sensors

MDPI

Review

Computational Diagnostic Techniques for Electrocardiogram Signal Analysis

Liping Xie *, Zilong Li, Yihan Zhou, Yiliu He and Jiaxin Zhu

College of Medicine and Biological Information Engineering, Northeastern University, Shenyang 110169, China; longzilipro@gmail.com (Z.L.); 20185423@stu.neu.edu.cn (Y.Z.); 20185457@stu.neu.edu.cn (Y.H.); 20175563@stu.neu.edu.cn (J.Z.)

* Correspondence: xielp@bmie.neu.edu.cn

Received: 8 September 2020; Accepted: 4 November 2020; Published: 5 November 2020

Abstract: Cardiovascular diseases (CVDs), including asymptomatic myocardial ischemia, angina, myocardial infarction, and ischemic heart failure, are the leading cause of death globally. Early detection and treatment of CVDs significantly contribute to the prevention or delay of cardiovascular death. Electrocardiogram (ECG) records the electrical impulses generated by heart muscles, which reflect regular or irregular beating activity. Computer-aided techniques provide fast and accurate tools to identify CVDs using a patient's ECG signal, which have achieved great success in recent years. Latest computational diagnostic techniques based on ECG signals for estimating CVDs conditions are summarized here. The procedure of ECG signals analysis is discussed in several subsections, including data preprocessing, feature engineering, classification, and application. In particular, the End-to-End models integrate feature extraction and classification into learning algorithms, which not only greatly simplifies the process of data analysis, but also shows excellent accuracy and robustness. Portable devices enable users to monitor their cardiovascular status at any time, bringing new scenarios as well as challenges to the application of ECG algorithms. Computational diagnostic techniques for ECG signal analysis show great potential for helping health care professionals, and their application in daily life benefits both patients and sub-healthy people.

Keywords: electrocardiogram; classification; feature engineering; deep learning; machine learning

1. Introduction

Cardiovascular disease (CVD) is the leading cause of death in the world, according to the statistics of the World Health Organization. More than 30% of global deaths are caused by cardiovascular disease each year, and more than 130 million adults are estimated to be suffered from cardiovascular disease by 2035 [1]. As a technique of recording electrophysiological signals of heart, electrocardiogram (ECG) with the advantages of non-invasiveness and real time, has been widely used in medical care, such as heartbeat recognition, blood pressure detection, and disease detection [2,3]. Since the first discovery of ECG by Augustus Desire Waller in 1887 [4], ECG signals have been widely used in the diagnosis of heart diseases, such as arrhythmia and myocardial infarction. Computer-assisted medical diagnoses based on ECG signals can give professional suggestions or even make a judgment directly by looking for characteristic patterns on the ECG signals. According to the number of leads, common ECG acquisition device can be divided into 1-lead ECG, 3-lead ECG, 6-lead ECG and 12-lead ECG [5]. 12-lead ECG is the most commonly used type in clinical practice [6], because of its ability to record the potential changes of 12 sets of electrode patches attached to the body in standardized locations at the same time. Compared with other types of ECG acquisition equipment, 12-lead ECG possesses more detailed information about heart activities and is often used in professional diagnosis and treatment in hospitals. With the advancement of software and hardware, single-lead ECGs and photoplethysmography (PPG) sensors which are more readily available, inexpensive, convenient, and easily integrated into portable devices have entered people's daily life. All these characteristics make it suitable for long-term monitoring and expand the application of ECG.

Analysis and interpretation of the ECG often perform by professional doctors [7], which largely depends on the doctor's training, certifications experience and knowledge [8]. However, even experts cannot get enough information from the ECG signals [9]. With the advancement of algorithm and physical hardware technology, automated diagnostic systems become to play an increasingly important role in the diagnosis of heart disease, transitioning from selecting potentially effective lesion features for doctors to independent decision making. ECG features are unique information extracted from ECG signals and are used to represent the state of the heart. In addition to morphological features that can be observed, features such as wavelet features and statistical features have also been proven to be effective in diagnosis [10–12]. In recent years, the fast development of machine learning methods has brought new opportunities to medical signal analysis [13], making ECG analysis more intelligent and efficient [14]. At present, wearable medical devices in healthcare can even replace 12-lead ECG acquisition devices [15,16], which can detect some common cardiovascular diseases, such as atrial fibrillation, arrhythmia, and stress.

There are many reviews on ECG in recent years [3], however, most of them focus on single processing techniques, such as ECG pre-processing, feature engineering techniques. Lyon et al. analyzed several commonly used classifier algorithms in ECG analysis, including support vector machine (SVM), random forest algorithm, Bayesian network, and neural network networks [17]. However, there are few systematic reviews on the application of deep learning in ECG analysis. With the enhancement of computing power and the development of cloud computing technology in recent years, machine learning and deep learning have shown great advantages in ECG analysis. More studies choose to use dimensionality reduction methods to extract effective features from a large feature sets, instead of only studying pre-selected specific features. With the development of End-to-End models represented by artificial neural networks, ECG analysis no longer relies on feature engineering. At present, end-to-end models featured by deep neural network's structure have aroused more and more researchers' attention. All of the parameters of the end-to-end models are trained jointly, rather than step by step. Without using any hand-crafted techniques, ECG analysis based on end-to-end model has great advantages in accuracy and robustness. This article summarizes the latest computational diagnostic techniques based on ECG signals for estimating CVD conditions (Figure 1). The procedure of ECG signals analysis based on machine learning is discussed from data preprocessing, feature extraction and selection, classification, and application. Notably, End-to-End models based on deep learning algorithms for ECG analysis have been summarized, which enable the analysis process no longer to require a feature extraction with hand-crafted techniques. We also discuss the development trends and challenges of computational diagnostic techniques for ECG analysis, which demonstrates great potential of ECG-assisted analysis in health diagnosis.

