A Review on Deep Learning Methods for ECG Arrhythmia Classification

Zahra Ebrahimi[a,*], Mohammad Loni[b,*], Masoud Daneshtalab[b], Arash Gharehbaghi[b,**]

[a]*Shahrood University of Technology, Shahroud, Iran*
[b] *Department of Innovation, Design and Technology, Mälardalen University, 72123 Västerås, Sweden*

Abstract

Deep Learning (DL) has recently become a topic of study in different applications including healthcare, in which timely detection of anomalies on Electrocardiogram (ECG) can play a vital role in patient monitoring. This paper presents a comprehensive review study on the recent DL methods applied to the ECG signal for the classification purposes. This study considers various types of the DL methods such as Convolutional Neural Network (CNN), Deep Belief Network (DBN), Recurrent Neural Network (RNN), Long Short-Term Memory (LSTM), and Gated Recurrent Unit (GRU). From the 75 studies reported within 2017 and 2018, CNN is dominantly observed as the suitable technique for feature extraction, seen in 52% of the studies. DL methods showed high accuracy in correct classification of Atrial Fibrillation (AF) (100%), Supraventricular Ectopic Beats (SVEB) (99.8%), and Ventricular Ectopic Beats (VEB) (99.7%) using the GRU/LSTM, CNN, and LSTM, respectively.

Keywords: Electrocardiogram, Deep Learning, Computer-Aided Diagnosis, Smart Health-care

1. Introduction

Cardiovascular Disease (CVD) is the main cause of human death, responsible for 31% of the worldwide deaths in 2016 (Benjamin et al., 2018), from which 85% happened due to heart attack. The annual burden of CVD on the European and American economy is estimated to be €210 billion and $555 billion, respectively (Wilkins et al., 2017; Benjamin et al., 2018). The traditional CVD diagnosis paradigm is based on individual patient's medical history and clinical examinations. These results are interpreted according to a set of the quantitative medical parameters to classify the patients based on the taxonomy of medical diseases.

In many cases, the traditional rule-based diagnosis paradigm is inefficient due to dealing with large amount of heterogeneous data, and requires significant analysis and medical expertise to achieve adequate accuracy in diagnosis. The problem will become more pronounced in places, where there is a lack of medical experts and clinical equipment, especially in developing countries. This motivates the requirement for a reliable, automatic, and low-cost system for monitoring and diagnosis. This requirement is becoming more demanded by the healthcare providers, such that appropriate medical assessments can be linked to utilizing Compute Aided Diagnosis Systems Computer-Aided Diagnosis (CADS). A CADS is composed of automatic monitoring procedures of health conditions working based on analysis of physiological signals for monitoring and evaluating functionality of the corresponding organ. CADSs provide individuals with portable and straightforward solutions to make them informed about their diseases.

Electrocardiogram (ECG) is a non-stationary physiological signal, representing electrical activity of heart. It is not only used to look for pathological patterns among the heartbeats, but also used to measure the beats' regularity as well as other conditions like mental stress.

Deep Neural Network (DNN) has been widely used for classification and prediction purposes in different domains. Recently, it has been noticed that DNNs are being developed sharply with a significant effect on the accuracy in clas-

*Equal Contribution
**Corresponding Author
Email address: `arash.ghareh.baghi@mdh.se` (Arash Gharehbaghi)
URL: `zahraebrahimi114@gmail.com` (Zahra Ebrahimi), `mohammad.loni@mdh.se` (Mohammad Loni), `masoud,daneshtalab@mdh.se` (Masoud Daneshtalab)

sification for a wide range of medical tasks. Modern CADS systems leverage DNNs to detect arrhythmia of captured ECG signal leading to decrease the cost of continuous heart monitoring and improving the quality of predictions. However, an ECG-based automatic arrhythmia classification is typically faced with several important challenges.

