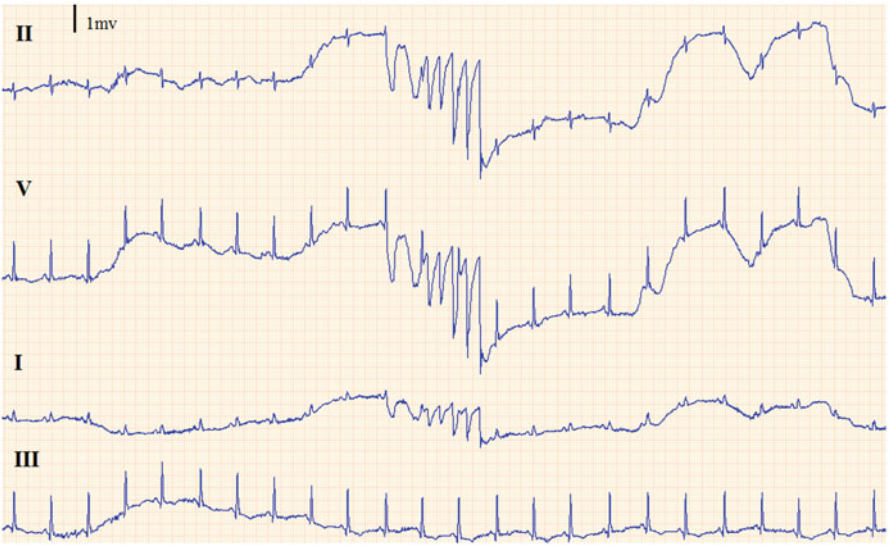

Fig. 1 Interference showing a pattern similar to ventricular tachycardia

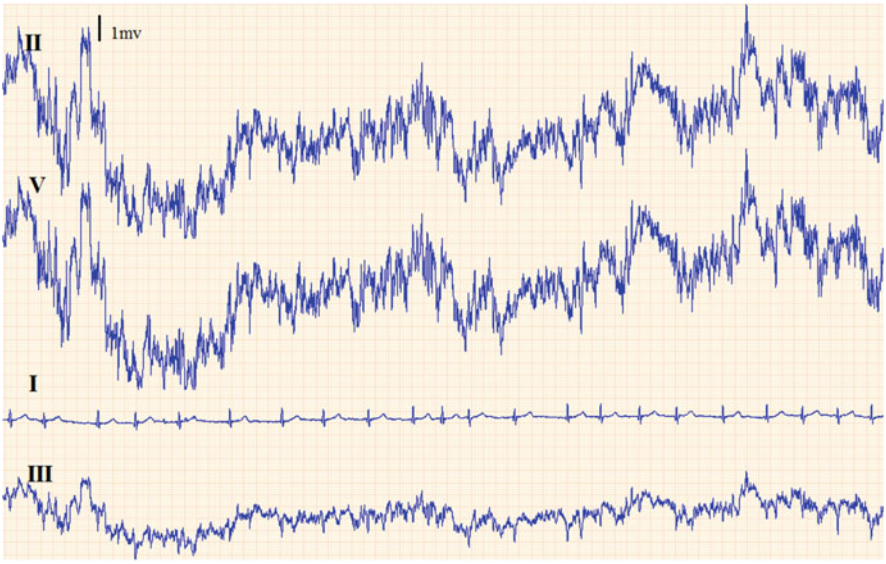

Fig. 2 Example of friction interference to electrode from patient motion

The function of multi-lead ECG joint analysis has a great advantage in this situation. Monitors without this function will generate incorrect heart rate and many arrhythmia false alarms.

At the same time, daily living activities such as brushing the teeth, lying and other activities may squeeze the ECG electrode. As illustrated in Fig. 3, the RA, V electrodes are squeezed, resulting in severe distortions in the I, II, V lead ECG waveforms, and QRS waves. At the same time, since the waveforms of these interferences have certain regularity in their pattern, and III recording is not disturbed, these artifacts can be removed successfully (Fig. 3).

During the sustained monitoring process, long-term adhesion of the electrode will cause skin irritation. If the patient scratches the surrounding area, artifacts in lead I, II, and V in Fig. 4 will show up in the ECG recordings. These artifacts have a similar shape to that of short-term ventricular tachycardia and cannot be eliminated by single lead ECG analysis. Only when the I, II, V, and III lead ECG recordings are jointly analyzed can the interference be completely removed, and the arrhythmia false alarm can be avoided.

3.1.3 Interference Caused by Electrode Dropping

During the monitoring process, if the monitoring time is too long, the electrode may get loose or even fall off. In this situation, the dropped lead will generate no signal. As shown in Fig. 5, the lead I, III, and V signals are lost, but the lead II signal is still normal and therefore can be used for arrhythmia detection. In this situation,

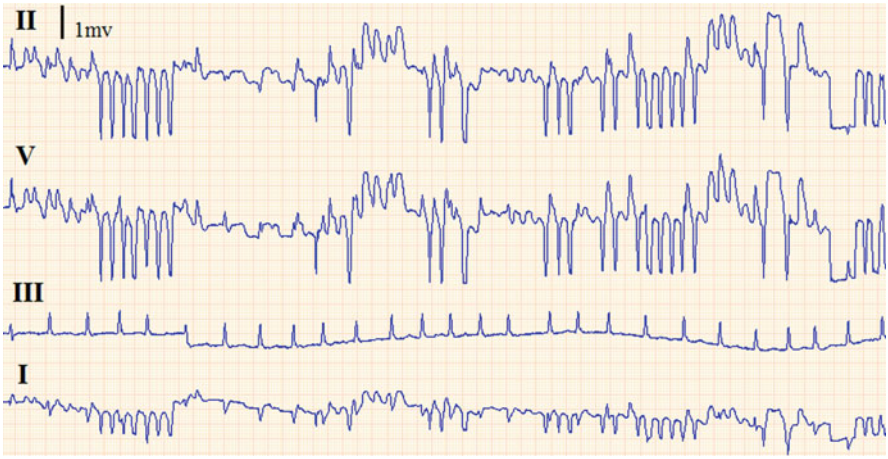

Fig. 3 Example of ECG interference from patient squeezing the electrodes

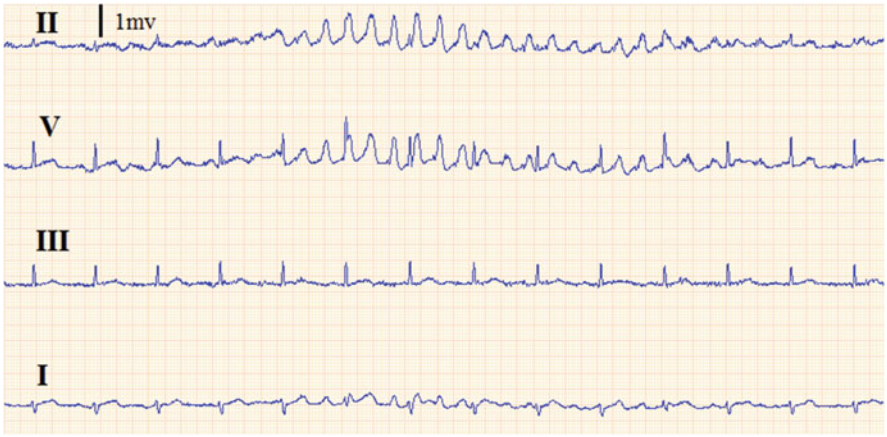

Fig. 4 Example of ECG interference from skin scratching

single-lead or two-lead ECG monitoring mode will generate a false asystole alarm. It is necessary for the nurse to manually switch to normal lead or reattach the electrode.

3.2 ECG False Alarm Solution

3.2.1 Multi-Lead Joint Analysis

One of the most effective ways to reduce ECG false alarms is multi-lead ECG analysis. As early as 1989, AHA suggested that the monitoring equipment should be

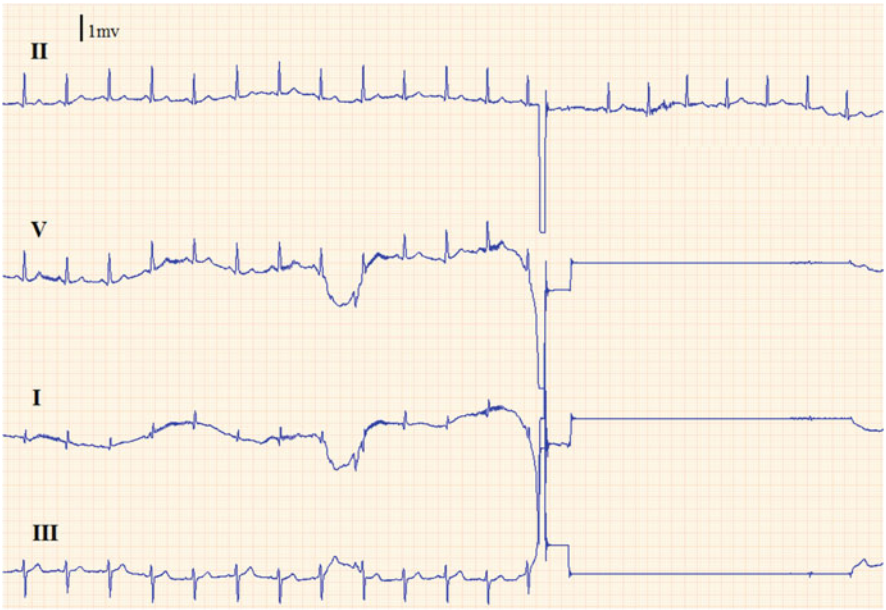

Fig. 5 Example of interference from electrode dropping

capable of jointly analyzing and simultaneously displaying two, preferably three or more leads. Multi-lead analysis is a promising way for reducing false alarms [10]. Compared with the single-lead ECG algorithm, using signals from multiple leads can reduce the impact of interference and improve the detection accuracy of the algorithm. Moreover, in the multi-lead mode, medical staff will no longer need to select the analysis lead as the algorithm will auto-select the lead. In addition to the increased ease of use, multi-lead joint ECG analysis will also reduce arrhythmia false negative and false detections due to inappropriate lead selection [11]. For example, in Fig. 1, the effect of interference can be removed by including lead III in the analysis.

3.2.2 Multi-parameter Joint Analysis

In addition to the multi-lead mode in ECG analysis, we can also use multi-parameter joint analysis to further reduce false alarms. In most clinical monitoring scenarios, multiple vital sign parameters are measured from the patient, such as ECG, photoplethysmogram (PPG) from oximetry and arterial blood pressure (IBP). Normally, in one cardiac cycle, ECG records the electrical activity of the heart, whereas the synchronized PPG and IBP record the pulsatile variation in the arterial blood resulted from the corresponding heart contraction. Therefore, PPG or IBP can be used to crosscheck the heartbeat detected in ECG and discriminate artifacts from arrhythmia, correcting arrhythmia false alarm [12]. As in Fig. 6, there is a segment of

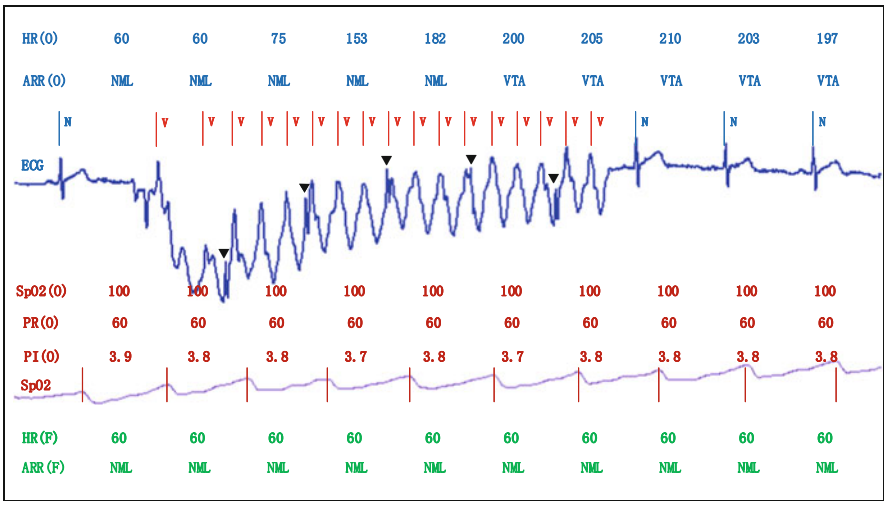

Fig. 6 Example of correcting ECG false alarms using the oximetry photoplethysmogram [12]

ECG showing the typical characteristics of ventricular tachycardia. However, the simultaneously measured PPG signal (marked as SpO2 in Fig. 6) shows no sign of arrhythmia, with the pulse amplitude and interval being uniform and consistent as in normal pulse cycles. Through the joint analysis of ECG and PPG, we recognize the ECG signal segment being artifact. Therefore, the ventricular tachycardia alarm is a false alarm and should be suppressed.

In the 2015 physionet/cinc challenge [13], the scheme of combining of ECG, PPG, and IBP signal features was proposed to suppress ICU arrhythmia false alarms. Many teams from all over the word participated in the challenge and proposed their false alarm suppression algorithms. For instance, Chengyu Liu team proposed a multifeature fusion algorithm consisting steps of multifeature extraction, feature screening, feature fusion, and a decision mechanism [14], as shown in Fig. 7. This algorithm is very effective in reducing false alarms and meanwhile simple for practical use. Machine learning was also used by many teams to solve the problem [15, 16]. However, multiple vital signs in the real clinical environment can demonstrate a huge variety of combinational patterns, posing a big challenge to data collecting and annotation.

4 ECG Meaningless Alarm and Its Solutions

As mentioned before, the alarms in the clinical setting can be divided into (1) clinically relevant alarms that correctly reflect the patient status and require clinical interventions and (2) meaningless alarms that reflect the patient status correctly but require nonclinical interventions [17–19]. As shown in Fig. 7, there is a segment of

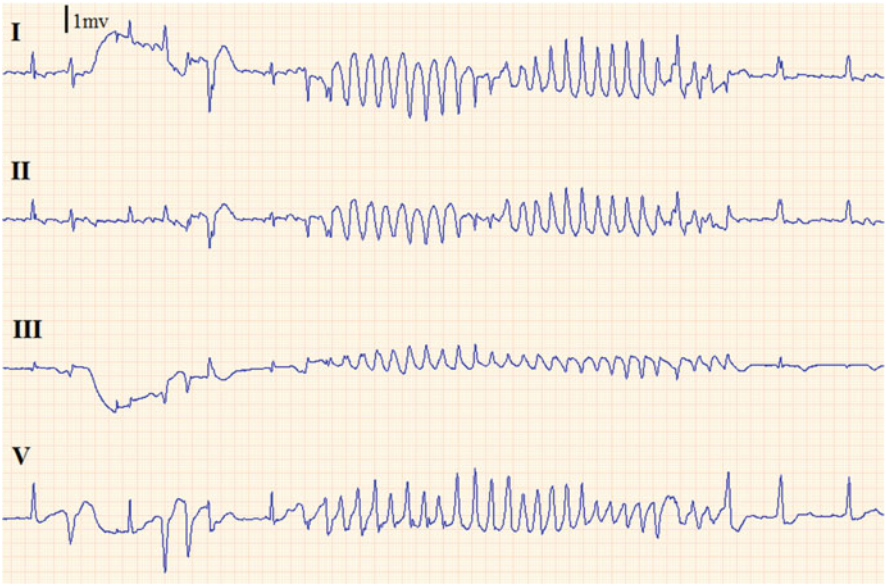

Fig. 7 Example of ventricular tachycardia alarm

torsade de pointes (TDP) ECG triggering alarm. This alarm reflects the patient's critical physiological state which needs to be treated and is a clinically relevant alarm.

In clinical practice, meaningless alarms consist of a large portion of the alarms, reducing medical staff's alarm compliance and increasing response time to the alarms, so that clinically relevant alarms might not get responded in time.

