International Journal of Cardiology 337 (2021) 71–78

Contents lists available at ScienceDirect

International Journal of Cardiology

journal homepage: www.elsevier.com/locate/ijcard

The application of deep learning in electrocardiogram: Where we came from and where we should go?

Jin-Yu Sun, Hui Shen, Qiang Qu, Wei Sun *, Xiang-Qing Kong *

Department of Cardiology, The First Affiliated Hospital of Nanjing Medical University, Nanjing 210000, China.

A R T I C L E I N F O

Article history:
Received 6 April 2021
Received in revised form 22 April 2021
Accepted 10 May 2021
Available online 14 May 2021

Keywords:
Deep learning
Artificial intelligence
Machine learning
Electrocardiogram
Diagnosis

A B S T R A C T

Electrocardiogram (ECG) is a commonly-used, non-invasive examination recording cardiac voltage versus time traces over a period. Deep learning technology, a robust artificial intelligence algorithm, can imitate the data processing patterns of the human brain, and it has experienced remarkable success in disease screening, diagnosis, and prediction. Compared with traditional machine learning, deep learning algorithms possess more powerful learning capabilities and can automatically extract features without extensive data pre-processing or hand-crafted feature extraction, which makes it a suitable tool to analyze complex structures of high-dimensional data. With the advances in computing power and digitized data availability, deep learning provides us an opportunity to improve ECG data interpretation with higher efficacy and accuracy and, more importantly, expand the original functions of ECG. The application of deep learning has led us to stand at the edge of ECG innovation and will potentially change the current clinical monitoring and management strategies. In this review, we introduce deep learning technology and summarize its advantages compared with traditional machine learning algorithms. Moreover, we provide an overview on the current application of deep learning in ECGs, with a focus on arrhythmia (especially atrial fibrillation during normal sinus rhythm), cardiac dysfunction, electrolyte imbalance, and sleep apnea. Last but not least, we discuss the current challenges and prospect directions for the following studies.

1. Introduction

Electrocardiogram (ECG) is an inexpensive, commonly used, and non-invasive diagnostic examination that illustrates cardiac voltage versus time traces from surface recordings over a period [1]. Since the development of ECG investigators in 1895 [2], much effort has been paid to expand or strengthen the diagnostic potential of ECG. To date, ECG has become a mature method assisting the diagnosis of multiple cardiovascular abnormalities, and it is regularly performed on patients for both cardiac and non-cardiac reasons [3]. Despite the rapid increase in digital ECG data, appropriate and efficient interpretation remains challenging since analysis by electrophysiologists or cardiologists is usually time-consuming and highly dependent on individual expertise.

Among the many artificial intelligence algorithms, deep learning is a powerful tool, which imitates the data processing patterns of the human brain and creates models for decision making [4]. Over the last 5 years, deep learning has shown remarkable success in medical applications, including disease screening [4], diagnosis [5], and prediction [6]. For digital ECG data, deep learning algorithms could detect subtle alterations in ECGs related to cardiac structural or functional abnormalities [7]. Studies have demonstrated that deep learning application provides substantial improvements for ECG data interpretation with high efficacy and accuracy.

Rapid algorithmic and computational advances allow us to revisit the role of deep learning in ECG analysis. Previous applications of artificial intelligence in ECG were focused on single aspects of ECG processing, while deep learning has shown great potential to strengthen and expand the original functions of ECGs, such as predicting left ventricular function [4,8], sleep apnea [9], and mortality [10] (Fig. 1). In this review, we introduce deep learning and compare it with traditional machine learning algorithms. Moreover, we provide an overview of the current application of deep learning in ECGs, focusing on arrhythmia, cardiac dysfunction, electrolyte imbalance, sleep apnea, and so forth [11–13] (Table 1). Last but not least, we discuss the current challenges and propose future directions.

2. Deep learning

Deep learning, also known as deep structured learning, is a subfield of machine learning algorithms that can extract raw data representations without information loss and perform complex tasks. Neural networks, which make up the backbone of deep learning, mimic the

* Corresponding authors.
E-mail addresses: weisun7919@njmu.edu.cn (W. Sun), kongxq@njmu.edu.cn (X.-Q. Kong).

https://doi.org/10.1016/j.ijcard.2021.05.017
0167-5273/

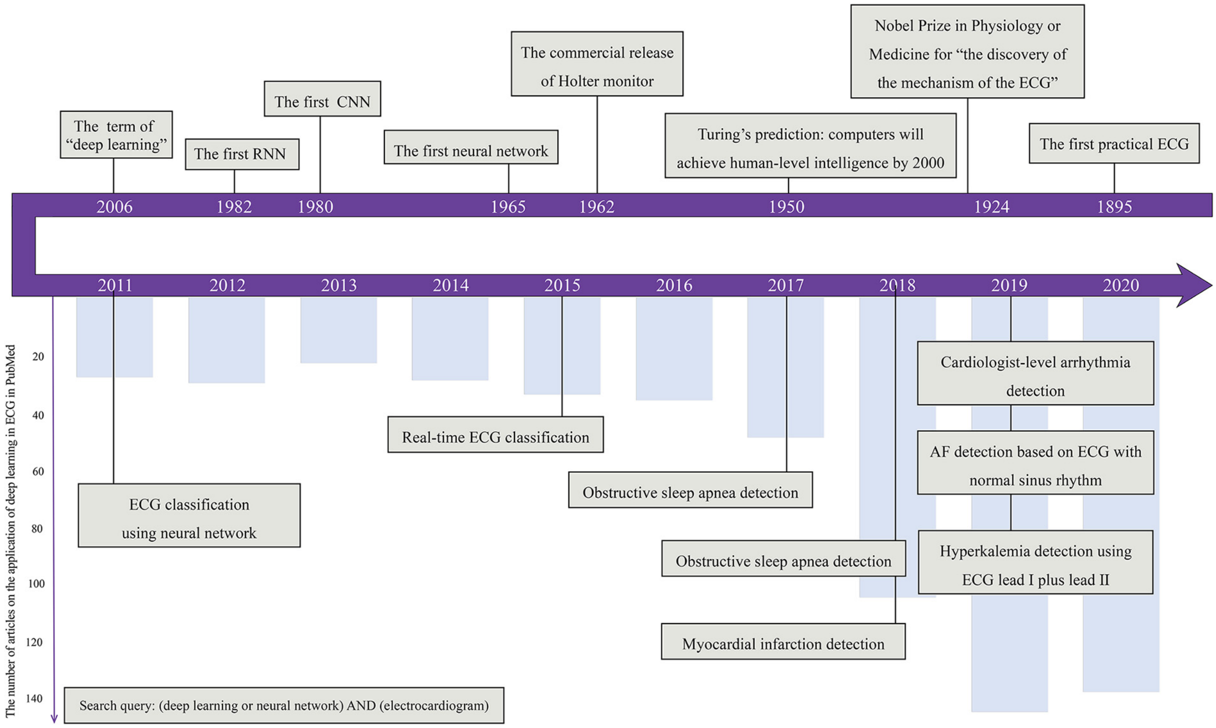

Fig. 1. Timeline showing the landmarks about deep learning and its application in the electrocardiogram. The timeline shows landmarks in the development of deep learning algorithms and their applications in ECGs. The bar graph illustrates the number of peer-reviewed articles in this area published per year from 2011 to 2020, as indexed by PubMed (https://pubmed.ncbi.nlm.nih.gov/). ECG: electrocardiogram; CNN: convolutional neural network; RNN: recurrent neural network; AF: atrial fibrillation.

information processing process and the distributed communication nodes architecture of the human brain (Fig. 2A). The word 'deep' in the term 'deep learning' indicates the depth of layers in a network, meaning that a neural network with ≥3 layers is considered a deep learning algorithm. Currently, multiple deep learning architectures (such as convolutional neural networks [CNN], deep neural networks

Table 1
Current application of deep learning in electrocardiogram.