1.1. Arrhythmia Classification Challenges

The main challenges of CADS in arrhythmia classification can be summarized as follows:

1. The symptoms of the arrhythmia might not be seen during the ECG signal capturing period (Ceylan & Özbay, 2007).
2. ECG signal properties (such as period, and amplitude) vary from person to person and depends on different factors such as age, gender, physical conditions, and lifestyle. Finding a generalized framework along with the related standards to be functional for general population is problematic (Ceylan & Özbay, 2007; Joshi et al., 2009).
3. Morphology of ECG signal is often not stationary even for one testing person because physical state such as running, walking, and sleeping.
4. The volume of data to be considered for ECG signal analysis is large. Hence there is a higher probability of having a false diagnosis of arrhythmia.
5. The noise, artifacts and interference can result in morphological variations and discrepancies in the captured ECG signal (Adams & Choi, 2012; Dinakarrao et al., 2019).

1.2. The Study Objectives

The main objective of this study is to cover a broad range of Deep Learning (DL) topics in arrhythmia classification. To this end, we first show a big picture of most common learning models used in the studied papers (see Section 2). Then, we present an overview of the arrhythmias from the medical perspective (see Section 3), performance evaluation metrics of ECG classifiers (see Section 5), and the existing ECG databases (see Section 4). The second objective is to provide a tabular representation to be used as a quick reference. Therefore, we categorized the studied papers according to ❶ their main focus to be on heart arrhythmia(s), ❷ their utilized DNNs for both feature extraction and classification, and ❸ variants of different Deep Learning methods for arrhythmia classification. The final objective of this review is to analyze arrhythmia classification methods in terms of technical limitations, performance, and the inference overhead (see Section 9).

1.3. Contributions of This Study Paper

We summarize and compare notable studies within 2017 and 2018 based on the DL-based methods to overcome the challenges exist in arrhythmia classification. The main contributions of this review are listed below:

1. We reviewed the structure of different popular DL-based methods employed in the related studies.
2. Presenting an overview of the characteristics of the notable heart arrhythmia considered in the reviewed papers.
3. Presenting the widely accepted datasets as well as the evaluation metrics exist in this community for detecting and comparing different arrhythmia.
4. ECG arrhythmia is presented in a categorized manner based on the classification method, dataset and the papers cited them.
5. analyzing different arrhythmia classification methods along with comparing them based on their reported performance.
6. Finally, discussing on the conclusions explicitly obtained from this paper by doing the following analysis:
    - Analyzing the contribution percentage of each learning method in the studied papers in order to find the most popular technique.
    - Analyzing the contribution percentage of each arrhythmia in the studied papers in order to target the most interesting and less considered applications.
    - Presenting the most accurate arrhythmia classification method along with reported accuracy in order to help researchers to select the technique depending on their needs.

1.4. Paper Organization

This paper is organized as follows: Section 2 presents a general overview of DL methods used in the throughout of this review. Section 3 gives the medical background, needed to gain sufficient understanding of ECG characteristics discussed in this paper. Section 4 describes ECG databases used for training and testing. Common metrics accepted in the community for measuring, and comparing the accuracy and quality of the results, are all presented in Section 5. Section 6 presents the research methodology of the paper. In Section 7, different taxonomies of the reviewed papers in terms of the DL-based categorization, and the heart diseases based categorization is presented. Section 8 reviews the outstanding methods in detail while summarizing all the other papers in Table 6 to Table 11. In addition, we present the search results in this section. Further discussions on the limitations, DL computational complexity, and future research trend for ECG arrhythmia classification are presented in Section 9. Finally, Section 10 concludes the paper.

1.5. Acronyms

The acronyms of cardiology and DL terms used in this paper are listed in glossary Section, Appendix A.

2. Deep Learning Techniques

The topic of Deep Learning (DL) refers to the studies on knowledge extraction, predictions, intelligent decision making, or in another term recognizing intricate patterns using a set of the data, so called training data.Comparing to the traditional learning techniques, DNNs are more scalable since higher accuracy is usually achieved by increasing the size of the network or the training dataset. Shallow learning models such as decision trees and Support Vector Machine (SVMs) are inefficient for many modern applications, meaning that they require a large number of observations for achieving generalizability, and imposing significant human labour to specify prior knowledge in the model (Goodfellow et al., 2016)(Loni et al., 2020).