4.1 Types of ECG Meaningless Alarm

4.1.1 Meaningless Arrhythmia Alarm

During ECG monitoring, most alarms are due to ventricular arrhythmia according to Drew et al. the number of ventricular arrhythmia alarms reached 900,000 times in a 31-day period in the ICU, accounting for about half of all the alarms in the ICU. This is partly because ventricular arrhythmia in previous years was thought to be the preceding signs of ventricular tachycardia or ventricular fibrillation and requires immediate intervention. However, CAST (cardiac arrhythmia suppression trial) in 1989 showed that the treatments on ventricular rhythm could lead to increased patient mortality. After this study, patients with ventricular arrhythmia but with no obvious symptoms or no hemodynamic changes are not recommended to receive medication therapy [20]. Nowadays, ventricular arrhythmia like this usually triggers intermediate-level alarms which are often left untreated [1]. For example, Fig. 8

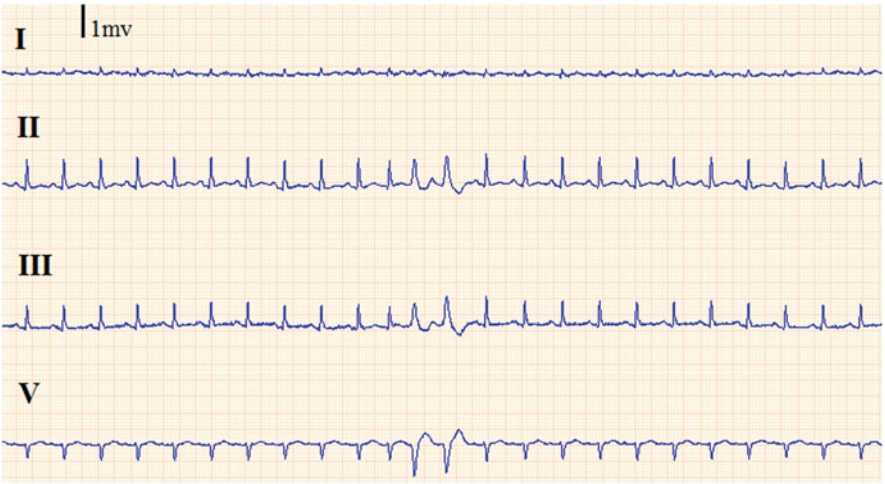

Fig. 8 Example of an occasional pair of premature

shows a pair of premature ventricular contraction which is commonly seen in the clinical monitoring. If the PVCs like this only occur occasionally, clinical interventions are usually not necessary.

4.1.2 Heart Rate Meaningless Alarm

Heart rate alarms consist of another large portion of monitor alarms. This kind of alarm is often directly triggered if heart rate exceeds the preset thresholds. Since the basal heart rate varies greatly across different patients, and for the same patient, his/her physiological state can change dramatically during the monitoring process, the over-the-limit scheme often generates a large number of false or meaningless alarms. For example, if a patient's basal heart rate is high or elevated due to drug effect, lots of over-the-limit false alarms will appear [21]. To reduce the number of false alarms, it is desirable to customize the threshold based on the patient's the basal heart rate level and adaptively adjust the threshold during the monitoring process according to the patient's condition change. In addition, the monitoring is performed using the default threshold configuration.

4.2 Solution to ECG Meaningless Alarm

4.2.1 Improvement of Clinical Alarm Management System

From the perspective of clinical application, in order to reduce the harm caused by excessive alarms from the clinical monitoring systems, the following measures are recommended for the medical institutions [5]:

1. Establish alarm system as a significant influencing factor for patient safety
2. Identify the key vital signs requiring close
3. Establish a set of standardized policies and procedures for alarm management
4. Educate the responsible medical staff and licensed independent practitioners about the importance and correct operation of alarm systems

Therefore, enhancing the alarm management regulation, and training staff on alarm management procedure can reduce the number of false alarms and relieve the impacts of alarm fatigue. For instance, with a customized threshold for each patient, heart rate over-limit alarm can be reduced effectively [22]. Furthermore, the parameter configuration, alarm mute-criteria, alarm priority rules, and sound/light function should all be set appropriately, and readjust during the monitoring process if needed. In a word, it is important to ensure that alarms are processed effectively and normatively.

4.2.2 Intelligent Alarm

The alarm fatigue problem has received more attention than before in the clinical setting. With the development of monitoring technology, intelligent alarm technology has been proposed to solve the alarm fatigue problem. Different from the traditional alarm system triggered by solely analyzing single vital sign, the intelligent alarm analyzes the patient's medical record, historical monitoring data, frequency of triggered alarms and the characteristic of user interaction so as to comprehensively evaluate the current alarm setting and patient status. The alarm strategy can be streamlined to suppress the insignificant alarms, alleviating the alarm fatigue problem. For example, for recurrent alarms for chronic AF, persistent tachycardia, bundle branch block, and pacing patient with sustained ST elevation, alarm fusion should be used to form the multiple repetitive short-term alarms into one significantly clinical relevant alarm to capture the caregiver's attention. In addition to alarm integration, multiparameter joint analysis also can help to detect the potential physiological critical events such as infection, shock, etc.

5 Discussion

Alarm fatigue is one of the most urgent problems in the ICU. The excessive alarms generated by the physiologic monitor had created a cacophonous environment, which desensitized the clinicians to the life-threatening alerts. Reducing the false alarm effectively has become the key to resolve the problem, and various solutions that had achieved excellent clinical results had been proposed. Compared with traditional single parameter analysis, multiparameter feature fusion analysis has shown great potential for false alarm rejection in the ICU.

Meanwhile, with the development of artificial intelligence, machine learning has been used to solve more and more clinical ECG monitoring problems. These AI-based approaches can be extended to solve other alarm fatigue problems. For example, machine learning can be used for disease diagnosis [23, 24], and with labeled ECG signatures, supervised learning algorithms can be used for arrhythmias identification [25]. Considering the variability of the waveform characteristics of ECG signals and the effects of noises, a large amount of expert-annotated data are required to train a viable model for use in clinical practice. With the development of big data and deep learning technology, data-driven algorithms are showing their potential for precise alarming. For example, deep learning has been used for arrhythmia detection in ECG monitoring [26, 27].

Smarter signal analysis reduces false alarms in patient monitoring and alleviates alarm fatigue. Patient monitor manufactures should work closely with clinical medical staff to develop an effective and easy-to-use alarm management system, in order to combat alarm fatigue and to improve patient safety.

References

1. Drew, B.J., Harris, P., Zègre-Hemsey, J.K., et al.: Insights into the problem of alarm fatigue with physiologic monitor devices: a comprehensive observational study of consecutive intensive care unit patients. PLoS One. **9**(10), e110274 (2014)
2. ECRI Institute: Top 10 health technology hazards for 2014. Health Dev. **42**(11), 354 (2013)
3. Kowalczyk L (2010) MGH death spurs review of patient monitors. Available from http://www.boston.com/news/health/articles/2010/02/21/mgh_death_spurs_review_of_patient_monitors/?page=1. Accessed 29 Jul 2014.
4. Joint Commission: Medical device alarm safety in hospitals. Sent. Event Alert. **50**, 1–3 (2013)
5. Joint Commission. 2016 National patient safety goals effective January 1, 2016. The Joint Commission website. http://www.jointcommission.org/assets/1/6/2016_NPSG_HAP.pdf. Accessed 16 Sep 2017.
6. Ruppel, H., Funk, M., Whittemore, R.: Measurement of physiological monitor alarm accuracy and clinical relevance in intensive care units. Am. J. Crit. Care. **27**(1), 11–21 (2018)
7. Bonafide, C.P., Lin, R., Zander, M., et al.: Association between exposure to nonactionable physiologic monitor alarms and response time in a children's hospital. J. Hosp. Med. **10**(6), 345–351 (2015)
8. Bonafide, C.P., Zander, M., Graham, C.S., et al.: Video methods for evaluating physiologic monitor alarms and alarm responses. Biomed. Instrum. Technol. **48**(3), 220–230 (2014)

9. Wu, J., Zhihong, Y., Xiangping, C.: Current status and influencing factors of bedside monitor alarms in ICU. J. Nurs. Sci. **30**(24), 20–22 (2015)
10. Mirvis, D.M., et al.: Instrumentation and practice standards for electrocardiographic monitoring in special care units. A report for health professionals by a Task Force of the Council on Clinical Cardiology. American Heart Association. Circulation. **79**, 464–471 (February 1989)
11. Jianwei, S., et al.: A four-lead real time arrhythmia analysis algorithm. Comput. Cardiol. **44**, 1–4 (2017)
12. Jianwei, S., et al.: Real-time fusion of ECG and SpO2 signals to reduce false alarms. Comput. Cardiol. **45**, 1–4 (2018)
13. Clifford, G.D., Silva, I., Moody, B., Li, Q., Kella, D., Shahin, A., Kooistra, T., Perry, D., Mark, R.G.: The PhysioNet/computing in cardiology challenge 2015: reducing false arrhythmia alarms in the ICU. Comput. Cardiol. 273–276 (2015)
14. Liu, C., Zhao, L., Tang, H.: Reduction of false alarms in intensive care unit using multi-feature fusion method. In: Computing in Cardiology. IEEE, Washington, DC (2015)
15. Kalidas, V., Tamil, L.S.: Enhancing accuracy of arrhythmia classification by combining logical and machine learning techniques. In: Computing in Cardiology. IEEE, Washington, DC (2015)
16. Eerikainen, L.M., Vanschoren, J., Rooijakkers, M.J., et al.: Decreasing the false alarm rate of arrhythmias in intensive care using a machine learning approach. In: Computing in Cardiology. IEEE, Washington, DC (2015)
17. Inokuchi, R., Sato, H., Nanjo, Y., et al.: The proportion of clinically relevant alarms decreases as patient clinical severity decreases in intensive care units: a pilot study. BMJ Open. **3**(9), e003354 (2013)
18. Siebig, S., Kuhls, S., Imhoff, M., Gather, U., Schölmerich, J., Wrede, C.E.: Intensive care unit alarms--how many do we need? Crit. Care Med. **38**(2), 451–456 (2010)
19. Siebig, S., Kuhls, S., Imhoff, M., et al.: Collection of annotated data in a clinical validation study for alarm algorithms in intensive care: a methodologic framework. J. Crit. Care. **25**(1), 128–135 (2010)
20. The Cardiac Arrhythmia Suppression Trial (CAST) Investigators: Preliminary report: effect of encainide and flecainide on mortality in a randomized trial of arrhythmia suppression after myocardial infarction. N. Engl. J. Med. **321**, 406–412 (1989)
21. Fidler, R.L., Pelter, M., Drew, B.J., et al.: Understanding heart rate alarm adjustment in the intensive care units through an analytical approach. PLoS One. **12**(11), e0187855 (2017)
22. Allan, S.H., Doyle, P.A., Sapirstein, A., et al.: Data-driven implementation of alarm reduction interventions in a cardiovascular surgical ICU. Jt. Comm. J. Qual. Patient Saf. **43**(2), 62–70 (2017)
23. Aurore, L., Rina, A., Ana, M., et al.: Distinct ECG phenotypes identified in hypertrophic cardiomyopathy using machine learning associate with arrhythmic risk markers. Front. Physiol. **9**, 213–226 (2018)
24. Simjanoska, M., Gjoreski, M., Gams, M., et al.: Non-invasive blood pressure estimation from ECG using machine learning techniques. Sensors. **18**(4), 1160–1174 (2018)
25. Casas, M., Avitia, R.L., Reyna, M.A., et al.: Evaluation of three machine learning algorithms as classifiers of premature ventricular contractions on ECG beats. In: 2016 Global Medical Engineering Physics Exchanges/Pan American Health Care Exchanges (GMEPE/PAHCE), pp. 1–6. IEEE, Washington, DC (2016)
26. Pourbabaee, B., Roshtkhari, M.J., Khorasani, K.: Deep convolutional neural networks and learning ECG features for screening paroxysmal atrial fibrillation patients. IEEE Trans. Syst. Man Cybernet. Syst. **48**(12), 1–10 (2018)
27. Wu, M.H., Chang, E.J., Chu, T.H.: Personalizing a generic ECG heartbeat classification for arrhythmia detection: a deep learning approach. In: IEEE Conference on Multimedia Information Processing and Retrieval (MIPR). IEEE, Washington, DC (2018)

Respiratory Signal Extraction from ECG Signal

Kejun Dong, Li Zhao, and Chengyu Liu

Abstract The extraction of respiration from physiological signals such as the electrocardiogram (ECG) and photoplethysmogram (PPG) has been explored for a long time. The proposed methods are mainly based on filters and features. However, the performances among methods are hardly compared and summarized. In this chapter, we focus on the studies of the typical feature-based ECG-derived respirations (EDR). The review of each method is given. The experiment is processed with rest ECG data and reference respiratory data collected synchronously over 60 s. Three parameters, waveform correlation C_1, interval correlation C_2, and respiratory rate RR, are introduced to evaluate each method under conditions of good and poor signal qualities. The results indicate that the optimal method should be determined by applications. For parameter C_1, trough envelope-based method provides the highest similarity with reference waveform (0.8426) when the signal quality is good. However, it is easily affected by the noise, decreasing to -0.3219. For parameter C_2, ECG area mean-based method has the highest similarity with intervals of reference signal (0.8162). Likewise, it performs no better than QRS complex area-based method (0.7013) when the signal quality is poor. In general, signal quality has an effect on the results of these methods.

Keywords Derived respiration · Electrocardiogram (ECG) · Feature-based

K. Dong
School of Information Science and Engineering, Southeast University, Nanjing, P. R. China

School of Instrument Science and Engineering, Southeast University, Nanjing, P. R. China

L. Zhao (✉)
School of Information Science and Engineering, Southeast University, Nanjing, P. R. China
e-mail: zhaoli@seu.edu.cn

C. Liu (✉)
The State Key Laboratory of Bioelectronics, School of Instrument Science and Engineering, Southeast University, Nanjing, China
e-mail: chengyu@seu.edu.cn

C. Liu, J. Li (eds.), *Feature Engineering and Computational Intelligence in ECG Monitoring*, https://doi.org/10.1007/978-981-15-3824-7_13

1 Introduction

Respiration is an important factor to monitor diseases, such as sleep apnea, depression, and cardiac arrest, which evaluates dysfunction by the respiratory rate [1]. In hospital, abnormal respiratory rate (RR) is recognized as an indicator of catastrophic deterioration. A study shows RR is a reliable indicator of cardiac arrest than heart rate, since 54% of all sufferers from cardiac arrest in an internal medicine unit have a high RR more than 27 breaths per minute before cardiac arrest attacks [2]. In addition, in primary care, the respiratory rate is a tool to predict the pneumonia. Many out of more than two million children dead of pneumonia every year can be prevented by early detection and treatment without delay [3]. Even though the respiratory rate provides vital information for clinical use, the measurement usually depends on the nurse counting the chest wall moves manually, which is time consuming and not accurate [4]. Not many automatic devices are available to fit the clinical robustness. Currently, the respiratory signal is usually detected using spirometry, pneumography, or plethysmography [5]. However, with the development of homecare, the devices with noninvasive and low-cost techniques are much in demand. Therefore, extracted respiratory signal from other physiological signals is more cost-effective, compared with respiration collected directly. Electrocardiogram (ECG) is a widely used signal to derive respiratory signal due to its noninvasive detection and stable signal quality [6]. The ECG-derived respiration (EDR) methods are commonly based on three respiratory modulations: amplitude modulation (AM), frequency modulation (FM), and baseline wander (BW) [7]. Heart rate increases during inspiration and decreases during expiration due to FM [8]. A more complex EDR method based on FM is to detect the duration of QRS complex [9]. The respiration effects on ECG morphology induced the methods based on peak and trough amplitudes [3, 10]. Other methods referring to morphology variation obtain respiration information from QRS complex area [11] and differences between the amplitudes of troughs and proceeding peaks [3]. Some methods based on the QRS slopes and R-wave angle variations have also been proposed [5]. We review eight EDR methods in the next parts and verify them using collected rest ECG signals. Then, we compare the result of each method with synchronized collected respiratory signal and among ECG signals with good and poor qualities.