Disease	ECG type	AUC	ACC (%)	SEN (%)	SPE (%)	Ref.
Atrial fibrillation	Standard 10-s, 12-lead ECGs with normal sinus rhythm	0.87	79.4	79.0	79.5	21
Atrial fibrillation	Segmented ECGs form Holter (30-s duration) with normal sinus rhythm	–	98.7	98.6	98.7	38
The 12 most common rhythm classes	Single-lead ECGs collected by a patch-based ambulatory ECG monitoring device	0.97	–	–	–	39
23 types of rhythm classes	Standard 10-s, 12-lead ECGs	0.96–0.99	–	0.88–1.00	0.39–0.99	41
LV dysfunction (EF ≤ 35%)	Standard 10-s, 12-lead ECGs	0.93	85.7	86.3	85.7	4
LV dysfunction (EF ≤ 35%)	Standard 10-s, 12-lead ECGs	0.92	86.5	86.8	82.5	22
LV dysfunction (EF ≤ 50%)	Standard 10-s, 12-lead ECGs	0.71	73.9	69.2	70.5	7
Heart failure with EF ≤ 40%	ECGs and demographic features	0.89	–	–	–	48
Heart failure with EF ≤ 50%	ECGs and demographic features	0.85	–	–	–	48
LV dysfunction (EF ≤ 40%)	Standard 10-s, 12-lead ECGs	0.83	76.1	72.8	77.8	49
Multiple electrolyte imbalance	Standard 10-s, 12-lead ECGs	0.81–0.87	–	–	–	51
Hyperkalemia	ECG lead I plus lead II (from Minnesota)	0.88	–	90.2	63.2	23
Hyperkalemia	ECG lead I plus lead II (from Florida)	0.86	–	91.3	54.7	23
Hyperkalemia	ECG lead I plus lead II (from Arizona)	0.85	–	88.9	55.0	23
Sleep apnea	Single-lead ECGs during polysomnography recording	–	0.96	96.0	96.0	9
Sleep apnea	Single-lead ECGs during polysomnography	–	0.93	93.0	94.0	9
Sleep apnea	60-s, single-lead ECGs	0.94	87.9	92.0	81.1	60
Sleep apnea	60-s, single-lead ECGs	0.95	87.6	90.3	83.1	59
Hypertension	Segmented ECGs from 24-h Holter	–	99.99	99.99	99.97	64
Myocardial infarction	Segmented ECGs by heartbeat	–	96.0	95.4	97.4	11
Hypertrophic cardiomyopathy	Standard 10-s, 12-lead ECGs	0.96	–	87	90	12
Myocardial scar	Standard 10-s, 12-lead ECGs	0.89	78.0	70.0	84.3	13
1-year all-cause mortality	Standard 10-s, 12-lead ECGs	0.88	–	–	–	10

ECG: electrocardiogram; AUC: area under the curve; ACC: accuracy; SEN: sensitivity; SPE: specificity; CNN: convolutional neural network; DNN: deep neural network; LV: left ventricular; EF: ejection fraction; —: not available.

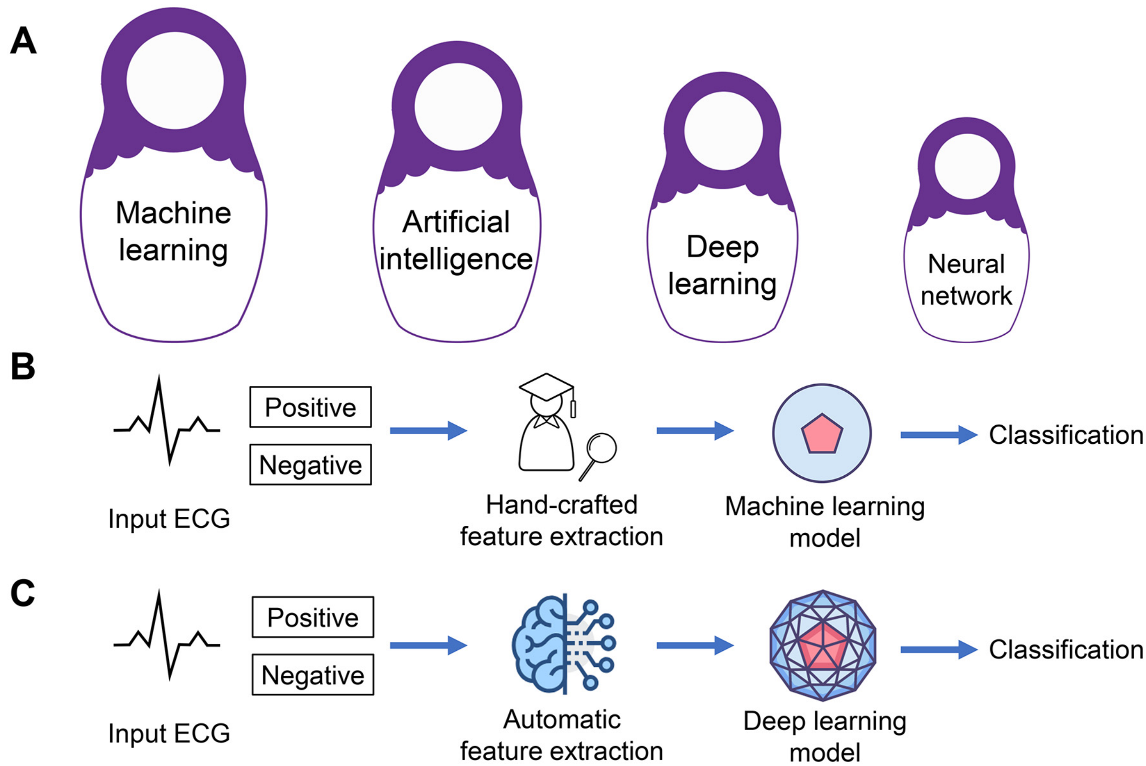

Fig. 2. The difference between deep learning and traditional machine learning. (A) The relationships of artificial intelligence, machine learning, deep learning, and neural network. Like the Russian nesting dolls, each term is a component of the prior term: artificial intelligence is a subfield of machine learning; deep learning is a branch of artificial intelligence; neural networks are the backbone of deep learning. (B) Traditional automatic ECG analysis models identify abnormal electrocardiographic activity based on the specific features in the raw ECG data, which human experts have previously assigned. (C) Deep learning-based methods could automatically extract and decide features related to cardiac structural or functional abnormalities without the feature extraction step by human experts. ECG: electrocardiogram.

[DNN], and recurrent neural networks) have shown remarkable performance in many application areas, including medical image analysis, bioinformatics, drug design, natural language processing, and computer vision [14–16]. The characteristics of these deep learning architectures are summarized in Table 2.