In the recent years, several Deep Learning (DL) models have been proposed to improve the accuracy of different learning tasks, including Multilayer Perceptron (MLP), Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), Long Short-Term Memory (LSTM), and Deep Belief Network (DBN). Time-growing neural network is an elaboration of time-delayed neural network, recently introduced to the context of learning theory (Gharehbaghi et al., 2014)(A Gharehbaghi, 2015). Although the idea of deep time growing neural network is well-tailored for biological signals, especially those with cyclic characteristics (Gharehbaghi & Lindén, 2018)(Gharehbaghi et al., 2019b)(Gharehbaghi et al., 2019a)(Gharehbaghi & Babic, 2018), application of this powerful method has not been studied for ECG classification, yet.

2.1. Multilayer Perceptron (MLP)

MLP is the most frequently used supervised neural network appearing effective in learning complex systems. The MLP architecture is variable, however, it consists of several layers of neurons connected to each other in a feed-forward manner. Each neuron is the weighted sum of its inputs passed through a non-linear function (Goodfellow et al., 2016).

2.2. Convolutional Neural Network (CNN)

CNN is one of the most popular DNN architecture usually trained by a gradient-based optimization algorithm (LeCun et al., 1998). In general, a CNN consists of multiple back-to-back layers connected in a feed-forward manner. The main layers are including convolutional layer, normalization layer, pooling layer, and fully-connected layer. Three first layers are responsible for extracting features, while fully-connected layers are in charge of classification. In Fig. 1, a general architecture of the CNN is represented for the classification task (Ciresan et al., 2012). Table 2 shows different popular CNN architectures where their efficiency has been proved for different problems (Appendix B).

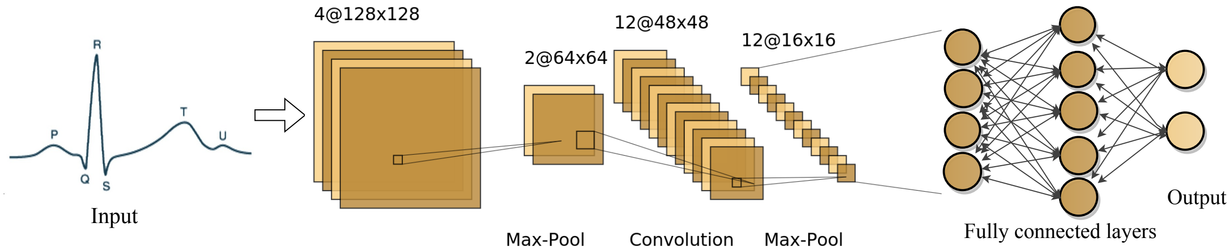


Figure 1: Illustration of Convolutional Neural Network (CNN) architecture.

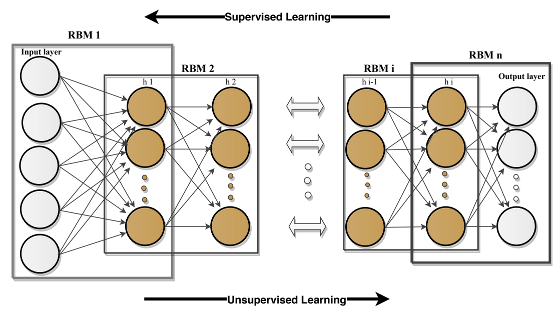


Figure 2: Illustration of Deep Belief Network (DBN) architecture (Qiao et al., 2018).

2.3. Deep Belief Network (DBN)

In 2006, Hinton proposed DBNs which are composed of multiple Restricted Boltzmann Machine (RBM) layers. DBN is a powerful learning model used to model evolving random variables over time. As Fig. 2 shown, the DBN layers are composed of RBMs. Each RBM, within a given layer, receives the inputs of the previous layer and feeds the RBM in the next layer. Training DBNs is conducted by training RBMs, layer by layer from bottom to up.