2 Typical EDR Methods

2.1 Amplitude Variation-Based Method

Amplitude variation is the technique to extract the difference between the amplitudes of troughs (Q wave) and proceeding peaks (R wave). Karlen et al. proposed to use this method on PPG [3]. The variation of PPG amplitude induced by the respiration is caused by the corresponding decrease in cardiac output due to reduced ventricular

filling. Einthoven et al. showed that the respiration is correlated with the ECG in some certain aspects [12]. The amplitude of ECG can be changed due to breathing. Peter et al. applied Karlen's method on ECG rather than on PPG [7]. The difference on amplitude between R and Q peaks is extracted for each beat. The formula is described as follows:

$$\text{Feature} = V_{\text{R}} - V_{\text{Q}} \tag{1}$$

where V_{R} and V_{Q} are the amplitudes of R wave and Q wave (Fig. 1).

2.2 Peak Interval-Based Method

Frequency method aims to measure the time interval between consecutive peaks. Heart rate is affected by the respiration, which increases during inspiration and decreases during expiration. Such cyclic variation of heart rate is also referred to as respiratory sinus arrhythmia (RSA). Karlen et al. [3] measured the synchronization of the heartbeat with respiratory rate in PPG waveform, corresponding to the frequency variation. Orphanidou et al. [6] showed that two aspects of ECG are associated with respiration, in which RSA is one. The heart rate variability (HRV) is modulated by respiration, causing the dominant frequency in HRV changes. Charlton et al. [7] extracted the interval between two consecutive R peaks in ECG signal and assessed it with other methods. The formula is described as follows:

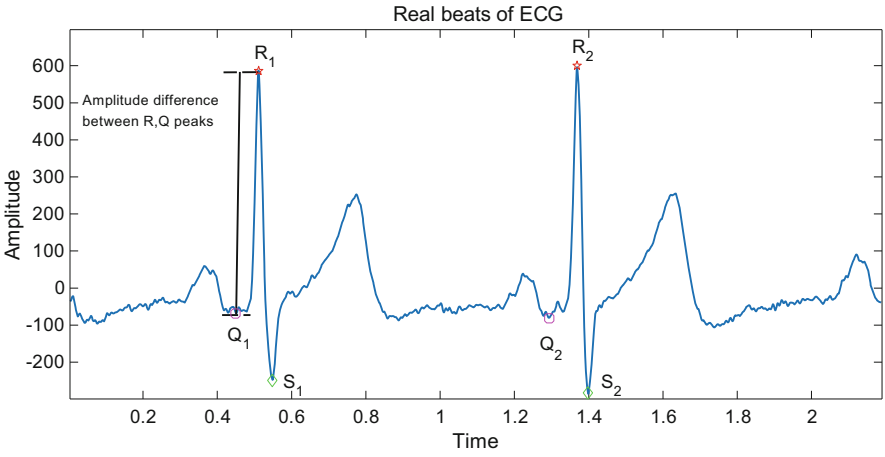

Fig. 1 Amplitude variation

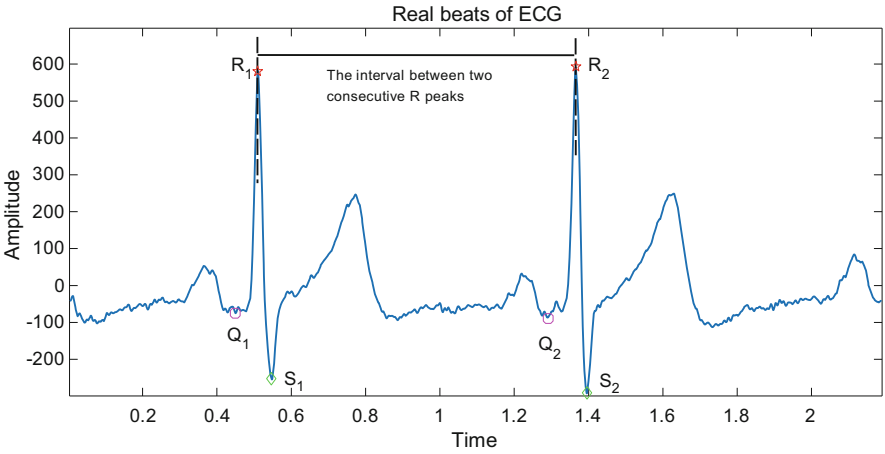

Fig. 2 Peak interval

$$\text{Feature} = T_{\text{R2}} - T_{\text{R1}} \tag{2}$$

where T_{R2} and T_{R1} are locations of the consecutive R peaks (Fig. 2).

2.3 ECG Area Mean-Based Method

The ECG mean is to obtain mean signal value over one heartbeat. Unlike the amplitude modulation method, it finds the envelope of the ECG waveform, which is greatly influenced by environment and artifact noises. Mean values show the baseline fluctuation of ECG caused by the breathing, which is more robust to noise. Ruangsuwana et al. [10] calculated these values during a window period starting from 40% of the current RR interval before a given R peak and ending before the P peak of the next beat. The baseline modulated by the respiration should observe the increase and decrease trends during the chosen window. Charlton et al. [7] modified the chosen window in their application on the assessment of different ECG-derived respiration methods. The offset is current Q peak and the end is next Q peak. The signal located in the Q-Q interval is calculated by the mean value. The calculation is based on Eq. (3).

$$\text{Feature} = \int_{Q_1}^{Q_2} V_{\text{ECG}} \tag{3}$$

where V_{ECG} is the amplitude of ECG signal (Fig. 3).

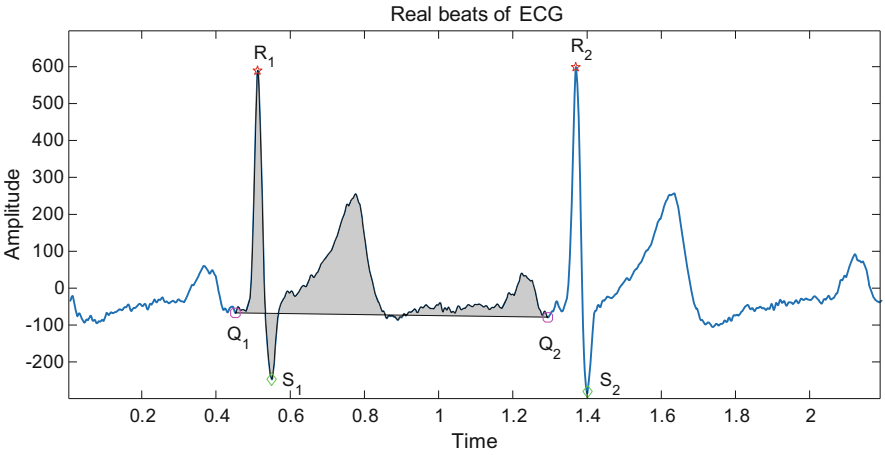

Fig. 3 The ECG mean

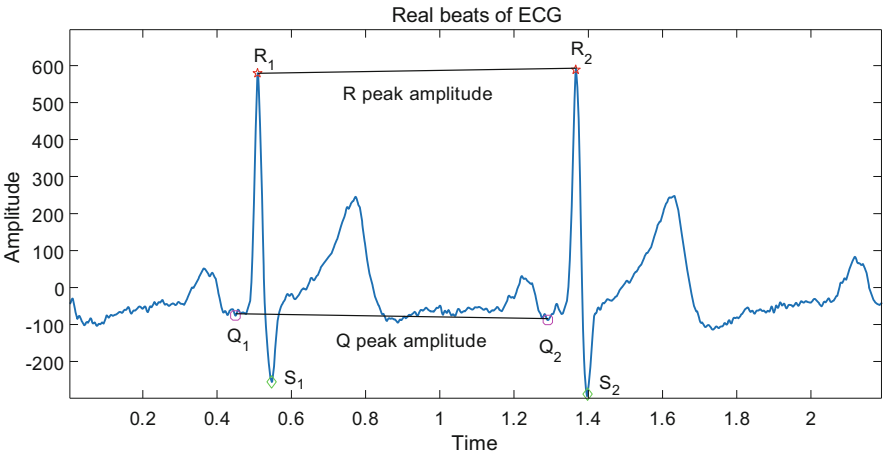

Fig. 4 Envelope method

2.4 Envelope-Based Method

The envelope method tries to capture the amplitude change under breathing works, similar to the mean value and amplitude variation methods. Karlen et al. [3] considered the amplitude variation of PPG peaks, defined as the respiratory-induced intensity variation. Same method was used on R peaks, corresponded peaks in ECG signal, by Charlton et al. [7]. Ruangsuwana et al. [10] also evaluated the QQ envelope, in addition to RR envelope (Fig. 4).

2.5 QRS Duration-Based Method

Ramya and Rajkumar [9] evaluated the ECG-derived respiration (EDR) by the duration of QRS complex, which is based on the moving window integration. Appropriate window length is important to determine the QRS complex; otherwise, other peak information is added in the integration waveform, causing difficulties in the detection process. The duration of QRS complex is equal to the width of the rising edge of the integration waveform. When this method was implemented, Charlton et al. [7] found the locations of the onset and end of the QRS complex to determine the duration. The onset is defined as the trough farthest before the R peak within the 0.1-s window. Similarly, the end is found at the trough nearest after the R peak in the 0.1 s. The time duration between the onset and end is extracted for each heartbeat. The result of the second method is shown in the following:

$$\text{Feature} = T_{\text{End}} - T_{\text{Onset}} \tag{4}$$

where T_{End} and T_{Onset} are the locations of end and onset points of QRS complex (Fig. 5).

2.6 QRS Complex Area-Based Method

The rotation of the electrical axis of the heart is caused by the heart apex stretching toward the abdomen and compressing toward the breast within the respiratory cycle [11], which leads to the change of beat morphology of ECG. Sobron et al. [11] selected the area of QRS complex as the method to obtain EDR signal. The area

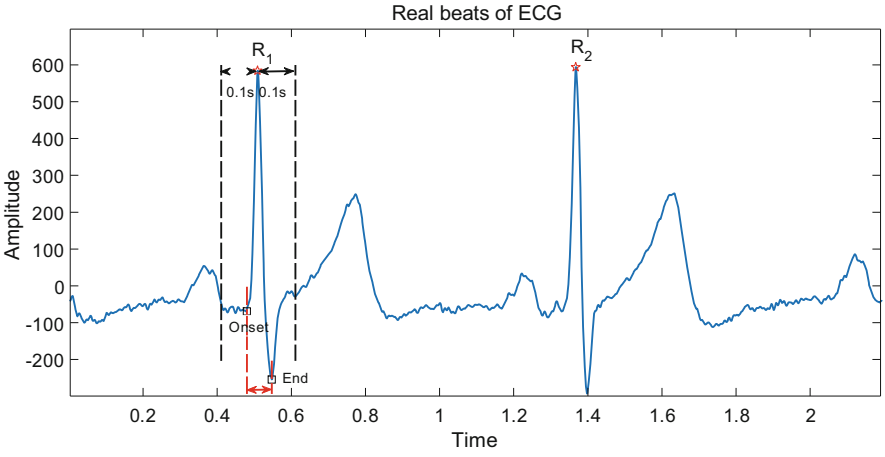

Fig. 5 QRS duration

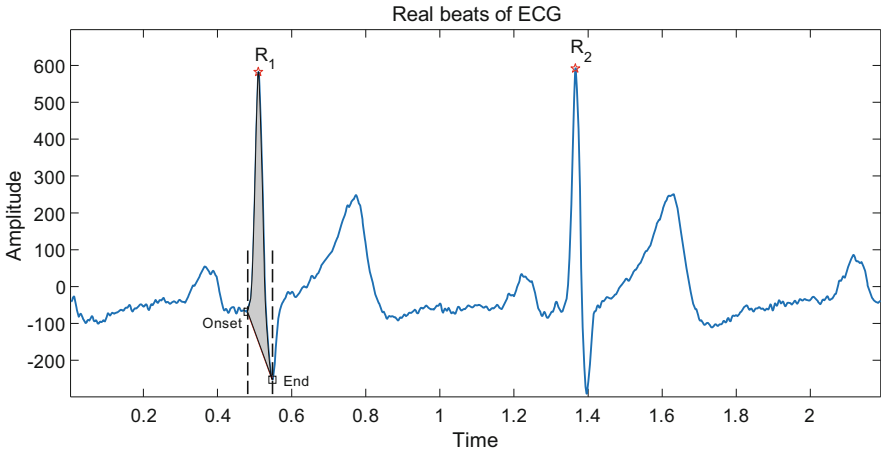

Fig. 6 QRS area

means the ECG signal within a 0.1-s window around R peak. A baseline defined as the mean value around each beat is removed. Taking this method as reference, the period between the onset and end around each beat is substituted for the 0.1-s window by Charlton et al. [7].

The definitions of onset and end refer to the duration of QRS complex. The line linking to the two points is recognized as the baseline. The area is surrounded by the ECG signal and baseline as the shadowed part in Fig. 6. The shadow area could be expressed as follows:

$$\text{Feature} = \int_{\text{Onset}}^{\text{End}} (V_{\text{ECG}} - l) \tag{5}$$

where l is the baseline linking onset and end points (Fig. 6).

2.7 *Slope-Based Method*

The slopes of QRS complex are proposed as the marks for evaluating ECG changes caused by myocardial ischemia, in particular for upward slope and downward slope of R wave [14]. It is mentioned that the respiration can affect the measurement of slopes through the amplitude modulation [13]. The algorithm used to calculate slopes is proposed in [14]. Three steps are described: first, the point, denoted n_{U} or n_{D}, is found as the maximum value of first derivate of the signal between the onset and R peak or R peak and end. Second, centered at the point, n_{U} or n_{D}, the least squares are computed within an 8-ms window. Last, the lines are fitted with maximal

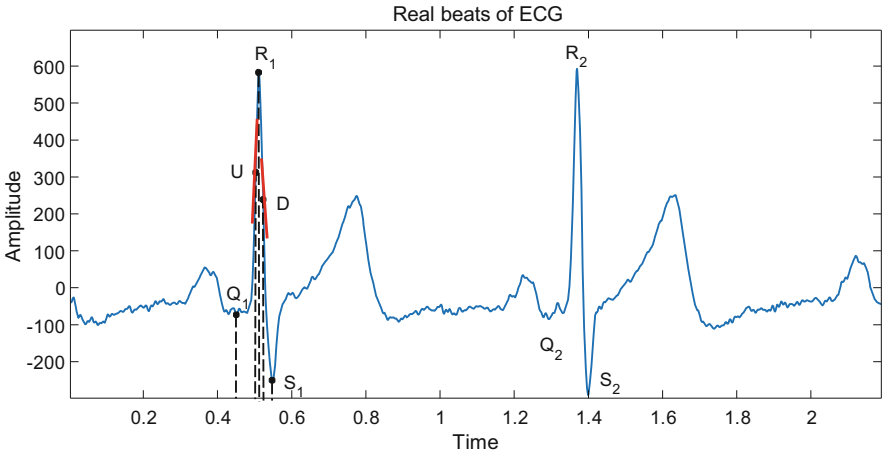

Fig. 7 A beat example with relevant upward and downward slopes

upslope or downslope to ECG signal. Figure 7 gives two fitted lines marked on the ECG signal. The derived respiration associated with two slopes is shown.

2.8 Angle-Based Method

Based on the respiration-induced influence on slope measurement, the angles formed by the lines of ECG are also associated with the respiration. Romero et al. [13] introduced three types of angles shown as follows:

θ_R: The R-wave angle, which is opposite to the R line l_R.
θ_U: The upstroke angle, opposite to the down line l_D.
θ_D: The downstroke angle, opposite to the up line l_U.