Traditionally, the computerized ECG interpretation methods are designed to extract features in the raw ECG data and match them with the specific 'expert features' previously assigned by human experts [17]. The design of feature extractor requires substantial manpower, and this interpretation strategy is limited by human expert knowledge and data quality [18,19]. It is likely to miss important information in this process since manually acquired ECG features possess only a fraction of the many informative features (Fig. 2B). In contrast, deep learning algorithms possess powerful learning capabilities. They can automatically extract and decide features related to cardiac abnormalities without extensive data pre-processing or hand-crafted feature extraction by human experts [20]. When provided with raw data, deep learning can analyze complex structures of high-dimensional data, identify necessary features, and create models to perform specific tasks without information loss or excessive manpower. For example, when using the traditional machine learning method, we usually input known electrophysiological parameters (e.g., QT interval, QRS duration, R wave axis, or T wave axis) for model training. Differently, deep learning algorithms can acquire features automatically from raw ECG data without additional optimization [4,21–23]. Still, proper data pre-processes, such as band-pass filter, discrete wavelet transform, or short-time Fourier transformation, are necessary for data training since they might improve the algorithm performance. These characteristics endow the deep learning algorithm with unique advantages in interpreting ECGs: deep learning can take raw ECGs as input and outputs diagnostic or predictive probabilities (Fig. 2C). Current data have revealed that deep learning-based ECG features are more informative than traditional expert features [24]. Deep learning showed comparable and, in some cases, even better performance than human experts in multiple diseases [25–28].

Table 2
Characteristics of three representative deep learning architectures.

Architectures	Characteristics
CNN	CNN is regularized versions of multilayer perceptron and is characterized with a mathematical operation called convolution, which is a specialized kind of linear operation. A series of convolutional layers in CNN architecture could convolve with a multiplication or other dot product.
DNN	DNN consists of multiple processing layers, and each layer in DNN architecture can learn the abstract and higher-level representations of the input data.
RNN	RNN is a class type of conventional neural network, which can accept variable sequence input and is particularly suited for sequential data. It performs well especially in speech signal processing and speech recognition fields.

CNN: convolutional neural network; DNN: deep neural network; RNN: recurrent neural networks.

Moreover, deep learning algorithms can analyze ECGs as numeric arrays instead of discrete components. They can identify relationships among the ECG signals and clinical variables, which is well beyond the traditional methods. Although the clinically used ECGs are usually presented as paper or PDF ECG tracings, the 'image form (like a regular photograph)' of ECG is only one way to present the raw data, and ECGs data are actually numeric arrays. It should be highlighted that the careful construction and utilization of the numeric array is what distinguishes deep learning from traditional machine learning technologies. Furthermore, since the training of the deep-learning model relies on the amount of input data, its performance is supposed to be optimized after assimilating vast amounts of ECG. When provided with sufficient training data, a deep learning model can identify multiple features in a data-driven manner, including the subtle features that have not been recognized by human experts [20]. A considerable amount of ECG data have been collected and stored in digital bases, and it provides opportunities for deep learning models to get well trained [29]. Additionally, properly designed continuous feedback and refinement of the deep learning model are likely to improve its classification ability continuously.

3. Arrhythmia

Owing to simplicity and low cost, ECG is the most widely-used method to detect cardiac arrhythmia. Usually, the diagnosis of arrhythmia is performed by automated software screening initially, followed by visual validation by cardiologists or cardiac technicians. This process is time-consuming, subjective, and tedious, even for well-trained experts. The diagnosis of cardiac arrhythmias is probably the most basic application of deep learning in ECG data [30–33].

Atrial fibrillation is a common (lifetime risk of 25–33% [34]) but frequently underdiagnosed class of arrhythmia [35]. The fleeting arrhythmia and cumbersome ambulatory cardiac rhythm monitoring (e.g., Holter monitoring) make it challenging to screen atrial fibrillation, while a deep learning algorithm provides a novel method to identify atrial fibrillation, even during normal sinus rhythm [21,36,37]. Attia et al. [21] recently developed a CNN-based method to diagnose paroxysmal atrial fibrillation during normal sinus rhythm using standard 10-s, 12-lead ECGs. After trained by 454,789 standard ECGs alone without any other information related to atrial fibrillation risk, the CNN model showed remarkable performance in the independent testing set with an area under the curve (AUC) of 0.87, sensitivity of 79.0%, and specificity of 79.5%. In another study by Erdenebayar et al. [38], they created another CNN model to diagnose atrial fibrillation using three datasets provided by PhysioNet (MIT-BIH Atrial Fibrillation Database, Paroxysmal Atrial Fibrillation Prediction Challenge Database, and MIT-BIH Normal Sinus Rhythm Database). After ECG data from the Holter monitor were segmented into 30-s duration, a total of 19,804 normal sinus rhythm ECG segments (11,882 from AF patients and 7922 from normal individuals) were collected and used to train a CNN model. In the testing set, this algorithm showed accuracy, sensitivity, and specificity of 98.7%, 98.6%, and 98.7%, respectively. These studies suggested that deep learning-based algorithms could perform well in identifying the presence of AF using short-term ECGs during normal sinus rhythm.

Previous application of deep learning in ECG is confined to limited single diagnostic tasks (e.g., atrial fibrillation or ventricular tachycardia), while Hannun et al. [39] created the first end-to-end deep learning method to interpret ECG and performed a comprehensive evaluation across a broad range of the most common ECG rhythm classes, including atrial fibrillation/flutter, atrioventricular block, bigeminy, ectopic atrial rhythm, idioventricular rhythm, junctional rhythm, noise, sinus rhythm, supraventricular tachycardia, trigeminy, ventricular tachycardia, and Wenckebach. They collected 91,232 single-lead ECGs from 53,549 patients using Zio monitor [40] (a single-lead, patch-based ambulatory ECG monitoring device) to train the DNN model. The network architecture has 34 layers, with 16 residual blocks and 2 convolutional layers per block. When evaluating in an independent testing dataset, the DNN model achieved an average class-weighted AUC of 0.97. With a specificity similar to the average level of cardiologists, the sensitivity of the DNN model exceeded the average sensitivity achieved by the cardiologist. In another study by Kashou and his colleges [41], they used 2,499,522 standard 10-s, 12-lead ECGs from 720,978 patients to develop and validate a CNN-enabled ECG algorithm. The model performed well in various types of arrhythmia with AUCs ranging from 0.960 to 0.999.

4. Cardiac dysfunction

Cardiac dysfunction is closely associated with a higher risk of hospitalization rate, heart failure progression, and death [42–44]. However, in the early stage of cardiac dysfunction, patients usually present mild signs and symptoms [45]. Deep learning allows the opportunity to develop a non-invasive and population-wide screening tool based on ECGs.

Using 35,970 paired ECGs (standard 10-s, 12-lead) and contemporaneous transthoracic echocardiograms, Attia et al. [4] first proposed a CNN-based screening method to identify patients with reduced left ventricular function (ejection fraction ≤35%). When tested in the dataset consisting of 52,870 ECGs, this algorithm showed an AUC of 0.93, sensitivity of 86.3%, specificity of 85.7% [4]. Interestingly, for patients identified as reduced ejection fraction by the algorithm but the contemporaneous echocardiogram revealed a normal EF (false positive), they represented a 4-fold risk to develop ventricular dysfunction in the next 5 years compared with individuals with a negative screen (age and sex-adjusted hazard ratio = 4.1, 95% confidence interval = 3.3–5.0, $P < 0.001$) [4]. In their following research [22], they further validated the algorithm in 16,056 adult patients, and the algorithm showed excellent diagnostic performance with an AUC, specificity, sensitivity of 0.92, 86.8%, and 82.5%, respectively. This study firstly demonstrated that a well-trained CNN model could detect subtle electrical activity abnormalities related to cardiac dysfunction based on ECG, even before ventricular dysfunction became clinically manifest. The proposed CNN model is now under prospectively evaluation to screen patients with reduced EF (NCT04000087) [46,47]. This ongoing trail serves as a pioneer real-world study to prospectively assess the value of deep learning, which profoundly impacts future implementation strategies [46]. Following Attia and his colleges, we proposed a CNN model to screen patients with left ventricular ejection fraction ≤50% using 26,786 standard 10-s, 12-lead ECGs from the Chinese population [7]. In the testing set, the CNN algorithm showed an overall accuracy, sensitivity, specificity, positive predictive value, and negative predictive value of 73.9%, 69.2%, 70.5%, 70.1%, and 69.9%, respectively.