RBM has been proposed in 1986 (Hinton et al., 1986). RBM is an undirected model for binary random variables used effectively in modeling distributions over binary-valued data. A Boltzmann machine is a particular type of Markov random fields that are composed of symmetric networks with binary random units (Keyvanrad & Homayounpour, 2015). Each RBM contains a layer of visible units that represent the data and a layer of hidden units that learn to represent features and capture higher-order correlations. As seen in Fig. 3, the two layers are connected by weighted connections, Wij, and there is no connection within a layer.

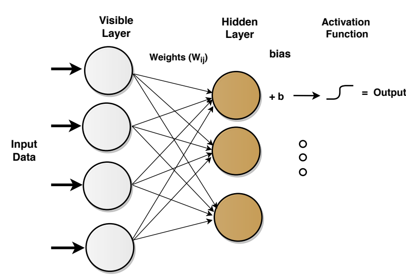


Figure 3: The architecture of RBM model. White nodes are *Visible Units* and brown nodes are *Hidden Units* (Hinton et al., 2006).

2.4. Recurrent Neural Network (RNN)

RNN is an extension of an Artificial Neural Network (ANN) whose weights are shared across time. RNN is the most proper learning model for learning sequential input data and the time-series data classification where the feedback and the present value is fed again into the network and the output contains the adding of values in the memory (Liu

& Kim, 2018). At each time step, the RNN receives an input, updates its hidden state, and makes a prediction. RNN uses gradient descent algorithm through time for training the weights. Fig. 4 illustrates the underlying architecture of the RNN. RNNs has highly dynamic behavior due to nonlinear activation functions used by the hidden units.

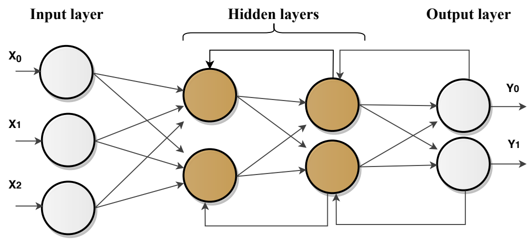


Figure 4: The architecture of Deep Belief Network (DBN).

2.5. *Long Short-Term Memory (LSTM)*

LSTM is a specific type of traditional RNN designed for temporal sequences and the long-range dependencies (Chung et al., 2014; LeCun et al., 2015). LSTM uses memory blocks instead of simple RNN units where each memory block includes one or more memory cells with a pair of adaptive multiplicative gates as the input and output (Fig. 5). A memory block places information and updates them across time-steps based on the input and output gates. The gates control the input and output flow of information to a memory cell.

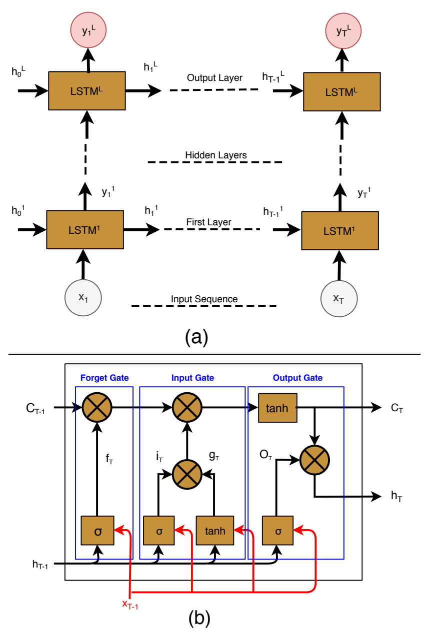


Figure 5: (a) General Structure of Long Short-Term Memory (LSTM) architecture. (b) Detailed structure of acrlonglstm (LSTM) functionality.

2.6. *Bidirectional Recurrent Neural Network (BRNN)*

The main goal of BRNN is to simultaneously get information from past and future states of the sequence by connecting two hidden layers of opposite directions to the same output (Schuster & Paliwal, 1997) (Fig. 6). LSTM-BRNN can be easily achieved by replacing the nonlinear units in Fig. 6 with the LSTM blocks.