The minimal angle, θ_R, is the feature extracted to derive respiratory signal in [5] (Fig. 8). The formula of angle calculation is described below:

$$\theta = \arctan\left(\left|\frac{S_U - S_D}{1 + S_U S_D}\right|\right) \tag{6}$$

However, for clinical purpose, the time axis and voltage axis should be rescaled to match the particular case of conventional ECG tracings. The formula is modified as follows:

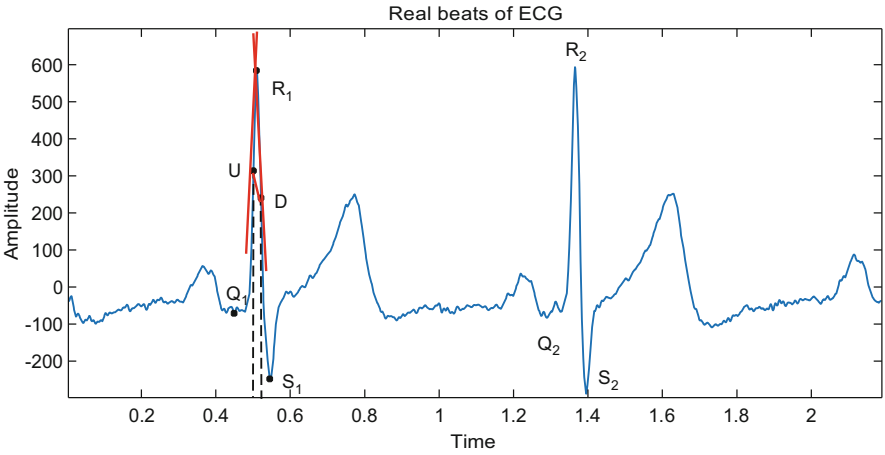

Fig. 8 The angles formed by the lines fitted with slopes of ECG

$$\theta_{\mathrm{R}} = \arctan\left(\left|\frac{S_{\mathrm{U}} - S_{\mathrm{D}}}{0.4(6.25 + S_{\mathrm{D}}S_{\mathrm{D}})}\right|\right) \tag{7}$$

3 Experiment

3.1 Datasets

ECG signals and a referenced respiratory signal are collected synchronously under 1-kHz sampling rate. Five ECG segments with good and poor qualities are picked up to verify these methods. Each segment is recorded over 60 s. The tolerance of methods to the quality of input signals can be illustrated by comparing the results.

3.2 Processing

Some noises are collected while the ECG signal is recorded by the electrodes, such as frequency interference and baseline drift. The movement of electrodes stuck to the skin results in the electrode contact noise. The polarization noise, muscle noise, and environmental artifacts are also present in the recorded ECG signal. Therefore, a low-pass filter with −3-dB cutoffs of 100 Hz is used to eliminate the very high frequencies. The main interference is eliminated by an additional 50-Hz notch filter [7]. After that, a common R-peak detection algorithm is applied on filtered ECG signal [15]. R peak is recognized as the maximum value in one beat. Q and S peaks

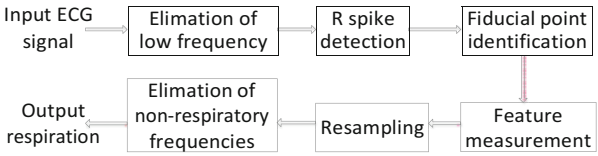

Fig. 9 The process of algorithm

are found within 0.1 s before and after detected R peak for each beat. "Respiration" is obtained after the determination of which EDR method is used. For elimination of additional noise and irregularity of the extracted signal, resampling under 5 Hz using linear interpolation is implemented. Last, nonrespiratory frequencies should be eliminated from the extracted respiration using a high-pass filter with −3-dB cutoff of 4 bpm. Figure 9 shows the structure of the algorithm process.

3.3 Evaluation Methods

The performances of extraction techniques are evaluated from three aspects:

1. A parameter, C_1, is defined as the correlation between each extracted respiratory waveform and the reference waveform. The extracted respiratory signal has been resampled at 5 Hz; however, the reference signal is collected simultaneously with the original ECG signal under a sampling rate of 1 kHz. For correlation calculation, the length of the extracted and reference signals with 60 s should stay the same.
2. A parameter, C_2, is defined as the correlation between the each peak interval of extracted respiratory signal and corresponding peak interval of the reference signal. Peaks are identified in the initial step. A moving window with the length of 1 s is used to search the peaks. The window starts at the beginning of the extracted signal and moves to the maxima found in the window range. The maxima is identified as the peak until it is the beginning of the moving window. Only the peaks closest to the peaks of corresponding beats of the reference signal are kept. The peak intervals are calculated based on the corrected peaks.
3. A parameter, RR, is defined as the respiratory rate of each extracted respiratory signal or the reference signal. Autoregressive (AR) spectral analysis is used on the fitting of respiratory frequency. Not like the Fast Fourier Transform (FFT), AR model only refers to the coefficients of the system, resulting in a smoother spectrum, which not only reduces some interference and noise but also keeps the characteristics of the spectrum. The power spectral density (PSD) based on the AR model with P coefficients is calculated by the following equation:

$$P(\omega) = \frac{\delta_\omega^2}{\left|1 + \sum_{K=1}^{P} \alpha^k e^{-jk\omega}\right|^2} \tag{8}$$

where δ_ω^2 is the PSD of white noise and α^k is the kth coefficient of the AR model.

4 Results

This section shows the derived respirations from a 60-s collected ECG signal with good and poor signal qualities using methods mentioned in above section. These results are compared with a 60-s reference respiration by three parameters, C_1, C_2, and RR. The result figures are shown below.

4.1 The Result from Signal with Good Quality

This section discusses the results derived from the ECG signal with good quality. The top figure is the reference respiratory signal collected within 60 s synchronously with the ECG signal. The others are the extracted respirations from ten methods detailed in Sect. 2. Comparing the parameters among these methods, positive values for C_1 and C_2 parameters are positive correlation, otherwise negative correlation. QRS complex area-based method and amplitude variation-based method give the same respiratory rate as reference. Also, the result waveform from QRS complex area-based method has the closest relation with the reference waveform. However, the peak intervals from downslope-based method take the highest similarity with reference, which is 0.9419 (Fig. 10).

4.2 The Result from Signal with Poor Quality

This section discusses the results derived from the ECG signal with poor quality. The orders of figures are same as in Fig. 10. Under the condition of poor signal quality, QRS complex area-based method is the optimal one, whose three parameters all achieve the highest values than any other method. However, the C_1 parameter under this method is only -0.2167, indicating that the big difference on morphology still exists. Since the highest parameter values are still lower than those in Sect. 4.1, the ECG signal quality does has an effect on the extracted results, particularly on the morphology (Fig. 11).

4.3 Statistical Results

The statistical results of parameters C_1 and C_2 are summarized in Tables 1 and 2, respectively. For each table, the first column represents ten EDR methods mentioned above. The parameters come from two-group data, which is derived from respirations under good and poor signal qualities. Each group includes five-segment ECG data collected over 60 s.

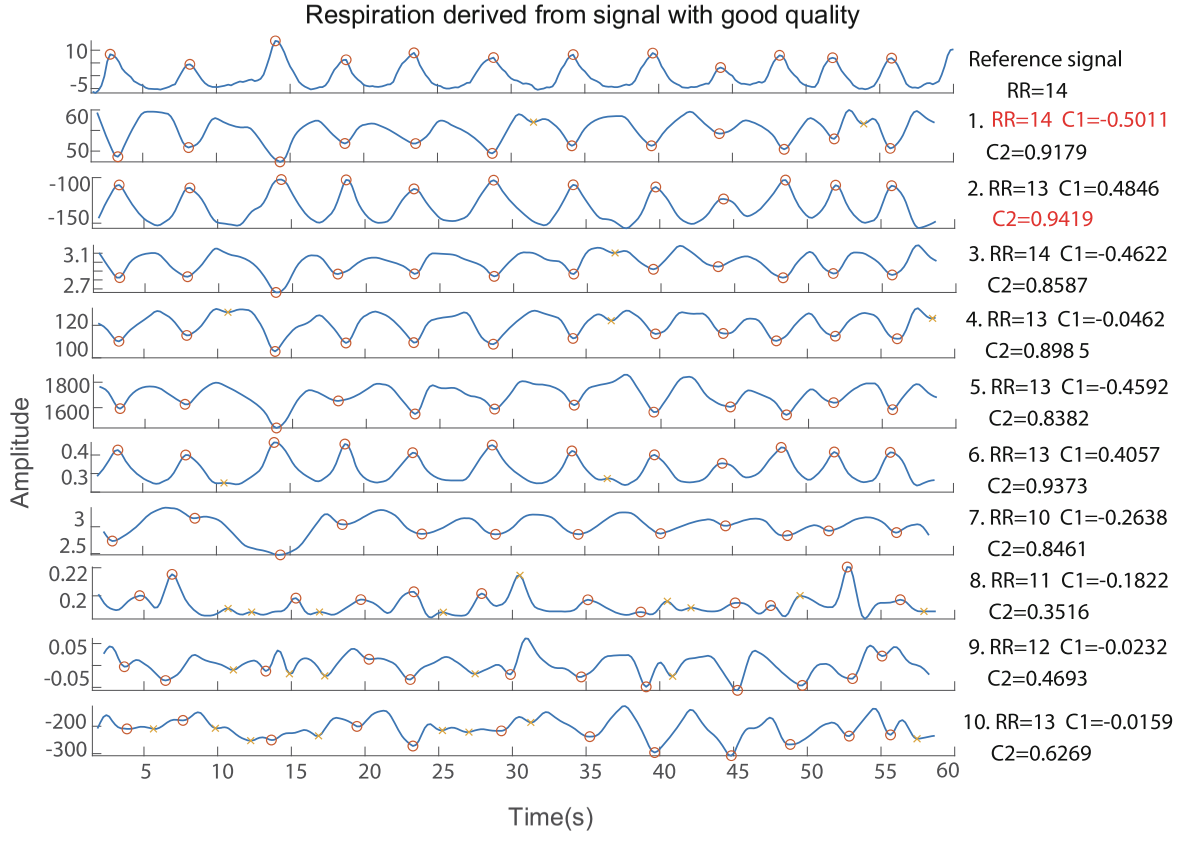

Fig. 10 The EDR signals and reference signal from good signal quality. 1. QRS complex area-based method; 2. downslope-based method; 3. amplitude variation-based method; 4. upslope-based method; 5. peak envelope-based method; 6. angle-based method; 7. peak interval-based method; 8. QRS duration-based method; 9. ECG area mean-based method; 10. trough envelope-based method. The red circles represent the selected peaks for interval calculation. The yellow crosses represent the wrong detected peaks. The peaks marked are not used for calculation of respiratory rate

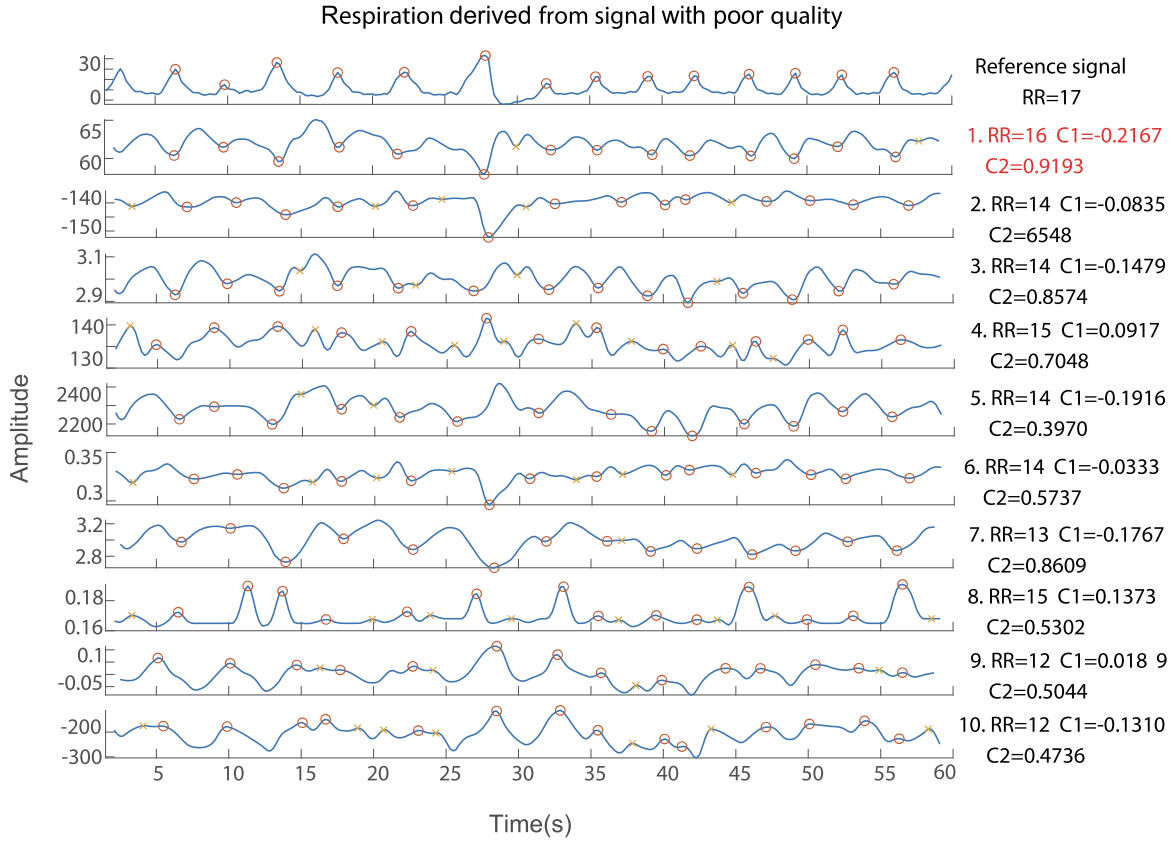

Fig. 11 The EDR signals and reference signal from poor signal quality. 1. QRS complex area-based method; 2. downslope-based method; 3. amplitude variation-based method; 4. upslope-based method; 5. peak envelope-based method; 6. angle-based method; 7. peak interval-based method; 8. QRS duration-based method; 9. ECG area mean-based method; 10. trough envelope-based method. The red circles represent the selected peaks for interval calculation. The yellow crosses represent the wrong detected peaks. The peaks marked are not used for calculation of respiratory rate

Table 1 The average and standard deviation values of parameter C_1 for each EDR method on good and poor signal conditions

	Good signal		Poor signal	
EDR methods	Mean	Std	Mean	Std
M1	−0.6508	0.1795	−0.3564	0.2130
M2	−0.0178	0.1217	0.1402	0.1170
M3	−0.4882	**0.0679**	−0.4656	0.1100
M4	−0.6105	0.2008	−0.3008	0.1956
M5	−0.1904	0.1451	−0.1017	0.1375
M6	0.2277	0.1202	0.2282	0.1300
M7	−0.7599	0.1049	**−0.5043**	0.1184
M8	−0.8038	0.0884	0.2777	**0.0992**
M9	0.4756	0.7755	−0.2790	0.2400
M10	**0.8426**	0.1006	−0.3219	0.1903

M1: QRS complex area-based method; M2: downslope-based method; M3: amplitude variation-based method; M4: upslope-based method; M5: peak envelope-based method; M6: angle-based method; M7: peak interval-based method; M8: QRS duration-based method; M9: ECG area mean-based method; M10: trough envelope-based method

The two bold values in the Mean columns indicate the respiration waveform derived from the methods and give the largest similarity with the reference waveform on good and poor signal conditions, respectively. In addition, the two bold values in the Std columns indicate parameter C_1 values of these two methods and give the relative stability compared to other methods

Table 2 The average and standard deviation values of parameter C_2 for each EDR method on good and poor signal conditions

	Good signal		Poor signal	
EDR methods	Mean	Std	Mean	Std
M1	0.4254	0.6120	**0.7013**	0.2118
M2	0.5283	0.2563	0.1491	0.4019
M3	0.6669	0.3166	0.5934	0.4019
M4	0.6793	0.4961	0.3959	0.4333
M5	0.5722	0.1916	0.4086	0.3369
M6	0.6240	**0.1597**	0.3918	0.4088
M7	0.7357	0.3648	0.5795	0.4194
M8	0.8148	0.3056	0.5460	0.1937
M9	**0.8162**	0.1877	0.5075	**0.1566**
M10	0.7263	0.4061	0.6478	0.1833

M1: QRS complex area-based method; M2: downslope-based method; M3: amplitude variation-based method; M4: upslope-based method; M5: peak envelope-based method; M6: angle-based method; M7: peak interval-based method; M8: QRS duration-based method; M9: ECG area mean-based method; M10: trough envelope-based method

Similar to Table 1, the two bold values in the Mean columns indicate the intervals of extracted respiration from the methods and are most similar with intervals of reference respiration. And the two bold values in the Std columns indicate parameter C_2 values of these two methods and give the relative stability compared to other methods

Table 1 reveals the performance of each method on waveform correlation. The two bold values in the Mean columns indicate the respiration waveform derived from the methods and give the largest similarity with the reference waveform on good and poor signal conditions, respectively. In addition, the two bold values in the Std columns indicate parameter C_1 values of these two methods and give the relative stability compared to other methods. Compared with good signal quality, the results under poor signal quality have worse performance in general. Peak interval-based method is the best one to deal with poor signals, but only half similar to the reference waveform. Trough envelope-based method provides a better waveform similarity when it is applied on good signals, whose correlation achieves 0.8426. However, it is easily affected by the signal quality. Amplitude variation-based method is the stable one not affected by the signal quality.