Although the models proposed by Attia et al. [4] and Sun et al. [7] relied purely on direct computer analysis of ECG signals, they provided high-level accuracy in identifying left ventricular dysfunction automatically without human interaction. In contrast, Kwon et al. [48] combined deep learning with traditional variables selection strategy. Various demographic and ECG features (including age, sex, weight, height, heart rate, presence of atrial fibrillation/flutter, QT interval, QRS duration, R wave axis, and T wave axis) were collected from 55,163 ECGs to develop and evaluate the DNN model to identify patients with reduced ejection fraction (≤ 40% or ≤ 50%). This DNN algorithm consists of 5 hidden layers, 45 nodes, as well as dropout layers. In the external validation, the AUCs were 0.889 for EF ≤ 40% and 0.850 for EF ≤ 50%. T-wave axis, weight, and the presence of atrial fibrillation/flutter play significant roles in this model, with the variable importance of 0.103, 0.087, and 0.073, respectively. However, as was highlighted by the authors [48], the training process heavily relied on the artificially assigned ECG features instead of raw digitized ECG signals, which was a significant limitation of this algorithm and potentially lowered its predictive value.

The application of deep learning in a specific population is another interesting topic. Jentzer et al. [49] apply their previously proposed CNN algorithm [4] patients to identify LV dysfunction in cardiac intensive care unit patients (acute, critically ill population). After trained by standard 10-s, 12-lead ECGs from 5680 patients, the algorithm showed good performance with an AUC of 0.83, specificity of 78%, negative predictive value of 85%, and overall accuracy of 76%. This study showed that deep learning could be used to identify LV dysfunction in resource-limited settings.

5. Electrolyte imbalance

Despite usually being asymptomatic, electrolyte imbalance, especially hyperkalemia, is associated with fatal arrhythmias and is potentially life threatening [50]. However, there lacks of a reliable and non-invasive screening tool for electrolyte imbalance. Deep learning showed remarkable potential in noninvasively detecting and monitoring plasm electrolyte imbalance.

To the best of our knowledge, Kwon et al. [51] developed the first deep learning-based tool to detect multiple electrolyte imbalance types using 83,449 standard 10-s, 12-lead ECGs. All the ECGs were acquired within 30 min before or after laboratory electrolyte examination. In the external test, the deep learning model showed remarkable performance in hyperkalemia, hypokalemia, hypernatremia, hyponatremia, hypercalcemia, and hypocalcemia with AUCs of 0.873, 0.857, 0.839, 0.856, 0.831, and 0.813, respectively.

Galloway et al. [23] proposed a DNN model to detect hyperkalemia (defined as a serum potassium level of ≥5.5 mEq/L) in patients with chronic kidney disease based on ECGs. They retrospectively collected 1,576,581 ECGs from 449,380 patients from Mayo Clinic (Minnesota, Florida, and Arizona), and the serum potassium levels were acquired within 12 h before or after ECGs. In the model using lead I plus lead II, the DNN showed significantly better performance than the traditional statistics method (AUC: 0.83 vs. 0.74). The addition of demographics information did not substantially improve the diagnostic accuracy (AUC: 0.83 vs. 0.85). Previous studies have reported that physician readers could diagnose hyperkalemia based on ECG abnormalities (e.g., peaking of T waves, QRS prolongation, and PR shortening) but with a poor sensitivity of 34–43% [52].

6. Sleep apnea

Sleep apnea significantly increases the risk of multiple cardiovascular diseases, including hypertension, heart failure, atrial fibrillation, and coronary artery disease [53,54]. Currently, polysomnography is a primary method to screen and diagnose sleep apnea, which is time-consuming, inconvenient, and expensive. Since sleep apnea increased the sympathetic and parasympathetic tone [55], ECGs have been proposed as an alternative method to screen sleep apnea [56–58]. With the application of deep learning, the accuracy of the ECG-based method has been substantially improved compared with conventional machine learning [59].

Recently, Erdenebayar et al. [9] applied deep learning methods to diagnose sleep apnea using single-lead ECG signals recorded at 200 Hz. In data pre-process, ECG signals were filtered by a 0.5–30 Hz band-pass to remove background noise and were subsequently segmented for event-based classification. A short-time Fourier transformation was used to convert the ECG singles into 2D spectrogram images to acquire 2D input signals. In the testing set, the 2-dimensional CNN model showed an accuracy of 0.96, sensitivity of 96.0%, and specificity of 96.0, whereas, in the DNN model, the accuracy, sensitivity, and specificity were 0.93, 93.0%, and 94.0%, respectively. Chang et al. [60] proposed another CNN-based sleep apnea detection system using single-lead one-dimensional ECG signals from the PhysioNet Apnea-ECG Database. For per-minute apnea detection, the algorithm achieved an AUC of 0.94, accuracy of 87.9%, specificity of 92.0%, and sensitivity of 81.1%. Moreover, in the study by Wang et al. [59], the LeNet-5 CNN showed a better performance in sleep apnea detection with an AUC, specificity, and sensitivity of 0.95, 90.3%, and 83.1%, respectively.

7. Current challenges and future perspectives

Digitized ECG is a common, low-cost, and widely available tool, which records time-series voltage information from the body surface. Computerized ECG analysis plays a fundamental role in clinical ECG workflow [18]. However, in developing countries or rural areas, inexperienced physicians may fail to interpret ECG data correctly and accept automated diagnosis by commercial interpretation algorithms without criticism, which potentially results in substantial misdiagnosis [18,61,62]. Deep learning algorithms mimic the information processing of the human brain and allow computers to detect generic spatial features directly from the vast amounts of input data, making it a well-suited method to analyze ECG [63]. Studies on the application of deep learning in ECG have been growing explosively in the last 5 years, potentially realizing the optimal utilization of existing ECG data. Deep learning can be used for arrhythmia detection and classification, but more importantly, expand the original functions of ECGs, such as diagnosing atrial fibrillation during normal sinus rhythm, cardiac dysfunction, sleep apnea, hypertension [64], 1-year all-cause mortality [10], and so forth.

Despite the promising performance of deep learning, several challenges remain. ECG data standardization is a vital problem for the following research since there is no standard ECG input type or data pre-process protocol. Most studies used 10-s, 12-lead ECGs performed in the supine position as input data, whereas some other studies used segmented Holter monitor data or non-standard ECG types (i.e., single lead patches and single lead wearable devices). Also, pre-process protocols of input ECG data vary among different studies. For example, many studies took raw ECG image matrix as input [4,21–23], while some studies used band-pass filter [38], discrete wavelet transform [38], short-time Fourier transformation [9], and other pre-processes [7]. It should be highlighted that the model performance may vary due to differences in the approach to signal acquisition and signal processing. However, the ECG data standardization is probably underestimated, and few studies provided the specific data forms (images, raw electrophysiological matrix, or other data forms) or detailed optimization pre-process. The inconsistent input data made the performance of the deep learning algorithm unable to be compared, and therefore, we are currently unaware which form or pre-process method could realize the full potential of deep learning technology.

Reproducibility or generalizability poses another significant barrier before clinical practice [65]. Deep learning performance heavily relies on the amount and quality of the input ECG data and might be notoriously inconsistent. However, most studies only collect data from one single center or publicly available digital dataset [66], which results in potential bias. The ECG acquisition is affected by ECG devices, operators, external electromagnetic signals, skeletal muscle electrical signals, electrode contact, electrode placement, body habitus, and races. Although the large size of input data can minimize this bias, small patient samples (< 100) in some studies is probably cover the actual algorithm performance by over-fit and weakens the promotion of capacity.