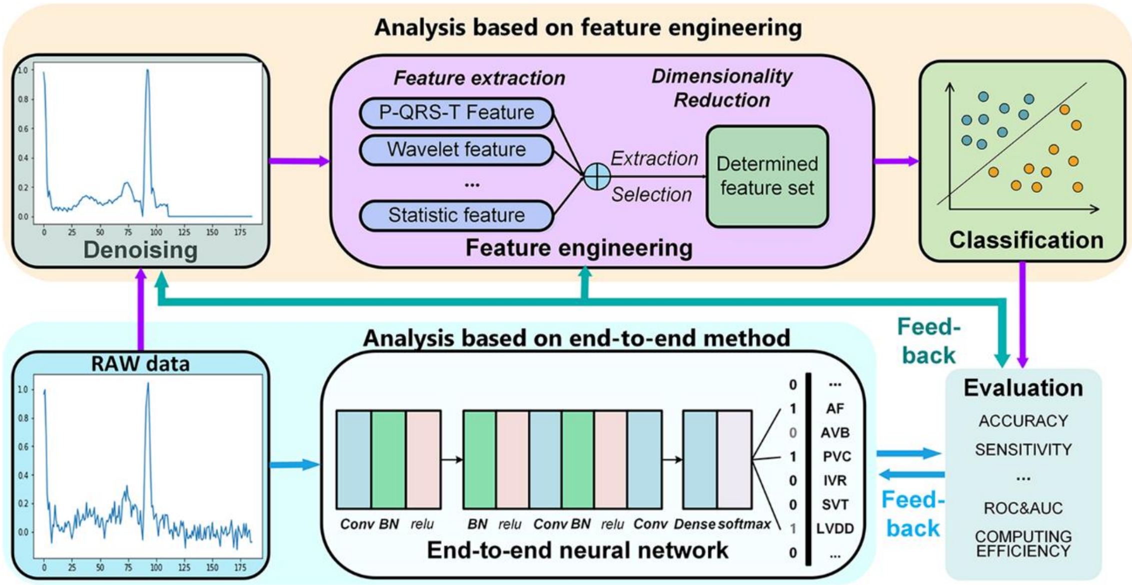

Figure 1. Process of computational diagnostic techniques for electrocardiogram signals.

2. Data Preprocessing

2.1. Noise in Electrocardiogram (ECG)

Normal ECG signals are time-varying signals with a small amplitude ranging from 10 μV to 5 mV. Their typical value is 1 mV and their frequencies range from 0.05–100 Hz, mainly concentrated in the 0.05~35 Hz range. Most ECG analysis systems require relatively noiseless ECG signals to achieve accurate and reliable CVDs diagnosis [18]. However, in practice ECG signals are often disturbed by various noises and artifacts, including baseline drift, electrode contact noise, power-line interference (PLI), and electromyographic (EMG) noise, which lead to ECG distortion and affect feature extraction [19–21].

Patient movements, poor electrode contact, and changes in electrode skin impedance cause baseline wander and abrupt drift noises. During Holter monitoring, ECG monitoring in moving ambulance or exercise, baseline wander plays a significant role in noises [22]. The frequency range of the motion noise is usually 1–10 Hz. The amplitude of baseline wander caused by respiration varies by ~15% of peak-to-peak ECG amplitude at frequencies ranging from 0.15 Hz to 0.3 Hz [19]. It is noted that the severe baseline wanders or motion artifacts can distort the ST-segment and other low-frequency components of ECG signals. The ST segment distortion may lead to the wrong diagnosis of myocardial infarction, Brugada syndrome, and other ST-segment related abnormalities.

Power-line interference (PLI) (50/60 Hz) is mainly contributed by inductive and capacitive couplings of ubiquitous power lines in the ECG signal acquisition circuitry [20]. The lower frequency noise components of the PLI are mixed with the frequency content of the ECG signal. The severely structured noises can distort the morphological features such as amplitude, duration, and shape of low-amplitude local waves of the ECG signal. In particular, the P-wave distortions can lead to the wrong diagnosis of atrial arrhythmias such as atrial enlargement and fibrillation [23]. Proper shielding, avoiding loose connection in wires and appropriate placement of electrodes can minimize the PLI.

EMG noise is contributed by the electrical activity of muscles during periods of contraction. The frequency distribution of EMG noise is basically within 0 Hz to 500 Hz and mainly concentrates in the range from 50–150 Hz. EMG noise with a high frequency above 100 Hz can be removed by a low-pass filter with an appropriate cut-off frequency. Previous studies demonstrated that EMG artifacts significantly altered the shapes of local waves of the ECG signal because the frequency of the noise was considerably superimposed with that of the ECG signals in the range of 0.05–100 Hz.