Table 2 reveals the performance of each method on interval correlation. Similar to Table 1, the two bold values in the Mean columns indicate the intervals of extracted respiration from the methods and are most similar with intervals of reference respiration. And the two bold values in the Std columns indicate parameter C_2 values of these two methods and give the relative stability compared to other methods. ECG area mean-based method and QRS complex area-based method provide the highest similarity with the peak intervals of reference signal under the situation of good and poor signals. Unlike the big influence of signal quality on waveform correlation, no big difference is achieved on the largest interval correlations of two signal qualities, which is 0.8162 and 0.7013, respectively. From the view of stability, ECG area mean-based method is still the relatively optimal one regardless of the signal quality.

5 Conclusion

In this chapter, we review some typical feature-based ECG-derived respiration methods. The dataset, which consists of the ECG data and respiratory signal collected synchronously, is classified into two groups: data with good and poor qualities. Five segments of each group are picked up to evaluate each method. Three parameters are introduced for the evaluation: waveform correlation, C_1; interval correlation, C_2; and respiratory rate, RR. The performances of methods keep consistency on waveform and interval correlations. For good signal quality, peak interval-based and QRS duration-based methods are in the top three ranking of both C_1 and C_2. However, they cannot fit the poor signal quality very well. In contrast, QRS complex area-based and amplitude variation-based methods perform better in the poor situation. If the stability is taken as the first factor to consider, amplitude variation-based method is relatively optimal one under parameter C_1, no matter what signal quality is. For parameter C_2, ECG area mean-based method is the relatively stable one in two-group data. According to the statistical results above, methods tend to provide high similarities with reference respiration versus stabilities under different signal qualities. Therefore, the method recognized as optimal should be determined by different applications.

In future studies, the size of each data group can be expanded to verify the performance of each method, which is more reliable. The data used to test are rest ECG data, which are not reliable for clinical applications since they are usually recorded in a short period of time. Therefore, dynamic data as a long-term data recorded continuously over 24 h or more should be focused to further explore each method in future analysis.

6. Appendix

6.1. The Least Square Method

The least square method finds the optimal estimation, which satisfies the minimum sum of squares with original data. It is described as follows.

Assume that a regression equation with two unknown parameters, β_0 and β_1, is defined as follows:

$$y = \beta_0 + \beta_1 x \tag{9}$$

Given a sequence of sample input data $\{x_1, x_2, \ldots, x_n\}$ and corresponding output data $\{y_1, y_2, \ldots, y_n\}$, where n denotes the length of data. The estimated regression equation is as follows:

$$\widehat{y} = b_0 + b_1 x \tag{10}$$

where b_0 and b_1 are the estimates of β_0 and β_1, respectively.

The aim is to get the parameter estimates satisfying the condition:

$$\min \sum (y_i - \widehat{y}_i)^2 \tag{11}$$

where y_i is the corresponding value of original data x, and $\widehat{y}_i$ is the estimated corresponding value of original data x.

Combining Eq. (10) with Eq. (11), the condition is rewritten as follows:

$$\min \sum [y_i - (b_0 + b_1 x_i)]^2 \tag{12}$$

Taking the partial derivative of b_0, we get the expression as follows:

$$b_0 = \overline{y} - b_1 \overline{x} \tag{13}$$

where $\overline{y}$ and $\overline{x}$ are the average values of y and x, respectively.

Substitute Eq. (13) for b_0 in Eq. (12), and the partial derivative of b_1 is obtained as follows:

$$b_1 = \frac{\sum(y_i - \overline{y})(x_i - \overline{x})}{\sum(\overline{x} - x_i)^2} \tag{14}$$

References

1. Gravelyn, T.R., Weg, J.G.: Respiratory rate as an indicator of acute respiratory dysfunction. JAMA. **244**(10), 1123–1125 (1980)
2. Birrenkott, D., MAF, P., Watkinson, P.J., Clifton, D.A.: A robust fusion model for estimating respiratory rate from photoplethysmography and electrocardiography. IEEE Trans. Biomed. Eng. **65**(99), 2033 (2018)
3. Walter, K., Srinivas, R., Mark, A.J., Dumont, G.A.: Multiparameter respiratory rate estimation from the photoplethysmogram. IEEE Trans. Biomed. Eng. **60**(7), 1946–1953 (2013)
4. Charlton, P., Birrenkott, D.A., Bonnici, T., Pimentel, M.A.F., Johnson, A.E.W., Alastruey, J., Tarassenko, L., Watkinson, P.J., Beale, R., Clifton, D.A.: Breathing rate estimation from the electrocardiogram and photoplethysmogram: a review. IEEE Rev. Biomed. Eng. **11**, 2–20 (2017)
5. Lázaro, J., Alcaine, A., Romero, D., Gil, E., Laguna, P., Pueyo, E., Bailón, R.: Electrocardiogram derived respiratory rate from QRS slopes and r-wave angle. Ann. Biomed. Eng. **42**(10), 2072–2083 (2014)
6. Orphanidou, C., Fleming, S., Shah, S.A., Tarassenko, L.: Data fusion for estimating respiratory rate from a single-lead ECG. Biomed. Signal Process. Contr. **8**(1), 98–105 (2013)
7. Charlton, P.H., Bonnici, T., Tarassenko, L., Clifton, D.A., Beale, R., Watkinson, P.J.: An assessment of algorithms to estimate respiratory rate from the electrocardiogram and photoplethysmogram. Physiol. Meas. **37**(4), 610–626 (2016)
8. Zhang, P.Z., Tapp, W.N., Reisman, S., Natelson, B.H.: Respiration response curve analysis of heart rate variability. IEEE Trans. Biomed. Eng. **44**(4), 321 (1997)
9. Ramya, K., Rajkumar, K.: Respiration rate diagnosis using single lead ECG in real time. Glob. J. Med. Res. **13**, 1 (2013)
10. Ruangsuwana, R., Velikic, G., Bocko, M.: Methods to extract respiration information from ECG signals. In: 2010 IEEE International Conference on Acoustics, Speech and Signal Processing, pp. 570–573. IEEE, Washington, DC (2010)
11. Sobron, A., Romero, I., Lopetegi, T.: Evaluation of methods for estimation of respiratory frequency from the ECG. Comput. Cardiol. **37**, 513 (2010)
12. Einthoven, W., Fahr, G., De Waart, A.: On the direction and manifest size of the variations of potential in the human heart and on the influence of the position of the heart on the form of the electrocardiogram. Am. Heart J. **40**(2), 163–211 (1950)
13. Daniel, R., Michael, R., Pablo, L., Olle, P., Esther, P.: Depolarization changes during acute myocardial ischemia by evaluation of QRS slopes: standard lead and vectorial approach. IEEE Trans. Biomed. Eng. **58**(1), 110–120 (2010)
14. Esther, P., Leif, S., Pablo, L.: QRS slopes for detection and characterization of myocardial ischemia. IEEE Trans. Biomed. Eng. **55**(2), 468–477 (2008)
15. Pan, J., Tompkins, W.J.: A real-time QRS detection algorithm. IEEE Trans. Biomed. Eng. **32**(3), 230–236 (2007)

Noninvasive Recording of Cardiac Autonomic Nervous Activity: What Is Behind ECG?

Yike Zhang, Chang Cui, and Minglong Chen

Abstract The autonomic nervous system, which is divided into the sympathetic nervous system and the parasympathetic nervous system, takes part in various physiological processes. The assessment of autonomic nervous activity is necessary for disease diagnosis and risk stratification for patients. Noninvasive testing approaches based on ECG, such as heart rate variability, heart rate turbulence, baroreflex sensitivity, and skin sympathetic nerve activity, are briefly introduced in this chapter to enlighten broader applications of these physiological signals in clinical work.

Keywords Autonomic nervous system · Heart rate variability · Skin sympathetic nerve activity

1 Introduction

Since the invention of electrocardiogram (ECG) in 1895, this technique has brought much useful information for doctors in clinical practice. Myocardial ischemia can be detected early when the sign of the depressed ST presented in the ECG. Patients with atrial fibrillation can start anticoagulation therapy soon when diagnosed by irregular RR intervals. ECG helps in the diagnosis with the advantages of simplicity, accuracy, and real-time performance. However, the information behind ECG is more than that. Data are available reflecting the activity of the autonomic nervous system, which plays an indispensable role in the course of different diseases. Herein, from the physiological point of view, we summarize several parameters and methods to measure the autonomic nervous functions via noninvasive approaches, shedding light on broader applications of ECG-based techniques.

Y. Zhang · C. Cui (✉) · M. Chen (✉)
Department of Cardiology, The First Affiliated Hospital of Nanjing Medical University, Nanjing, China
e-mail: cuichang@njmu.edu.cn; chenminglong@njmu.edu.cn

C. Liu, J. Li (eds.), *Feature Engineering and Computational Intelligence in ECG Monitoring*, https://doi.org/10.1007/978-981-15-3824-7_14

2 The Autonomic Nervous System (ANS) and Cardiovascular Diseases

2.1 Functions of ANS

The autonomic nervous system, which innervates internal organs unconsciously, controls the homeostasis of the human body. It consists of two branches—the sympathetic nerve and the parasympathetic nerve. Heart, lungs, stomach, intestine, pupils, endocrine, and exocrine glands are innervated by ANS. Activation of the sympathetic nerves leads to a state of "fight or flight" response, referring to overall elevated activity and attention with increased blood pressure and heart rate, bronchodilation, and glycogenolysis. On the other hand, the parasympathetic nerves promote the "rest and digest" processes, with lower heart rate and blood pressure, bronchoconstriction, gastrointestinal peristalsis, etc. Normally, both branches have nerve tonic all the time but can fluctuate depending on different conditions [1].

The sympathetic nervous system has a distinguished feature. The postganglionic neurons of the sympathetic system distribute with spinal nerves. Sympathetic fibers account for 8% of the fibers of a spinal nerve. In consequence, the effectors in the skin, such as blood vessels and sweat glands, are also innervated by sympathetic nerve fibers.

2.2 ANS in the Cardiovascular System

ANS controls the balance of an organism and is critical to maintain sinus rhythm and sustain blood circulation. The heart and brain form interconnected cardio–neural hierarchy to modulate the physiological functions of the heart—chronotropy, dromotropy, inotropy, and lusitropy [2].

The sympathetic nerves controlling the heart locate in the stellate ganglion [3]. Playing the role like a transfer station, it receives preganglionic sympathetic impulses from the spinal cord and outputs efferent neural responses toward the heart through middle cervical ganglion. Efferent postganglionic fibers travel along the coronary vasculature, into the epicardial regions, and toward the endocardium and then control the atrial and the ventricular nerves. Stellate ganglion stimulation results in increased heart rate, conduction velocity, and contraction.

The parasympathetic nerve functions mainly by the cervical vagal nerve, which are divided into the superior and inferior cardiac nerves, entering the heart via the cardiac plexus and innervating the sinoatrial node, the atrioventricular node, and the ventricle. The parasympathetic nerves slow the heart rate and conduction, and reduce the blood pressure to balance the effect of sympathetic nerves.

2.3 ANS Dysfunction and Related Diseases

ANS dysfunction can derive from both primary or secondary causes [3]. Primary causes, such as diabetes mellitus, directly impair myelinated vagus branches, sympathetic nerves axons and ganglia, resulting in autonomic neuropathy. Many organs and systems could be involved such as digestive, cardiovascular, and urinary system. Symptoms such as diarrhea, constipation, nausea, dizziness, and night sweats could present in these patients. Besides, syndromes of primary autonomic disorder or failure (pure autonomic failure, Parkinson's disease, multiple system atrophy, etc.) also damage the ANS, manifested by orthostatic hypotension and other symptoms. The activation of the autoimmune system, which targets ganglionic receptors, might be the underlying pathogenesis for these primary disorders.

Secondary causes such as myocardial infarction (MI), heart failure, and sleep apnea, bring about the secondary alterations in ANS. These are mostly due to the compensation of the organism to maintain the balance. In cases such as MI or cardiomyopathy, the ANS acutely responses to the dysfunction of the circulation system, increasing the sympathetic nerve activity and decreasing the vagus nerve activity to ensure the cardiac output. However, elevated catecholamines and cytokines are relatively persistent, which aggravate the primary diseases.

3 Parameters for the Function Analysis of ANS

ECG provides us with the physiological information of the heart, including heart rate, rhythm, and cardiac conduction. To go one step further, we can extract information on cardiac ANS. Herein, we list several parameters to assess ANS function and introduce each one from physiology basis, analysis methods to clinical applications.

3.1 Heart Rate Variability (HRV)

3.1.1 Introduction

HRV refers to the oscillation in the intervals between consecutive beats [4]. The study on HRV has a long history. In 1965, Hon and Lee [5] first found that there were alterations in interbeat intervals before fetal distress. In 1970s, to identify the autonomic neuropathy diabetic patients, Ewing et al. [6] devised tests to measure the short-term RR variation. After that, as a convenient method to detect autonomic nerve function, HRV was broadly used in different clinical circumstances, such as coronary artery diseases, ventricular arrhythmia, heart failure, sleep apnea, asthma,

fetal distress, and so on. HRV data are also included in a 24-h Holter monitoring report as routine.

3.1.2 Physiological Basis

The disturbance of the body can affect the center (e.g., vasomotor and respiratory center) or peripheral (e.g., oscillation in arterial pressure and respiratory movements) oscillators. These oscillators dominate the tone and balance of the sympathetic nerve and vagal nerve so as to affect the rhythm fluctuations of the sinus node. As a result, the variability of heart rhythm can provide us with information about the modulating effect of the autonomic nervous system [4].

3.1.3 Analysis Methods

The broad applications of HRV are attributed to its advanced analysis approaches, including the time-domain, the frequency-domain, and nonlinear analysis.

The time-domain analysis of HRV calculates the RR intervals by mathematical or statistical operations to output SDNN (standard deviation of normal RR intervals), RMMD (root mean square of RR interval differences), and PNN50 (the percentage of normal RR intervals that differ by more than 50 ms). The frequency-domain analysis of HRV is based on the spectral analysis of a sequence of RR intervals, and demonstrates information on the power distributed as a function of frequency. Three main spectral components are distinguished in a spectrum analysis: very low frequency (VLF), low frequency (LF), and high frequency (HF), representing the balance of sympathetic and vagal nerves. Clinical procedures and physiological experiments such as electrical vagal stimulation, muscarinic receptor blockade, and vagotomy have shown that the efferent vagal activity is a major contributor to the HF component. On the other hand, the LF component is more controversial because both sympathetic and vagal nerves can affect it. Therefore, normalization like LF/HF ratio is required to study the sympathetic activity. Except for the time-domain and frequency-domain analysis, nonlinear analysis approaches, such as detrended fluctuation analysis (DFA), are also feasible but have not been widely used.