The imbalanced data remain a common problem in many studies, which might result in misleading algorithm performance. For example, Attia et al. [4] created a CNN-based screening method to identify patients with left ventricular dysfunction, while only about 7.8% of individuals had an ejection fraction of ≤35%. The imbalanced data distribution impairs the learning efficiency and limits the test efficiency, and the reliable accuracy (83.5%) was likely owing to the excellent performance in normal participants (a positive predictive value of 33.8; a negative predictive value of 98.7%). Besides, the algorithm performed in patients with other cardiovascular diseases (e.g., myocardial infarction, arrhythmia, or cardiomyopathy) should be investigated since these patients presented characteristic ECG changes. In the following research, it is necessary to validate the diagnostic or predictive performance in specific populations, such as patients with basic cardiovascular diseases.

Explainability remains one of the key issues to be solved to achieve the trust of clinicians and insert deep learning algorithm into clinical workflow. An explainable algorithm model should provide a detailed and understandable process on how a decision or prediction is made. However, the current deep learning algorithms are primarily programmed to find solutions directly, and it is ambiguous how the algorithms are operated, which makes the system a 'black box'. ECG variables are conventionally well-defined by human experts, including heart rate, propagation velocity, repolarization, impulse creation, activation/repolarization sequence, and so forth. Although deep learning has been demonstrated to add important diagnostic and prognostic information to the ECG interpretation, even for those identified as normal by physicians, it remains unclear what features or physiological bases are used to make the prediction. The disturbances captured by surface ECGs are beyond human recognition, but deep learning algorithms can detect these subtle alterations. Nevertheless, the classification is biologically plausible since the disorders in the heart affect electrical impulse creation, propagation, and repolarization in heart tissues by various mechanisms. For example, in the deep-learning algorithm to diagnose atrial fibrillation during normal sinus rhythm, myocyte hypertrophy, fibrosis, and chamber enlargement are likely to form the subtle ECG changes for underlying atrial fibrillation prediction. Wavelets smaller than the observable P wave might also reflect the regional non-sinus electrical activity in these patients. However, human experts would not routinely recognize these features or formally report the observation. The development of explainable deep learning algorithms remains in the primary stage, and many studies are still under ongoing investigation [67–69]. Until recently, Jo et al. [70] developed and validated the first explainable artificial intelligence to analyzed ECG data. In the following research, one important strategy is to apply deep learning to select the interpretable features in ECGs, whereas another primary solution is to provide visible explanations of the algorithm output [71].

Currently, most input ECG types for network training were standard 10-s, 12-lead ECGs or segmented Holter signals, while the importance of other ECG types might be underestimated, especially the ECG data collected by wearable devices. It was reported that about 13% of the US population owned smartwatches, with 40% showing interest to get one, which suggested a large amount of potential ECG data [72]. Recently, many wearable devices have been developed and tested for continuous cardiac electrical monitoring [73–77]. For example, the FDA (DEN180044) approved Apple Watch to record, store, and display a single-lead ECG like a Lead I ECG [78,79]. Moderate to a strong agreement was reported to exist between lead-1 waveforms in Apple Watch and traditional 12-lead ECG [80]. With AliveCor KardiaBand accessory, Apple Watch can detect atrial fibrillation with high sensitivity and specificity [81]. The combination of deep learning and commercial wearable devices could substantially increase the accessibility of ECG signals and provide opportunities for long-term cardiac surveillance and management [82]. The following research would be necessary to extend the deep learning model to non-standard ECG data collected by wearable devices. These algorithms could also be incorporated into the electronic health record system to identify suspected cardiac diseases, which would support cardiologists to perform further examination. Additionally, most deep learning algorithms are designed based on only the raw ECG samples without other ECG- or patient-related features, while the addition of certain features assigned by human experts might improve model performance, including demographic, clinical, metabolic, and ECG factors.

Before clinical application, all deep learning algorithms should be trained, tested, and prospectively trialed on diverse populations across various medical facilities in different countries. Owing to not requiring substantial manpower, the deep learning algorithm can be developed relatively quickly, making extensive validation a possible strategy to demonstrate the reproducibility and generalizability of deep learning algorithms. Therefore, it is vital to build a publicly available dataset designed for algorithm training or validation with balanced data. After prospective validation, deep learning will contribute to revolutionizing the ECG-based clinical monitoring and management strategies.

Finally, despite the rapid development in deep learning, there is currently no evidence suggesting that human experts' roles will ever be removed in ECG analysis. It should be highlighted that the deep learning algorithms are designed to perform the computer-assisted interpretation as an adjunct decision support system for human experts instead of replacing their roles. The central role of well-trained cardiologists in ECG analysis remains immovable in the foreseeable future.

8. Conclusion

Owing to the increased computational power and growing digital ECG data availability, the application of deep learning in ECG has been widely studied. Despite the promising performance, the application of deep learning in ECG remains in its infancy stage, and there are many problems to be solved before clinical usage. Nevertheless, the application of deep learning in ECG has led us to a new area of electrocardiology and made us stand at the edge of ECG innovation.

Data availability statement

NA.

Funding

None.

Declaration of competing interest

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Acknowledgment

Qiang Qu acknowledged the guidance given by Prof. Xin-Li Li (The First Affiliated Hospital of Nanjing Medical University) during the postgraduate stage. Jin-Yu Sun sincerely acknowledged Prof. Xin-Li Li, Dr. Fang Zhou, Dr. Yuan He, and Dr. Shen-Gen Liao for their generous help and support during his resident training.