3.1.4 Clinical Applications

Decreased HRV has an association with both sudden cardiac death and non-sudden death in myocardial infarction and chronic left ventricular dysfunction, independent of the ejection fraction of left ventricle [7]. In a recent study of 9550 patients, reduced RMSSD was also associated with elevated risk across a range of established cardiovascular risk factors, including inflammatory markers, blood lipids and glucose, as well as blood pressure [8]. However, at present, HRV cannot precisely

predict the risk of sudden cardiac death. The pathophysiological relationship between reduced HRV and elevated mortality remains unclear.

There are also many pitfalls and limitations of HRV [9]. First, long-term HRV is highly dependent on the daily activity of the patient. Second, the cutoff value of 0.15 Hz to separate LF and HF is also questionable. Besides, respiratory sinus arrhythmia (RSA) also interferes with the frequency component of HRV. As a result, HRV provides us with a measurement approach of ANS control of the heart; however, the modulation of ANS is not that simple, and more work is required to explain the HRV data precisely.

3.2 Heart Rate Turbulence (HRT)

3.2.1 Introduction

Heart rate turbulence refers to the fluctuations of heart rate after a premature ventricular contraction (PVC). This term was first defined by Schmidt in 1999 [10].

3.2.2 Physiology Basis

PVCs interrupt the hemodynamics evidently because the premature beat pumps much lower volume. The decreased cardiac output leads to lower blood pressure, which is then detected by the baroreceptor in the carotid artery. Therefore, such a baroreflex stimulates the sympathetic nerve and inhibit the vagus nerve, leading to blood pressure rising. This overcompensation results in the increased tone of the vagus nerve and decreased tone of sympathetic nerves. The modulation of the autonomic nervous system is reflected in the fluctuation after PVC and termed as HRT.

3.2.3 Analysis Methods

Turbulence onset (TO) and turbulence slope (TS) are mostly used to evaluate the chronotropic response of sinus rhythm to PVCs. TO refers to the immediate initial acceleration quantified by the relative change of RR intervals immediately after a PVC compared with before a PVC. TS, the steepest regression line between RR interval count and RR interval duration, is used to quantify the speed of the subsequent deceleration following a PVC. A 24-h Holter, which can monitor both the PVC and the heart rate, is an accessible tool to evaluate HRT.

3.2.4 Clinical Applications

Schmidt et al. [10] suggested that HRT was a powerful tool to predict the mortality after acute myocardial infarction. Patel et al. [11] found that HRT was significantly associated with incident congestive heart failure in asymptomatic, older adults. The association between HRT and atrial fibrillation symptoms has also been reported, indicating that compromised baroreflex sensitivity could come with severe AF symptoms [12]. Except for the wide use in cardiovascular diseases, abnormal HRT has also been reported to be associated with liver cirrhosis deterioration [13] and systemic lupus erythematosus (SLE) [14].

3.3 Baroreflex Sensitivity (BRS)

3.3.1 Introduction

Baroreflex plays a pivotal role in regulating cardiovascular homeostasis. BRS, which quantifies the magnitude of the cardiac cycle changes following arterial pressure variations, evaluates the efficiency of this reflex and implies the regulation of both the vagus and the sympathetic nerves.

3.3.2 Physiology Basis

The circuit of baroreflex is achieved via baroreceptor at carotid sinus and aorta arch. The signal of reduced blood pressure is sent to the central nervous system, reflectively activating the sympathetic nerves, thus bringing about the rise in ventricular contractility, heart rate, and peripheral vascular resistance to maintain normal blood pressure. On the contrary, elevated blood pressure leads to the activation of the parasympathetic nerve, with decreased contractility, heart rate, and peripheral vascular resistance. There is a few seconds delay for sympathetic nerves, but no time lag for the vagus nerve, which means the responses to elevated and reduced arterial pressure are asymmetrical in man [15]. As a result, when it comes to BRS, the baroreflex slopes obtained by decreased arterial pressure are lower compared with increased arterial pressure.

3.3.3 Analysis Methods

There are several ways to measure baroreflex or the heart rate change due to blood pressure. First, vasoactive drugs, such as phenylephrine, are used to stimulate the alteration in blood pressure. To avoid the usage of drugs, noninvasive methods such as Valsalva maneuver and mechanical manipulation of carotid baroreceptors are also used. Besides, spontaneous fluctuation of heart rate and blood pressure can also be

monitored to assess BRS with advanced wearable devices. Similar to the analysis of HRV, the time-domain and frequency-domain analysis are two basic methods to study BRS [16].

3.3.4 Clinical Applications

At the beginning, BRS was used to study the risk of patients with previous myocardial infarction. The ATRAMI study, which enrolled nearly 1300 patients to study the relationship between cardiac mortality and autonomic nerve dysfunction, showed that reduced BRS was an independent predictor of total cardiac mortality [17]. BRS has also been established with prognostic value in chronic heart failure [18, 19].

Overactivation of autonomic nerves can also be found by BRS. Postural tachycardia syndrome (POTS) is a drastic hemodynamic alteration after a postural change, which means the increase of the heart rate is over 40 bpm, or the maximum heart rate is more than 120 bpm. Li et al. [20] found that BRS, which was positively correlated with HR change in POTS group ($r = 0.304$, $P < 0.05$), could be a predictor for the short-term outcome of POTS in children.

Population divergence in BRS was revealed in a recent cross-sectional study. Athletes ($n = 30$) and nonathlete boys ($n = 30$) aged between 10 and 19 years were recruited. The BRS index for athletes was higher than that in nonathlete boys, implying that exercise may improve the autonomic nervous function [21].

3.4 Skin Sympathetic Nerve Activity (SKNA)

3.4.1 Introduction

Skin sympathetic nerve activity (SKNA) is a novel method to record the signals of sympathetic nervous noninvasively. Direct nerve recordings are the gold standard to measure nerve activity, but anesthesia and invasive procedures prevent its widespread application in conscious populations. Recently, it has been found that it is feasible to test the sympathetic nerve activity by electrodes in the subcutaneous space of the left thorax. Subcutaneous nerve activity (SCNA) is in correlation with the stellate ganglion nerve activity, which can become a micro-invasive method to evaluate the nerve function [22]. Furthermore, they observed that SKNA, which can be monitored noninvasively like ECG, correlated well with the stellate ganglion nerve activity (SGNA), providing a brand new avenue to monitor the real-time autonomic nerve activity.

The feasibility was first proved in the canine experiment. SKNA correlated morphologically with heart rate acceleration and preceded the onsets of ventricular arrhythmia. Based on these animal studies, they hypothesized that it is practical to record both SKNA and ECG simultaneously in humans. They tested on four

different groups of patients to stimulate or inhibit their ANS and determined that SKNA could become a useful tool to estimate sympathetic tone in humans [23].

3.4.2 Physiological Basis

The postganglionic neurons of the sympathetic system travel along with the 31 pairs of spinal nerves, which means sympathetic nerve fibers are also distributed in the skin. The upper thorax skin is innervated by sympathetic nerves extended from the stellate ganglion, which is in charge of the sympathetic fibers to the heart. As a result, as an alternative way to measure SGNA, SKNA can be recorded totally noninvasively by conventional ECG electrodes simultaneously with ECG in patients. The bursts of SKNA suggest the activation of sympathetic nerves. SKNA, ECG, and heart rate can be illustrated in a single time axis, and the disturbance of SKNA can be analyzed with the fluctuation of heart rate, providing a clear view of ANS controlling the heart rhythm.

3.4.3 Analysis Methods

By conventional ECG electrodes, the SKNA signal can be recorded simultaneously with ECG. The sampling rate of SKNA is 10,000 Hz. The same electrical signal filtered between 0.5 and 150 Hz displays ECG, while the bandpass from 500 to 1000 Hz shows SKNA.

For the quantitative analysis of SKNA, the average voltage of SKNA (aSKNA) and the integrated SKNA (iSKNA) are often used. The aSKNA is obtained by dividing the total voltage of SKNA by the number of digitized samples over a time window. For example, the total voltage of all samples in 10 s was 200,000 μV (sampling rate 10,000 Hz), then the aSKNA was 2 μV. Similarly, iSKNA was obtained by integrating the voltage of digitized data over a time window of 100 ms, which is used in the figure, not in statistical analysis [23].

SKNA recordings can sometimes be affected by artifacts. First, the movement artifacts can produce evident morphological changes within the signal that can disturb data processing and analysis. Patients should be restrained when tested to ensure the quality of the signals. Timestamps can be used to record the movements of the patients. Artifacts related to this disturbance will be evaluated in the analysis process. If no obvious ECG alteration is shown, this piece of the signal will be counted as artifacts and then eliminated. Second, muscle artifact can also interfere with the SKNA signals. The surface electromyography (EMG) signal is between 0 and 500 Hz. Even though most muscle activities had a frequency of <100 Hz, a small part of muscle activities were over 400 Hz. To increase the specificity of detection, bandpass filtering can be useful to acquire SKNA signals. However, no data to date shows if there is any overlap between EMG and SKNA. Besides, SKNA varies depending on electrode contact, electromagnetic interference, and the site of

the recording. Therefore, a uniform method to measure SKNA has to be established [24].

3.4.4 Clinical Applications

SKNA was a novel way to evaluate the real-time sympathetic nervous activity in a noninvasive way. Cold water pressor test (CPT) and Valsalva maneuver are two common ways to stimulate the autonomic nervous system. SKNA significantly increased during the CPT and Valsalva maneuver and then reduced significantly during recovery [23]. Furthermore, the fluctuation of ANS could be displayed in a timeline in other situations, such as tilt table test (TTT) [24], sedative drugs [25], and vagal nerve stimulation [26], providing us with a clear understanding of ANS physiology.

SKNA was also employed widely in recent clinical studies. SKNA was first applied in patients with arrhythmias to identify the relationship between ANS and the progress of arrhythmias. SKNA was observed to increase at onset and termination of atrial tachycardia (AT) and atrial fibrillation (AF) [27]. The acceleration of the ventricular rate in AF patients is also associated with SKNA bursts, which indicated that ANS could be the therapeutic target for rate control in AF [28]. On the other hand, in patients with ventricular arrhythmias, the average SKNA (aSKNA) was higher than the control group, and SKNA was an independent predictor of ventricular arrhythmia recurrence [29]. Also, a shared phenomenon was seen in both atrial and ventricular tachy-arrhythmias that large and sustained sympathetic nerve activities are associated with the temporal clustering of arrhythmia, indicating that neuromodulation could be an effective therapy.

4 Expectations

The noninvasive methods to assess ANS function based on ECG are far from perfect. HRV is only available only in the period of sinus rhythm, which restricts its use in patients with irregular RR interval (e.g., persistent AF, PVC). The linking of the sympathetic and parasympathetic nerves to LF and HF is also controversial. For BRS and HRT, extra stimulation is often needed to disturb the ANS to assess its function. SKNA also has the problems of signal noise and the lack of information on parasympathetic tones. These methods have to be improved for further clinical application.

The evolution of wearable devices makes it easy to have access of massive information and biological signals. However, in some cases, we have difficulties in explaining those data. There seems to be a gap between the information and the clinical events. The study and assessment of ANS function have added an extra dimension to physiological signals. The knowledge of the physiological basis may be helpful to process and explain the information. With advanced engineering and

algorithm, data can be translated into clinical benefits for patients and doctors, improving primary, secondary, and tertiary prevention of diseases.

References

1. McCorry, L.K.: Physiology of the autonomic nervous system. Am. J. Pharm. Educ. **71**, 78 (2007). https://doi.org/10.5688/aj710478
2. Herring, N., Kalla, M., Paterson, D.J.: The autonomic nervous system and cardiac arrhythmias: current concepts and emerging therapies. Nat. Rev. Cardiol. **16**, 707–726 (2019). https://doi.org/10.1038/s41569-019-0221-2
3. Goldberger, J.J., Arora, R., Buckley, U., Shivkumar, K.: Autonomic nervous system dysfunction: JACC focus seminar. J. Am. Coll. Cardiol. **73**, 1189–1206 (2019). https://doi.org/10.1016/j.jacc.2018.12.064
4. Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology: Heart rate variability: standards of measurement, physiological interpretation and clinical use. Circulation. **93**, 1043–1065 (1996). https://doi.org/10.1161/01.Cir.93.5.1043
5. Hon, E.H., Lee, S.T.: Electronic evaluation of the fetal heart rate. VIII. Patterns preceding fetal death, further observations. Am. J. Obstet. Gynecol. **87**, 814–826 (1963)
6. Ewing, D.J., Martyn, C.N., Young, R.J., Clarke, B.F.: The value of cardiovascular autonomic function tests: 10 years experience in diabetes. Diabetes Care. **8**, 491–498 (1985). https://doi.org/10.2337/diacare.8.5.491
7. Odemuyiwa, O., et al.: Comparison of the predictive characteristics of heart rate variability index and left ventricular ejection fraction for all-cause mortality, arrhythmic events and sudden death after acute myocardial infarction. Am. J. Cardiol. **68**, 434–439 (1991). https://doi.org/10.1016/0002-9149(91)90774-f
8. Jarczok, M.N., Koenig, J., Wittling, A., Fischer, J.E., Thayer, J.F.: First evaluation of an index of low vagally-mediated heart rate variability as a marker of health risks in human adults: proof of concept. J. Clin. Med. **8**, E1940 (2019). https://doi.org/10.3390/jcm8111940
9. Hayano, J., Yuda, E.: Pitfalls of assessment of autonomic function by heart rate variability. J. Physiol. Anthropol. **38**, 3 (2019). https://doi.org/10.1186/s40101-019-0193-2
10. Schmidt, G., et al.: Heart-rate turbulence after ventricular premature beats as a predictor of mortality after acute myocardial infarction. Lancet. **353**, 1390–1396 (1999). https://doi.org/10.1016/s0140-6736(98)08428-1
11. Patel, V.N., et al.: Association of Holter-derived heart rate variability parameters with the development of congestive heart failure in the cardiovascular health study. JACC Heart Fail. **5**, 423–431 (2017). https://doi.org/10.1016/j.jchf.2016.12.015
12. Makimoto, H., et al.: Reduced heart rate response after premature ventricular contraction depending on severity of atrial fibrillation symptoms - analysis on heart rate turbulence in atrial fibrillation patients. Int. J. Cardiol. Heart Vasc. **18**, 33–38 (2018). https://doi.org/10.1016/j.ijcha.2018.02.004
13. Jansen, C., et al.: Severe abnormal Heart Rate Turbulence Onset is associated with deterioration of liver cirrhosis. PLoS One. **13**, e0195631 (2018). https://doi.org/10.1371/journal.pone.0195631
14. Poliwczak, A.R., Waszczykowska, E., Dziankowska-Bartkowiak, B., Kozirog, M., Dworniak, K.: The use of heart rate turbulence and heart rate variability in the assessment of autonomic regulation and circadian rhythm in patients with systemic lupus erythematosus without apparent heart disease. Lupus. **27**, 436–444 (2018). https://doi.org/10.1177/0961203317725590
15. La Rovere, M.T., Mortara, A., Schwartz, P.J.: Baroreflex sensitivity. J. Cardiovasc. Electrophysiol. **6**, 761–774 (1995). https://doi.org/10.1111/j.1540-8167.1995.tb00452.x

16. La Rovere, M.T., Pinna, G.D., Raczak, G.: Baroreflex sensitivity: measurement and clinical implications. Ann. Noninvasive Electrocardiol. **13**, 191–207 (2008). https://doi.org/10.1111/j.1542-474X.2008.00219.x
17. Rovere, M.T.L., Bigger, J.T., Marcus, F.I., Mortara, A., Schwartz, P.J.: Baroreflex sensitivity and heart-rate variability in prediction of total cardiac mortality after myocardial infarction. Lancet. **351**, 478–484 (1998). https://doi.org/10.1016/s0140-6736(97)11144-8
18. Pinna, G.D., et al.: Applicability and clinical relevance of the transfer function method in the assessment of baroreflex sensitivity in heart failure patients. J. Am. Coll. Cardiol. **46**, 1314–1321 (2005). https://doi.org/10.1016/j.jacc.2005.06.062
19. Mortara, A., et al.: Arterial baroreflex modulation of heart rate in chronic heart failure: clinical and hemodynamic correlates and prognostic implications. Circulation. **96**, 3450–3458 (1997). https://doi.org/10.1161/01.cir.96.10.3450
20. Li, H., et al.: Baroreflex sensitivity predicts short-term outcome of postural tachycardia syndrome in children. PLoS One. **11**, e0167525 (2016). https://doi.org/10.1371/journal.pone.0167525
21. Subramanian, S.K., Sharma, V.K., Arunachalam, V., Rajendran, R., Gaur, A.: Comparison of baroreflex sensitivity and cardiac autonomic function between adolescent athlete and non-athlete boys - a cross-sectional study. Front. Physiol. **10**, 1043 (2019). https://doi.org/10.3389/fphys.2019.01043
22. Jiang, Z., et al.: Using skin sympathetic nerve activity to estimate stellate ganglion nerve activity in dogs. Heart Rhythm. **12**, 1324–1332 (2015). https://doi.org/10.1016/j.hrthm.2015.02.012
23. Doytchinova, A., et al.: Simultaneous noninvasive recording of skin sympathetic nerve activity and electrocardiogram. Heart Rhythm. **14**, 25–33 (2017). https://doi.org/10.1016/j.hrthm.2016.09.019
24. Kumar, A., et al.: Skin sympathetic nerve activity as a biomarker for syncopal episodes during a tilt table test. Heart Rhythm. (2019) (in press). https://doi.org/10.1016/j.hrthm.2019.10.008
25. Liu, X., et al.: Effects of anesthetic and sedative agents on sympathetic nerve activity. Heart Rhythm. **16**, 1875–1882 (2019). https://doi.org/10.1016/j.hrthm.2019.06.017
26. Yuan, Y., et al.: Left cervical vagal nerve stimulation reduces skin sympathetic nerve activity in patients with drug resistant epilepsy. Heart Rhythm. **14**, 1771–1778 (2017). https://doi.org/10.1016/j.hrthm.2017.07.035
27. Uradu, A., et al.: Skin sympathetic nerve activity precedes the onset and termination of paroxysmal atrial tachycardia and fibrillation. Heart Rhythm. **14**, 964–971 (2017). https://doi.org/10.1016/j.hrthm.2017.03.030
28. Kusayama, T. et al. Skin sympathetic nerve activity and ventricular rate control during atrial fibrillation. Heart Rhythm. **17**, 544–552 (2020). https://doi.org/10.1016/j.hrthm.2019.11.017
29. Zhang, P., et al.: Characterization of skin sympathetic nerve activity in patients with cardiomyopathy and ventricular arrhythmia. Heart Rhythm. **16**, 1669–1675 (2019). https://doi.org/10.1016/j.hrthm.2019.06.008

A Questionnaire Study on Artificial Intelligence and Its Effects on Individual Health and Wearable Device

Tiange Bu and Fangyuan Li

Abstract Our goal is to better understand what a professional person thinks about the current hot topics such as Artificial Intelligence (AI), wearable devices, and individual health. We designed a questionnaire to survey the professionals. The questionnaire focused on three main subjects, AI, and its effects on individual health and wearable device. In the AI-related questions, since AI can play an extremely important role in signal processing, it is widely used by respondents in the field of biological signal processing. In the individual health-related questions, the view that the local residents in the city where they live have healthy dietary habits is accepted by 43.75% of the respondents, and 21.88% of the respondents consider that the dietary habits of the local residents in the city where they live is not healthy. In the wearable devices-related questions, 59.38% of the respondents think the current wearable devices are far from meeting the need for the early disease detection and health monitoring, and only 7.81% of the respondents believe the current wearable devices are very useful. The societal impact of the AI revolution will be significant as it is beginning to affect most aspects of our lives and work, shaped our shopping and entertainment habits, as well as our employment patterns.

Keywords Artificial intelligence · Wearable devices · Individual health

1 Introduction and Motivation

With the progress of technology, artificial intelligence (AI) has emerged frequently in science and public consciousness in recent years. Artificial intelligence is a branch of computer science that attempts to understand the essence of intelligence and produce a new kind of intelligent machine that can respond in a similar way to

T. Bu (✉)
British Columbia Academy, Nanjing Foreign Language School, Nanjing, China

F. Li
School of Instrument Science and Engineering, Southeast University, Nanjing, China

C. Liu, J. Li (eds.), *Feature Engineering and Computational Intelligence in ECG Monitoring*, https://doi.org/10.1007/978-981-15-3824-7_15

human intelligence. AI techniques are poised to influence nearly every aspect of the human condition. With recent progress in digitized data acquisition, machine learning, and computing infrastructure, AI applications are expanding into areas that were previously thought to be only the province of human experts [2]. Over the past decade, several machine-learning techniques have been used in fields of medicine, biomedical signal processing, and chemistry. A healthy lifestyle is a habit of behavior that is good for your health, which concludes scientific diet, reasonable work and rest, and appropriate exercise. Individual lifestyle has an essential effect on the occurrence and development of chronic diseases, thus plays a main role in the individual health. The core of healthy lifestyle management is to develop good habits. For a long time, people made a series of health plans by themselves, and it was difficult to monitor their health status in real time. With the rise of mobile Internet and wearable devices, healthy lifestyle management methods have also changed. A series of mobile Internet health management tools provide a lot of convenience for people, making the development of lifestyle more interesting and more motivated. With the development of mobile Internet, technological progress, and the launch of high-performance and low-power processing chips, some wearable devices have gone from conceptualization to commercialization, and new wearable devices are coming out. Wearables as medical technologies are becoming an integral part of personal analytics, measuring physical status, recording physiological parameters, or informing schedule for medication. These continuously evolving technology platforms do not only promise to help people pursue a healthier life style but also provide continuous medical data for actively tracking metabolic status, diagnosis, and treatment [4].

Our goal is to better understand what a professional person thinks about these current hot topics.

2 Method

The questionnaire focused on three main subjects: AI, its effects on individual health, and wearable device. There are ten questions in total: questions 1 to 3 are multiple choices; question 4 is indicating possibilities; questions 5 to 10 are multiple choices. In the first subject Artificial Intelligence, there are three questions; then in the second subject, there are three questions. In the last subject, there are four questions. At the beginning of the questionnaire, we designed a form to ask the respondents to fill in their details of work and research, and also their nationalities.

2.1 Basic Information

Name: _____ Gender: □ Female/□ Male
Nationality: _____ Age: _____

Research topic: (based on the conference classification for the topics of interest, please tick one opinion)

□ Medical Imaging Technology and Application
□ Biomedical Signal Processing and Medical Information
□ Biomechanics and Biomechanical Engineering
□ Molecular Biology
□ Chemistry, Pharmacology, and Toxicology
□ Other Topics

2.2 Questionnaire Survey

Part I: AI

AI is growing so fast currently. It has been used widely in many areas, such as autonomous vehicles, playing games, finance, economics, art, and of course, the healthcare. It is undeniable that AI has a profound impact, not only on our research but also on the people's daily life.

1. Assume the participants of the questionnaire have the quantitative assessment point on this question. Concerning the research work you engaged in, how often/how much do you use the AI-related techniques?
   Level: □ 0% (Never) □ 25% □ 50% □ 75% □ 100% (Very often)
2. Assume the participants of the questionnaire have the quantitative assessment point on this question. In your opinion, how often/how much the AI-related techniques are applied in the people's daily life in the city where you live?
   Level: □ 0% (Never) □ 25% □ 50% □ 75% □ 100% (Very often)
3. Due to the rapid rise of AI and the boundless possibilities existing, many people talk about the topic of "AI will replace the doctor one day." How long do you think the AI-related techniques can greatly surpasses the performance of professional human in your profession?
   □ Within 5 years □ 5–10 years □ 10–20 years □ 20–50 years □ Never
4. Assume the participants of the questionnaire have the quantitative assessment point on this question. How positive or negative would be the overall impact of AI on humanity, in the long run? Please indicate a probability for each option. (The sum should be equal to 100%.)

Impact	***Probability***
Extremely good	_____
On balance good	_____
More or less neutral	_____
On balance bad	_____
Extremely bad (existential catastrophe)	_____
	Total: 100%

Part II: Individual Health

Chronic, noncommunicable diseases, such as cardiovascular diseases, chronic respiratory disease, and cancer, are the leading causes of both death and the burden of disease in the world. It reported that individual lifestyle has an important effect on the occurrence and development of chronic diseases, and thus plays an important role on the individual health.

5. Assume the participants of the questionnaire have the quantitative assessment point on this question. For the dietary habits of the residents in the city where you live, do you think it is healthy?
   Level: □ 0% (Totally unhealthy) □ 25% □ 50% □ 75% □ 100% (Totally healthy)
6. Assume the participants of the questionnaire have the quantitative assessment point on this question. For the sports and fitness of the residents in the city where you live, how often do you think in the average level?
   Level: □ 0% (Never) □ 25% □ 50% □ 75% □ 100% (Very often)
7. Assume the participants of the questionnaire have the quantitative assessment point on this question. In your opinion, how important the individual lifestyle (including dietary habit, sport, fitness, etc.) performs on the individual health?
   Level: □ 0% (Not at all important) □ 25% □ 50% □ 75% □ 100% (Very important)

Part III: Wearable Device

Many wearable devices have been developed for the individual health monitoring currently, such as the daily ECG, blood pressure, exercise, sleep monitoring, generating benefit for the early disease detection, and health management, especially when working with the AI-related techniques. However, for the useful clinical application, challenges exist. And here come the questions:

8. Assume the participants of the questionnaire have the quantitative assessment point on this question. How do you think the current wearable devices/techniques can meet the need for the early disease detection and health monitoring?
   Level: □ 0% (Never useful) □ 25% □ 50% □ 75% □ 100% (Very useful)
9. How long do you think the wearable devices/techniques can perform very well for the individual health monitoring, and thus significantly improve the individual health management?
   □ Within 5 years □ 5–10 years □ 10–20 years □ 20–50 years □ Never
10. Wearable techniques can generate lots of individual data. Opening and sharing the wearable data (with hiding the personal information) can significantly benefit the research work, while it has the risk to reveal personal privacy. Assume the participants of the questionnaire have the quantitative assessment point on this question. How would you like to support the open source data?
    Level: □ 0% (Never support) □ 25% □ 50% □ 75% □ 100% (Fully support)

3 Results

The questionnaire survey has 65 respondents, consisting of 45 males and 20 females. They come from six countries, including China, the United States, South Korea, the United Kingdom, Turkey, and Israel. The respondents were mainly from the fields of medical imaging technology and application, biomedical signal processing and medical information, biomechanics and biomechanical engineering, molecular biology, chemistry, pharmacology, and toxicology (Fig. 1).

3.1 Results on AI-Related Questions

Artificial Intelligence is used frequently by 36.92% of the respondents, and only 3.08% of the respondents never use AI technology, suggesting that AI has been widely applied in the fields of medical imaging, biological signal processing, biomechanics, molecular biology, chemistry, and pharmacy. Only the AI-related techniques can greatly surpass the performance of human in their profession within 5 years, and 35.38% of the respondents believe that AI will never surpass the performance of human in their profession. It shows that although a number of respondents have used AI technology in scientific research, most of them believe that AI cannot surpass human performance in the short term. Only 1.54% of the respondents consider that the AI-related techniques are never applied in the people's daily life in the city where they live, suggesting that AI is gradually changing people's daily life in the city where respondents live. The overall impact of AI on humanity would be positive, and this view is agreed by 73.84% of the respondents. It follows that most respondents deem the impact of AI positive in the long run.

The respondents were mainly from the fields of medical imaging technology and application, biomedical signal processing and medical information, biomechanics and biomechanical engineering, molecular biology, chemistry, pharmacology, and toxicology. Figure 2 shows that the frequencies of AI used in various research topics. AI-related techniques are used more frequently in the fields of medical imaging and biological signal processing, especially in the field of biological signal processing. Medical imaging examination relies mainly on the manual reading of doctors, which requests a high standard of the doctor's professional skills and clinical experience. At present, AI techniques are used to extract useful information from massive medical image data for assisting clinical diagnosis and clinical decision. The promise of artificial intelligence (AI) and machine learning in cardiology is to provide a set of tools to augment and extend the effectiveness of the cardiologist [1]. Similarly, since AI can play an extremely important role in signal processing, it is widely used by respondents in the field of biological signal processing.

From Fig. 3, we can conclude that most respondents agree that AI is difficult to surpass human performance in their fields. Although AI technology is frequently used in the fields of medical imaging and biological signal processing, the largest

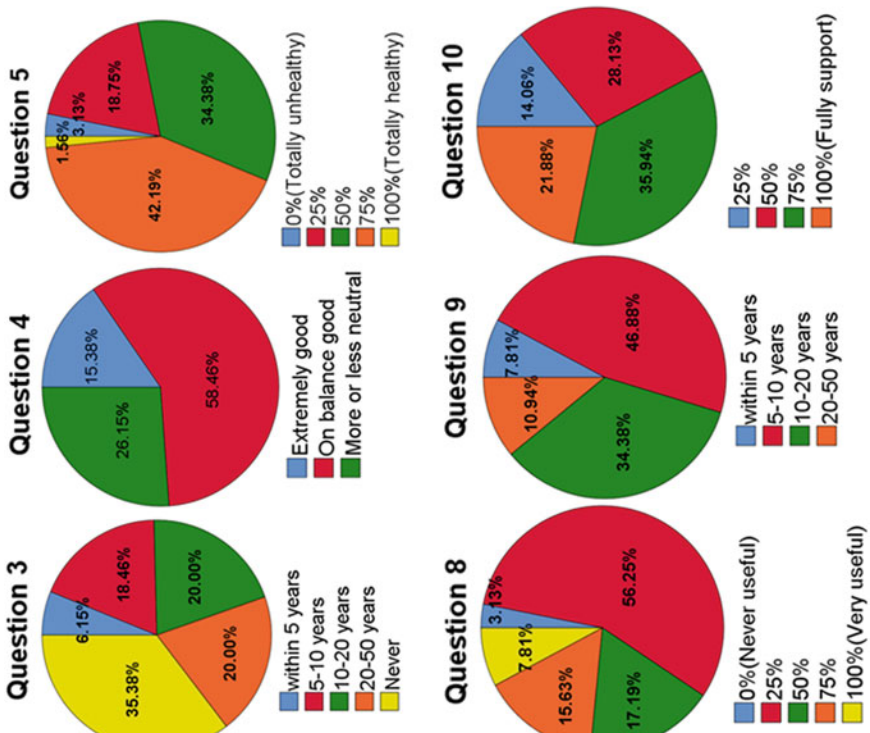

Fig. 1 The pie chart of statistical results of questionnaire survey

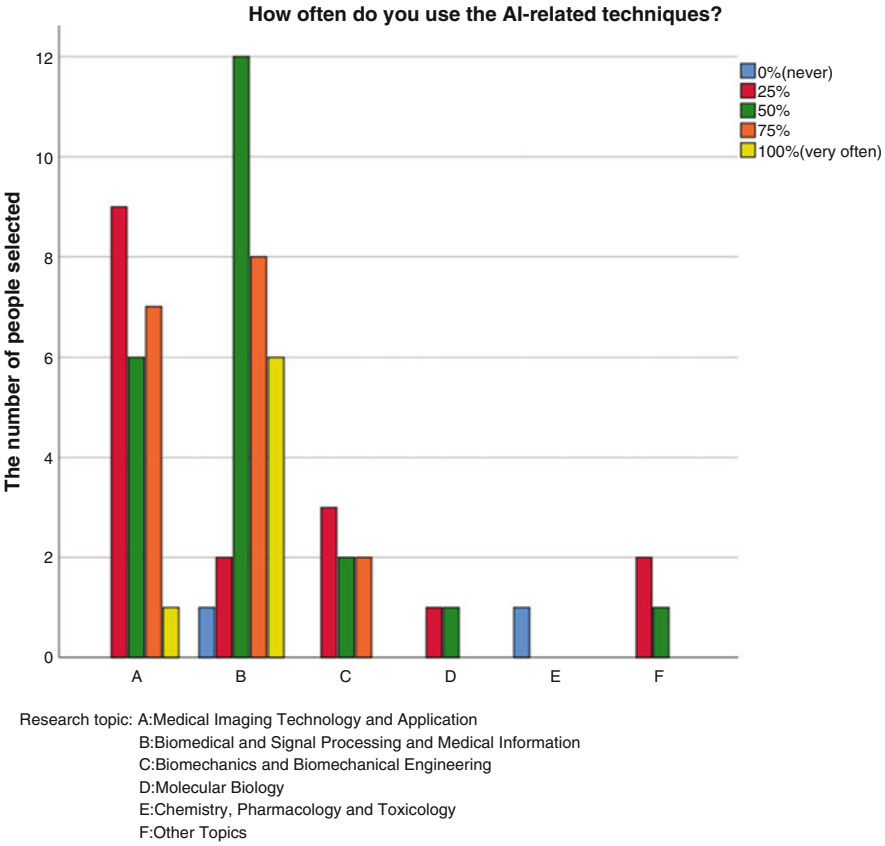

Fig. 2 The bar chart of the frequencies of AI used in various research topics

number of respondents believed that AI could never surpass human performance in their fields. It can be seen that Although AI promises to revolutionize many research fields, a number of technical challenges lie ahead. The availability of large amounts of high-quality training data on which AI-related techniques rely heavily is the first challenge [2]. Although several high-performing machine-learning models can achieve better-than-human performance, results generated by them are difficult to interpret by unassisted human [2].