References

[1] W.B. Fye, A history of the origin, evolution, and impact of electrocardiography, Am. J. Cardiol. 73 (1994) 937–949.
[2] S.S. Barold, Willem Einthoven and the birth of clinical electrocardiography a hundred years ago, Card. Electrophysiol. Rev. 7 (2003) 99–104.
[3] H. Holst, M. Ohlsson, C. Peterson, L. Edenbrandt, A confident decision support system for interpreting electrocardiograms, Clin. Physiol. 19 (1999) 410–418.
[4] Z.I. Attia, S. Kapa, F. Lopez-Jimenez, P.M. McKie, D.J. Ladewig, G. Satam, P.A. Pellikka, M. Enriquez-Sarano, P.A. Noseworthy, T.M. Munger, S.J. Asirvatham, C.G. Scott, R.E. Carter, P.A. Friedman, Screening for cardiac contractile dysfunction using an artificial intelligence-enabled electrocardiogram, Nat. Med. 25 (2019) 70–74.
[5] Y. Qian, Y. Qiu, C.C. Li, Z.Y. Wang, B.W. Cao, H.X. Huang, Y.H. Ni, L.L. Chen, J.Y. Sun, A novel diagnostic method for pituitary adenoma based on magnetic resonance imaging using a convolutional neural network, Pituitary 23 (2020) 246–252.
[6] Y. Jiang, C. Jin, H. Yu, J. Wu, C. Chen, Q. Yuan, W. Huang, Y. Hu, Y. Xu, Z. Zhou, G.A. Fisher Jr., G. Li, R. Li, Development and validation of a deep learning ct signature to predict survival and chemotherapy benefit in gastric cancer: a multicenter, retrospective study, Ann. Surg. (2020) https://doi.org/10.1097/SLA.0000000000003778 Online ahead of print.
[7] J.Y. Sun, Y. Qiu, H.C. Guo, Y. Hua, B. Shao, Y.C. Qiao, J. Guo, H.L. Ding, Z.Y. Zhang, L.F. Miao, N. Wang, Y.M. Zhang, Y. Chen, J. Lu, M. Dai, C.Y. Zhang, R.X. Wang, A method to screen left ventricular dysfunction through ECG based on convolutional neural network, J. Cardiovasc. Electrophysiol. 32 (4) (2021) 1095–1102, https://doi.org/10.1111/jce.14936.
[8] A.H. Ribeiro, M.H. Ribeiro, G.M.M. Paixao, D.M. Oliveira, P.R. Gomes, J.A. Canazart, M.P.S. Ferreira, C.R. Andersson, P.W. Macfarlane, W. Meira Jr., T.B. Schon, A.L.P. Ribeiro, Automatic diagnosis of the 12-lead ECG using a deep neural network, Nat. Commun. 11 (2020) 1760.
[9] U. Erdenebayar, Y.J. Kim, J.U. Park, E.Y. Joo, K.J. Lee, Deep learning approaches for automatic detection of sleep apnea events from an electrocardiogram, Comput. Methods Prog. Biomed. 180 (2019) 105001.
[10] S. Raghunath, A.E. Ulloa Cerna, L. Jing, D.P. van Maanen, J. Stough, D.N. Hartzel, J.B. Leader, H.L. Kirchner, M.C. Stumpe, A. Hafez, A. Nemani, T. Carbonati, K.W. Johnson, K. Young, C.W. Good, J.M. Pfeifer, A.A. Patel, B.P. Delisle, A. Alsaid, D. Beer, C.M. Haggerty, B.K. Fornwalt, Prediction of mortality from 12-lead electrocardiogram voltage data using a deep neural network, Nat. Med. 26 (2020) 886–891.
[11] W. Liu, M. Zhang, Y. Zhang, Y. Liao, Q. Huang, S. Chang, H. Wang, J. He, Real-time multilead convolutional neural network for myocardial infarction detection, IEEE J. Biomed. Health Inform. 22 (2018) 1434–1444.
[12] W.Y. Ko, K.C. Siontis, Z.I. Attia, R.E. Carter, S. Kapa, S.R. Ommen, S.J. Demuth, M.J. Ackerman, B.J. Gersh, A.M. Arruda-Olson, J.B. Geske, S.J. Asirvatham, F. Lopez-Jimenez, R.A. Nishimura, P.A. Friedman, P.A. Noseworthy, Detection of hypertrophic cardiomyopathy using a convolutional neural network-enabled electrocardiogram, J. Am. Coll. Cardiol. 75 (2020) 722–733.
[13] N. Gumpfer, D. Grun, J. Hannig, T. Keller, M. Guckert, Detecting myocardial scar using electrocardiogram data and deep neural networks, Biol. Chem. (2020) https://doi.org/10.1515/hsz-2020-0169 Online ahead of print.

[14] D. Dey, P.J. Slomka, P. Leeson, D. Comaniciu, S. Shrestha, P.P. Sengupta, T.H. Marwick, Artificial intelligence in cardiovascular imaging: JACC state-of-the-art review, J. Am. Coll. Cardiol. 73 (2019) 1317–1335.
[15] A. de Marvao, T.J. Dawes, J.P. Howard, D.P. O'Regan, Artificial intelligence and the cardiologist: what you need to know for 2020, Heart 106 (2020) 399–400.
[16] A.A. Miyazawa, Artificial intelligence: the future for cardiology, Heart 105 (2019) 1214.
[17] S. Hong, Y. Zhou, J. Shang, C. Xiao, J. Sun, Opportunities and challenges of deep learning methods for electrocardiogram data: a systematic review, Comput. Biol. Med. 122 (2020) 103801.
[18] J. Schlapfer, H.J. Wellens, Computer-interpreted electrocardiograms: benefits and limitations, J. Am. Coll. Cardiol. 70 (2017) 1183–1192.
[19] J.S. Alpert, Can you trust a computer to read your electrocardiogram? Am. J. Med. 125 (2012) 525–526.
[20] Y. LeCun, Y. Bengio, G. Hinton, Deep learning, Nature 521 (2015) 436–444.
[21] Z.I. Attia, P.A. Noseworthy, F. Lopez-Jimenez, S.J. Asirvatham, A.J. Deshmukh, B.J. Gersh, R.E. Carter, X. Yao, A.A. Rabinstein, B.J. Erickson, S. Kapa, P.A. Friedman, An artificial intelligence-enabled ECG algorithm for the identification of patients with atrial fibrillation during sinus rhythm: a retrospective analysis of outcome prediction, Lancet 394 (2019) 861–867.
[22] Z.I. Attia, S. Kapa, X. Yao, F. Lopez-Jimenez, T.L. Mohan, P.A. Pellikka, R.E. Carter, N.D. Shah, P.A. Friedman, P.A. Noseworthy, Prospective validation of a deep learning electrocardiogram algorithm for the detection of left ventricular systolic dysfunction, J. Cardiovasc. Electrophysiol. 30 (2019) 668–674.
[23] C.D. Galloway, A.V. Valys, J.B. Shreibati, D.L. Treiman, F.L. Petterson, V.P. Gundotra, D.E. Albert, Z.I. Attia, R.E. Carter, S.J. Asirvatham, M.J. Ackerman, P.A. Noseworthy, J.J. Dillon, P.A. Friedman, Development and validation of a deep-learning model to screen for hyperkalemia from the electrocardiogram, JAMA Cardiol. 4 (2019) 428–436.
[24] S. Hong, Y. Zhou, M. Wu, J. Shang, Q. Wang, H. Li, J. Xie, Combining deep neural networks and engineered features for cardiac arrhythmia detection from ECG recordings, Physiol. Meas. 40 (2019), 054009.
[25] G.D. Clifford, C. Liu, B. Moody, L.H. Lehman, I. Silva, Q. Li, A.E. Johnson, R.G. Mark, AF classification from a short single lead ECG recording: the PhysioNet/computing in cardiology challenge 2017, Comput. Cardiol. 2017 (2010) 44.
[26] J. Park, Can artificial intelligence prediction algorithms exceed statistical predictions? Korean Circ. J. 49 (2019) 640–641.
[27] H. Bleijendaal, L.A. Ramos, R.R. Lopes, T.E. Verstraelen, Baalman SWE, M.D. Oudkerk Pool, Tjong FVY, F.M. Melgarejo-Meseguer, F.J. Gimeno-Blanes, J.R. Gimeno-Blanes, A.S. Amin, M. Winter, H.A. Marquering, Kok WEM, A.H. Zwinderman, Wilde AAM, Y.M. Pinto, Computer versus cardiologist: is a machine learning algorithm able to outperform an expert in diagnosing a phospholamban p.Arg14del mutation on the electrocardiogram? Heart Rhythm. 18 (1) (2021) 79–87, https://doi.org/10.1016/j.hrthm.2020.08.021.
[28] S.W. Smith, B. Walsh, K. Grauer, K. Wang, J. Rapin, J. Li, W. Fennell, P. Taboulet, A deep neural network learning algorithm outperforms a conventional algorithm for emergency department electrocardiogram interpretation, J. Electrocardiol. 52 (2019) 88–95.
[29] M. Smith, R. Saunders, L. Stuckhardt, J.M. McGinnis (Eds.),Best Care at Lower Cost: The Path to Continuously Learning Health Care in America Washington (DC), 2013.
[30] S.S. Xu, M.W. Mak, C.C. Cheung, Towards end-to-end ECG classification with raw signal extraction and deep neural networks, IEEE J. Biomed. Health Inform. 23 (2019) 1574–1584.
[31] X. Zhang, K. Gu, S. Miao, X. Zhang, Y. Yin, C. Wan, Y. Yu, J. Hu, Z. Wang, T. Shan, S. Jing, W. Wang, Y. Ge, Y. Chen, J. Guo, Y. Liu, Automated detection of cardiovascular disease by electrocardiogram signal analysis: a deep learning system, Cardiovasc Diagn Ther. 10 (2020) 227–235.
[32] J. Zhang, A. Liu, M. Gao, X. Chen, X. Zhang, X. Chen, ECG-based multi-class arrhythmia detection using spatio-temporal attention-based convolutional recurrent neural network, Artif. Intell. Med. 106 (2020) 101856.
[33] F. Murat, O. Yildirim, M. Talo, U.B. Baloglu, Y. Demir, U.R. Acharya, Application of deep learning techniques for heartbeats detection using ECG signals-analysis and review, Comput. Biol. Med. 120 (2020) 103726.
[34] D.M. Lloyd-Jones, T.J. Wang, E.P. Leip, M.G. Larson, D. Levy, R.S. Vasan, R.B. D'Agostino, J.M. Massaro, A. Beiser, P.A. Wolf, E.J. Benjamin, Lifetime risk for development of atrial fibrillation: the Framingham Heart Study, Circulation 110 (2004) 1042–1046.
[35] T. Sanna, H.C. Diener, R.S. Passman, V. Di Lazzaro, R.A. Bernstein, C.A. Morillo, M.M. Rymer, V. Thijs, T. Rogers, F. Beckers, K. Lindborg, J. Brachmann, C.A. Investigators, Cryptogenic stroke and underlying atrial fibrillation, N. Engl. J. Med. 370 (2014) 2478–2486.
[36] W. Cai, Y. Chen, J. Guo, B. Han, Y. Shi, L. Ji, J. Wang, G. Zhang, J. Luo, Accurate detection of atrial fibrillation from 12-lead ECG using deep neural network, Comput. Biol. Med. 116 (2020) 103378.
[37] S. Kwon, J. Hong, E.K. Choi, E. Lee, D.E. Hostallero, W.J. Kang, B. Lee, E.R. Jeong, B.K. Koo, S. Oh, Y. Yi, Deep learning approaches to detect atrial fibrillation using photoplethysmographic signals: algorithms development study, JMIR Mhealth Uhealth 7 (2019), e12770.
[38] U. Erdenebayar, H. Kim, J.U. Park, D. Kang, K.J. Lee, Automatic prediction of atrial fibrillation based on convolutional neural network using a short-term normal electrocardiogram signal, J. Korean Med. Sci. 34 (2019), e64.
[39] A.Y. Hannun, P. Rajpurkar, M. Haghpanahi, G.H. Tison, C. Bourn, M.P. Turakhia, A.Y. Ng, Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network, Nat. Med. 25 (2019) 65–69.
[40] M.P. Turakhia, D.D. Hoang, P. Zimetbaum, J.D. Miller, V.F. Froelicher, U.N. Kumar, X. Xu, F. Yang, P.A. Heidenreich, Diagnostic utility of a novel leadless arrhythmia monitoring device, Am. J. Cardiol. 112 (2013) 520–524.
[41] A. Kashou, W.Y. Ko, Z.I. Attia, M. Cohen, P. Friedman, P. Noseworthy, A comprehensive artificial intelligence-enabled electrocardiogram interpretation program, J. Am. Coll. Cardiol. 75 (2020) 3504.
[42] T.A. McDonagh, K. McDonald, A.S. Maisel, Screening for asymptomatic left ventricular dysfunction using B-type natriuretic peptide, Congest Heart Fail. 14 (2008) 5–8.
[43] P. Jong, S. Yusuf, M.F. Rousseau, S.A. Ahn, S.I. Bangdiwala, Effect of enalapril on 12-year survival and life expectancy in patients with left ventricular systolic dysfunction: a follow-up study, Lancet 361 (2003) 1843–1848.
[44] H. Mehta, A. Armstrong, K. Swett, S.J. Shah, M.A. Allison, B. Hurwitz, S. Bangdiwala, R. Dadhania, D.W. Kitzman, W. Arguelles, J. Lima, M. Youngblood, N. Schneiderman, M.L. Daviglus, D. Spevack, G.A. Talavera, A. Raisinghani, R. Kaplan, C.J. Rodriguez, Burden of systolic and diastolic left ventricular dysfunction among hispanics in the United States: insights from the echocardiographic study of Latinos, Circ. Heart Fail. 9 (2016), e002733.
[45] M.M. Redfield, S.J. Jacobsen, J.C. Burnett Jr., D.W. Mahoney, K.R. Bailey, R.J. Rodeheffer, Burden of systolic and diastolic ventricular dysfunction in the community: appreciating the scope of the heart failure epidemic, JAMA. 289 (2003) 194–202.
[46] X. Yao, R.G. McCoy, P.A. Friedman, N.D. Shah, B.A. Barry, E.M. Behnken, J.W. Inselman, Z.I. Attia, P.A. Noseworthy, ECG AI-Guided Screening for Low Ejection Fraction (EAGLE): rationale and design of a pragmatic cluster randomized trial, Am. Heart J. 219 (2020) 31–36.
[47] X. Yao, R.G. McCoy, P.A. Friedman, N.D. Shah, B.A. Barry, E.M. Behnken, J.W. Inselman, Z.I. Attia, P.A. Noseworthy, Clinical trial design data for electrocardiogram artificial intelligence-guided screening for low ejection fraction (EAGLE), Data Brief 28 (2020) 104894.
[48] J.M. Kwon, K.H. Kim, K.H. Jeon, H.M. Kim, M.J. Kim, S.M. Lim, P.S. Song, J. Park, R.K. Choi, B.H. Oh, Development and validation of deep-learning algorithm for electrocardiography-based heart failure identification, Korean Circ J. 49 (2019) 629–639.
[49] J.C. Jentzer, A.H. Kashou, Z.I. Attia, F. Lopez-Jimenez, S. Kapa, P.A. Friedman, P.A. Noseworthy, Left ventricular systolic dysfunction identification using artificial intelligence-augmented electrocardiogram in cardiac intensive care unit patients, Int. J. Cardiol. 326 (2021) 114–123.
[50] L.M. Einhorn, M. Zhan, V.D. Hsu, L.D. Walker, M.F. Moen, S.L. Seliger, M.R. Weir, J.C. Fink, The frequency of hyperkalemia and its significance in chronic kidney disease, Arch. Intern. Med. 169 (2009) 1156–1162.
[51] J.M. Kwon, M.S. Jung, K.H. Kim, Y.Y. Jo, J.H. Shin, Y.H. Cho, Y.J. Lee, J.H. Ban, K.H. Jeon, S.Y. Lee, J. Park, B.H. Oh, Artificial intelligence for detecting electrolyte imbalance using electrocardiography, Ann. Noninvasive Electrocardiol. 26 (3) (2021), e12839, https://doi.org/10.1111/anec.12839.
[52] K.D. Wrenn, C.M. Slovis, B.S. Slovis, The ability of physicians to predict hyperkalemia from the ECG, Ann. Emerg. Med. 20 (1991) 1229–1232.
[53] E. Marulanda-Londono, S. Chaturvedi, The interplay between obstructive sleep apnea and atrial fibrillation, Front. Neurol. 8 (2017) 668.
[54] J.R. Tietjens, D. Claman, E.J. Kezirian, T. De Marco, A. Mirzayan, B. Sadroonri, A.N. Goldberg, C. Long, E.P. Gerstenfeld, Y. Yeghiazarians, Obstructive sleep apnea in cardiovascular disease: a review of the literature and proposed multidisciplinary clinical management strategy, J. Am. Heart Assoc. 8 (2019), e010440.
[55] V.A. Rossi, J.R. Stradling, M. Kohler, Effects of obstructive sleep apnoea on heart rhythm, Eur. Respir. J. 41 (2013) 1439–1451.
[56] A.H. Khandoker, C.K. Karmakar, M. Palaniswami, Automated recognition of patients with obstructive sleep apnoea using wavelet-based features of electrocardiogram recordings, Comput. Biol. Med. 39 (2009) 88–96.
[57] L. Chen, X. Zhang, H. Wang, An obstructive sleep apnea detection approach using kernel density classification based on single-lead electrocardiogram, J. Med. Syst. 39 (2015) 47.
[58] E. Urtnasan, J.U. Park, K.J. Lee, Multiclass classification of obstructive sleep apnea/hypopnea based on a convolutional neural network from a single-lead electrocardiogram, Physiol. Meas. 39 (2018), 065003.
[59] T. Wang, C. Lu, G. Shen, F. Hong, Sleep apnea detection from a single-lead ECG signal with automatic feature-extraction through a modified LeNet-5 convolutional neural network, PeerJ 7 (2019), e7731.
[60] H.Y. Chang, C.Y. Yeh, C.T. Lee, C.C. Lin, A sleep apnea detection system based on a one-dimensional deep convolution neural network model using single-lead electrocardiogram, Sensors (Basel) 20 (2020).
[61] T.S. Gruca, T.H. Pyo, G.C. Nelson, Providing cardiology care in rural areas through visiting consultant clinics, J. Am. Heart Assoc. 5 (2016).
[62] A.P. Shah, S.A. Rubin, Errors in the computerized electrocardiogram interpretation of cardiac rhythm, J. Electrocardiol. 40 (2007) 385–390.
[63] K. Shameer, K.W. Johnson, B.S. Glicksberg, J.T. Dudley, P.P. Sengupta, Machine learning in cardiovascular medicine: are we there yet? Heart 104 (2018) 1156–1164.
[64] D.C.K. Soh, E.Y.K. Ng, V. Jahmunah, S.L. Oh, R.S. Tan, U.R. Acharya, Automated diagnostic tool for hypertension using convolutional neural network, Comput. Biol. Med. 126 (2020) 103999.
[65] P.A. Noseworthy, Z.I. Attia, L.C. Brewer, S.N. Hayes, X. Yao, S. Kapa, P.A. Friedman, F. Lopez-Jimenez, Assessing and mitigating Bias in medical artificial intelligence: the effects of race and ethnicity on a deep learning model for ECG analysis, Circ. Arrhythm. Electrophysiol. 13 (2020), e007988.
[66] A.L. Goldberger, L.A. Amaral, L. Glass, J.M. Hausdorff, P.C. Ivanov, R.G. Mark, J.E. Mietus, G.B. Moody, C.K. Peng, H.E. Stanley, PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals, Circulation 101 (2000) E215–E220.