3.2 Results on Individual Health-Related Questions

Dietary habits and physical activity are significant indicators to evaluate whether an individual lifestyle is healthy; 85.94% of the respondents agree that a healthy lifestyle plays a vital role in individual health. The respondents of the questionnaire

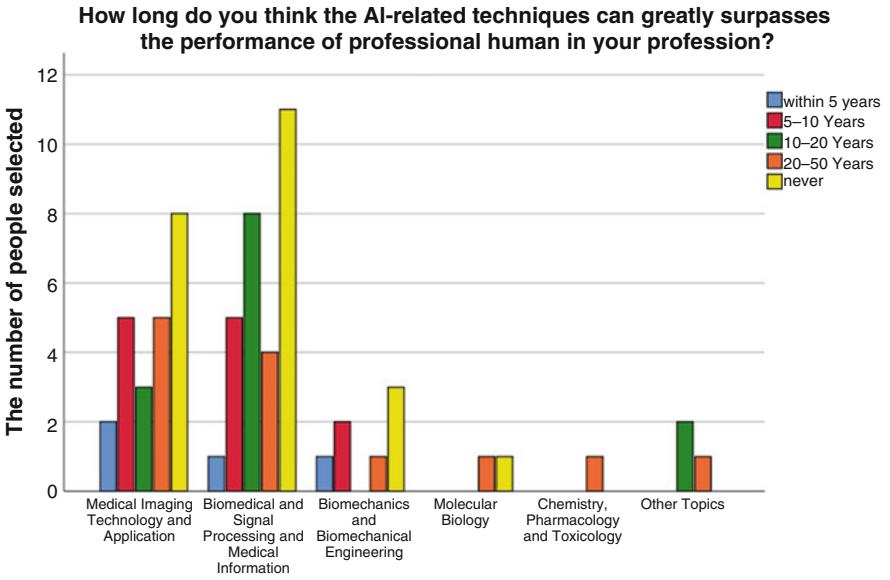

Fig. 3 The bar chart of the views on when AI will surpass human performance in various research topics

survey come from six countries, including China, the United States, South Korea, the United Kingdom, Turkey, and Israel. The view that the local residents in the city where they live have healthy dietary habits is accepted by 43.75% of the respondents, and 21.88% of the respondents consider that the dietary habits of the local residents in the city where they live is not healthy. Physical inactivity is one of the leading modifiable risk factors for global mortality, with an estimated 20–30% increased risk of death compared with those who are physically active [3]. Abundant scientific evidence has demonstrated that physically active people of all age groups and ethnicities have higher levels of cardiorespiratory fitness and health, and a lower risk for developing several chronic medical illnesses, compared with those who are physically inactive [3]. For the frequency of sports and fitness of the local residents in the city where they live, 34.38% of the respondents believe it is often, and 23.44% of the respondents hold the opposite view.

3.3 Results on Wearable Devices-Related Questions

With the growing consumer desire for health awareness and the development of wearable technologies, the past decade has seen the emergence of several wearable devices, including several that have been widely adopted by both physicians and consumers. Wearable devices are increasingly helping people to better monitor their health involving acquisition of data on daily activities, sport performance, and health

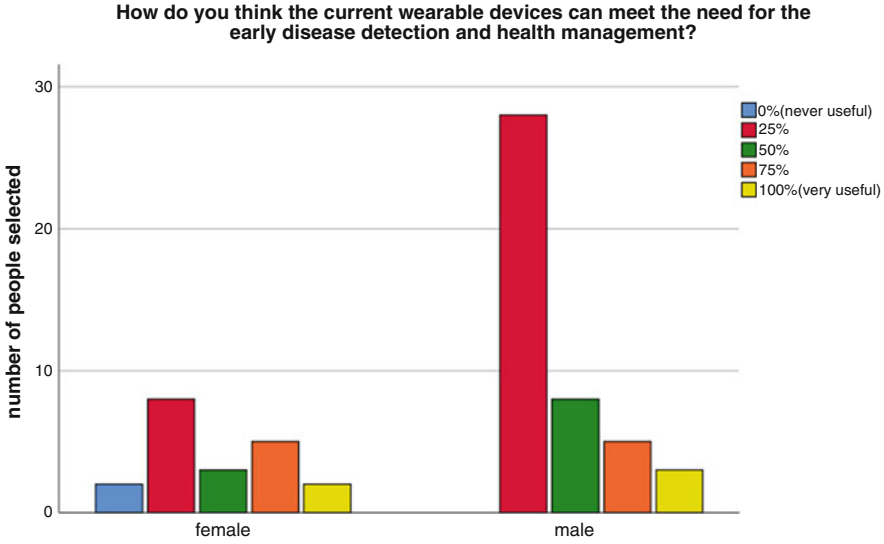

Fig. 4 The bar chart of the views on whether wearable devices can meet the demand of different gender

status. At present, wearable devices include smart watches, smart wristband, and electronic textiles [4]. In all, 59.38% of the respondents think the current wearable devices are far from meeting the need for the early disease detection and health monitoring, and only 7.81% of the respondents believe that the current wearable devices are very useful. It shows that most people are not satisfied with the performance of current wearable devices. From Fig. 4, we can conclude that both male and female are not satisfied with the performance of current wearable devices, especially males. The cause of this result is that males have a greater need than females for the ability to acquire sport performance using wearable devices in daily life. Figure 5 shows that the views on whether wearable devices can meet the demand of different age group. Respondents over 50 are more dissatisfied with the performance of wearable devices, reflecting the elderly have higher requirements on wearable devices and need better health monitoring. Due to the fact that elderly are more likely to have health problems than youngsters, wearable devices for the elderly need better performance in early disease detection and health management.

The key to development of wearable devices can be attributed to several factors, such as sensor signal acquisition, processing, and artificial intelligence algorithms. The next few years present a set of challenges in wearable devices, and to overtake them, an improvement of these technologies should be made. The view that wearable devices can perform very well for the individual health monitoring and management in 5–10 years is agreed by 46.88% of the respondents, and only 7.81% of the respondents think wearable devices can perform very well for the individual health monitoring and management within 5 years. It suggests that most respondents are confident in the future development of wearable devices in the long run, but

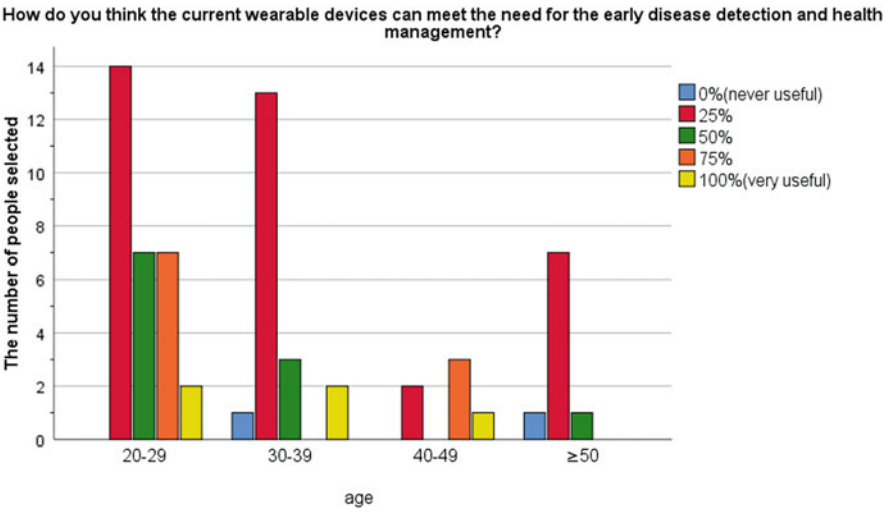

Fig. 5 The bar chart of the views on whether wearable devices can meet the demand of different age group

wearable devices can hardly meet the needs for the early disease detection and health monitoring within 5 years. The term big data refers to extremely large datasets that cannot be analyzed, searched, interpreted, or stored using traditional data-processing methods. Big data include data from mobile phone applications, wearable technology, and precision medicine platforms [5]. Wearable techniques can generate lots of individual data. Opening and sharing the wearable data (with hiding the personal information) can significantly benefit the research work, while it has the risk to reveal personal privacy. For example, currently some wearable device service providers limit their user's access to the collected and stored data. These service providers charge their users' fees to access their row data, which is also acquired by third-party companies. The open-source data is supported by 64.07% of the respondents, and only 14.06% of the respondents do not support the open source. Although data safety is a potential risk, open-source data enables AI algorithms more effective in early disease detection and health monitoring.

4 Discussion

With the current popularization of the Internet, the universal existence of sensors, the emergence of big data, development of e-commerce, rise of the information community, and the interconnection and fusion of data, and knowledge with society, physical space, and cyberspace, the information environment for AI development has been changed profoundly [6]. The societal impact of the AI revolution will be significant as it is beginning to affect most aspects of our lives and work, shaped our

shopping and entertainment habits, as well as our employment patterns. At present, AI-related technologies are applied in the fields of medical imaging, biological signal processing, biomechanics, molecular biology, chemistry, and pharmacy more or less. In the above fields, AI-related techniques are used more frequently in the fields of medical imaging and biological signal processing, especially in the field of biological signal processing. AI can also be used to process large amounts of data from images and signals (i.e., information about the attributes of a particular physical phenomenon). Data produced by motion and sound are common examples of signals. Steps in image and signal processing algorithms typically include signal feature analysis and data classification using tools such as artificial neural networks (ANNs) [7]. Most of the medical data come from medical imaging; the medical imaging data are still growing year by year, but the growth rate and work efficiency of imaging doctors are not enough to cope with this growth trend, which will bring great pressure to doctors. Medical imaging has become the most popular direction in the field of artificial intelligence. Artificial intelligence (AI) is already widely employed in various medical roles, and ongoing technological advances are encouraging more widespread use of AI in imaging [8]. AI techniques are used to extract useful information from massive medical image data for assisting clinical diagnosis and clinical decision. Moreover, AI methods have been gradually introduced into biological signal processing in recent years. Although AI-related techniques are widely used in many fields of scientific research, the majority of scholars consider that AI technology is difficult to surpass human performance in their profession, at least in the short term. Algorithm, computing power, and data are considered the three core elements of artificial intelligence. The growth of data volume, the improvement of computing power, and the optimization of deep learning algorithm are the current problems that we need to break through. Substantial uncertainty will be brought by AI; most people, however, believe the impact of AI positive in the long run. Individual health, as demonstrated by scientific practice, depends on heredity, environment, and lifestyle, the most important of which is the individual lifestyle. With the improvement of living standards, people worldwide are aware of the importance of individual health. Healthy dietary habits and adequate exercise are pursued by people all over the world. The awareness of a healthy individual lifestyle is the driving force behind wearable technologies involving acquisition of data on daily activities, sport performance, and health status. At the same time, the demand for wearable devices has increased. The consumer-directed wearable devices market is rapidly growing, which includes devices that can be worn from head to toe. At present, the main wearable products on the market have different forms, such as smart watches, smart bracelets, health wear, and so on. AI analytics support the practice of precision medicine, especially in the difficult setting of chronic diseases characterized by multiorgan involvement, erratic acute events, and long illness progression latencies [9]. However, few of people think the current wearable devices can meet the need for the early disease detection and health monitoring. Currently, the development of wearable devices is in the initial stage, and there are still many problems that need to be solved, among which the most important is the diversity of functions, accuracy of results, and security of data. In order to accurately measure

physiological indicators, comprehensive application of multiple sensors and cross-validation of results are required by wearable devices. The main target of wearable devices is to integrate several biosensors, intelligent processing, and alerts to detect early disease and monitor health while interacting with health providers. With the development of artificial intelligence, 5G network, and sensor technology, wearable devices can perform very well for the individual health monitoring, and thus significantly improve the individual health management in the long run. Moreover, the rapid development of machine-learning techniques has promised to bring forth even more useful applications from big data to disease diagnosis and health monitoring. However, the open-source data is still controversial: while many people support open data, some hope conservative attitudes. The concept of privacy is notoriously difficult to define. One currently prominent view connects privacy to context. There are contextual rules about how information can flow that depend on the actors involved, the process by which information is accessed, the frequency of the access, and the purpose of that access [10]. Therefore, users' privacy data should be properly handled in the future.

References

1. Johnson, K.W., Torres Soto, J., Glicksberg, B.S., Shameer, K., Miotto, R., Ali, M., Ashley, E., Dudley, J.T.: Artificial intelligence in cardiology. J. Am. Coll. Cardiol. **71**(23), 2668–2679 (2018)
2. Yu, K.-H., Beam, A.L., Kohane, I.S.: Artificial intelligence in healthcare. Nat. Biomed. Eng. **2** (10), 719 (2018)
3. Fletcher, G.F., Landolfo, C., Niebauer, J., Ozemek, C., Arena, R., Lavie, C.J., et al.: Promoting physical activity and exercise. J. Am. Coll. Cardiol. **72**(14), 1622–1639 (2018)
4. Yetisen, A.K., Martinez-Hurtado, J.L., Ünal, B., Khademhosseini, A., Butt, H.: Wearables in medicine. Adv. Mater. **30**(33), 1706910 (2018). https://doi.org/10.1002/adma.201706910
5. Krittanawong, C., Zhang, H., Wang, Z., et al.: Artificial intelligence in precision cardiovascular medicine. J. Am. Coll. Cardiol. **69**(21), 2657–2664 (2017)
6. Li, B.-h., Hou, B.-c., Yu, W.-t., Lu, X.-b., Yang, C.-w.: Applications of artificial intelligence in intelligent manufacturing: a review. Front. Inf. Technol. Electron. Eng. **18**(1), 86 (2017)
7. Wahl, B., Cossy-Gantner, A., Germann, S., Schwalbe, N.R.: Artificial intelligence (AI) and global health: how can AI contribute to health in resource-poor settings? BMJ Glob. Health. **3** (4), e000798 (2018)
8. Fazal, M.I., Patel, M.E., Tye, J., Gupta, Y.: The past, present and future role of artificial intelligence in imaging. Eur. J. Radiol. **105**, 246 (2018)
9. Miller, D.D., Brown, E.W.: Artificial intelligence in medical practice: the question to the answer. Am. J. Med. **131**, 129 (2018)
10. Price II, W.N., Cohen, I.G.: Privacy in the age of medical big data. Nat. Med. **25**(1), 37 (2019)