[67] Q. Xie, M. Zhou, Q. Zhao, Z. Xu, D. Meng, MHF-Net: an interpretable deep network for multispectral and hyperspectral image fusion, IEEE Trans. Pattern Anal. Mach. Intell. (2020) https://doi.org/10.1109/TPAMI.2020.3015691 Online ahead of print.

[68] N. Strodthoff, P. Wagner, T. Schaeffter, W. Samek, Deep learning for ECG analysis: benchmarks and insights from PTB-XL, IEEE J. Biomed. Health Inform. 25 (5) (2021) 1519–1528, https://doi.org/10.1109/JBHI.2020.3022989.

[69] S.W.E. Baalman, F.E. Schroevers, A.J. Oakley, T.F. Brouwer, W. van der Stuijt, H. Bleijendaal, L.A. Ramos, R.R. Lopes, H.A. Marquering, R.E. Knops, J.R. de Groot, A morphology based deep learning model for atrial fibrillation detection using single cycle electrocardiographic samples, Int. J. Cardiol. 316 (2020) 130–136.

[70] Y.Y. Jo, Y. Cho, S.Y. Lee, J.M. Kwon, K.H. Kim, K.H. Jeon, S. Cho, J. Park, B.H. Oh, Explainable artificial intelligence to detect atrial fibrillation using electrocardiogram, Int. J. Cardiol. 328 (2021) 104–110.

[71] M. Sermesant, H. Delingette, H. Cochet, P. Jais, N. Ayache, Applications of artificial intelligence in cardiovascular imaging, Nat. Rev. Cardiol. (2021) https://doi.org/10.1038/s41569-021-00527-2 Online ahead of print.

[72] J.M. Raja, C. Elsakr, S. Roman, B. Cave, I. Pour-Ghaz, A. Nanda, M. Maturana, R.N. Khouzam, Apple watch, wearables, and heart rhythm: where do we stand? Ann. Transl. Med. 7 (2019) 417.

[73] A. Samol, K. Bischof, B. Luani, D. Pascut, M. Wiemer, S. Kaese, Single-lead ECG recordings including einthoven and wilson leads by a smartwatch: a new era of patient directed early ECG differential diagnosis of cardiac diseases? Sensors (Basel) 19 (2019).

[74] A. Behzadi, A. Sepehri Shamloo, K. Mouratis, G. Hindricks, A. Arya, A. Bollmann, Feasibility and reliability of smartwatch to obtain 3-lead electrocardiogram recordings, Sensors (Basel) 20 (2020).

[75] J. Lazaro, N. Reljin, M.B. Hossain, Y. Noh, P. Laguna, K.H. Chon, Wearable armband device for daily life electrocardiogram monitoring, IEEE Trans. Biomed. Eng. 67 (2020) 3464–3473.

[76] M. Shao, Z. Zhou, G. Bin, Y. Bai, S. Wu, A wearable electrocardiogram telemonitoring system for atrial fibrillation detection, Sensors (Basel) 20 (2020).

[77] O.E. Santala, J.A. Lipponen, H. Jantti, T.T. Rissanen, J. Halonen, I. Kolk, H. Pohjantahti-Maaroos, M.P. Tarvainen, E.S. Valiaho, J. Hartikainen, T. Martikainen, Necklace-embedded electrocardiogram for the detection and diagnosis of atrial fibrillation, Clin. Cardiol. 44 (5) (2021) 620–626.

[78] A. Samol, K. Bischof, B. Luani, D. Pascut, M. Wiemer, S. Kaese, Recording of bipolar multichannel ECGs by a smartwatch: modern ECG diagnostic 100 Years after einthoven, Sensors (Basel) 19 (2019).

[79] M. Strik, T. Caillol, F.D. Ramirez, S. Abu-Alrub, H. Marchand, N. Welte, P. Ritter, M. Haissaguerre, S. Ploux, P. Bordachar, Validating QT-interval measurement using the Apple Watch ECG to enable remote monitoring during the COVID-19 pandemic, Circulation 142 (2020) 416–418.

[80] Apple Watch Series 4 vs. 12-Lead ECG: A Comparison Of Manual Electrocardiographic Waveform Analysis. 2021. https://doi.org/10.1016/S0735-1097%2820%2931005-6 Last accessed:2021-06-02.

[81] K. Rajakariar, A.N. Koshy, J.K. Sajeev, S. Nair, L. Roberts, A.W. Teh, Accuracy of a smartwatch based single-lead electrocardiogram device in detection of atrial fibrillation, Heart 106 (2020) 665–670.

[82] J. Wasserlauf, C. You, R. Patel, A. Valys, D. Albert, R. Passman, Smartwatch performance for the detection and quantification of atrial fibrillation, Circ. Arrhythm. Electrophysiol. 12 (2019), e006834.