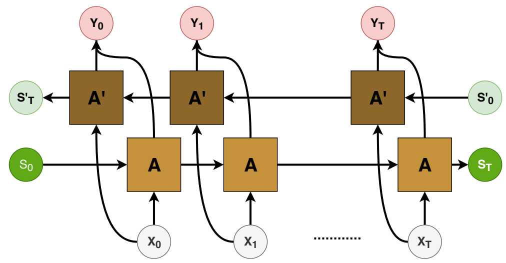


Figure 6: The illustration of the Bidirectional Recurrent Neural Network (BRNN) architecture.

2.7. Gated Recurrent Unit (GRU)

GRU is an improved version of LSTM with faster training process(Chung et al., 2014) (Fig. 7). It is simpler than LSTM with less computational complexity. GRU consists of gates that are collectively involved in balancing the interior flow of units' information. Input gate and forget gate are combined and formed a new gating unit typically called as update gate. The update gate mainly focuses on balancing the state between the previous activation and the candidate activation.

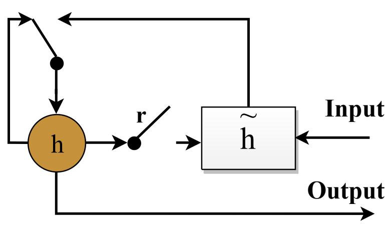


Figure 7: The architecture of Gated Recurrent Unit (GRU).

3. Medical Background

This section gives a overview about the heart diseases that can be commonly detected from the ECG signal. The ECG morphology reflects the heart status Kasper et al. (2018). In general, ECG provides two primary types of information. First, by measuring time intervals on ECG, a cardiologist can determine how long the electrical wave takes to pass through electrical conduction system of the heart. This information helps to find out if the electrical activity is regular or irregular, fast or slow. Second, by measuring the strength of electrical activity, a cardiologist is able to find out if parts of the heart are too large or are overworked. Fig. 8 shows a normal ECG heartbeat sample with different meaningful segments, including three important waves showing *atrial depolarization (*P-wave*)*, *ventral depolarization (*QRS *complex wave)*, and *repolarization (*T-wave*)*. Any disorder in electrical activity of heart neural cells affects ECG signals, known as arrhythmia. The most common types of arrhythmia are breifly described in the following sequels:

3.1. Atrial Fibrillation (AF)

AF occurs when action potentials fire very rapidly within the atrium, resulting in a rapid atrial rate (roughly 400-600 beats/minute). Therefore, *P* waves will not be seen since the atrial rate is so fast with low amplitude level (TRIAL, 2011) (Fig. 9.b).

3.2. Right Bundle Branch Block (RBBB) and Left Bundle Branch Block (LBBB)

Bundle Branch Block is an interruption in the regular conduction system that leads to abnormal *QRS* morphology. Typically, the right bundle depolarizes the Right Ventricle (RV). In an RBBB, the right bundle does not activate. The right ventricle is instead depolarized by spreading the impulse from the left bundle through the Left Ventricle

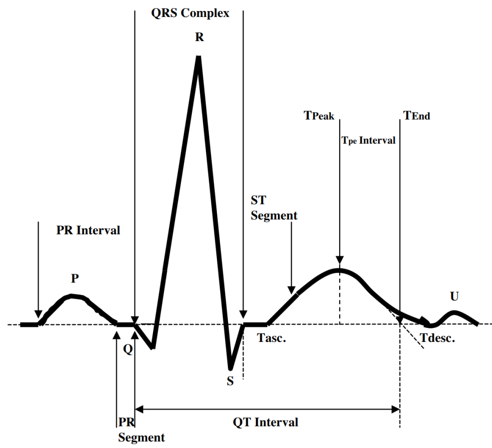


Figure 8: Influential segments and various usual intervals of a A pseudo Normal Sinus Rhythm (NSR). Source: (Pater, 2005).

(LV) and then to the RV. This pattern of electrical spread creates an aberrant *QRS* morphology. Typically, the left bundle depolarizes the LV. In an LBBB, the left bundle does not activate. The LV is instead depolarized by spread of impulse from the right bundle through the RV and then to the LV. This pattern of electrical spread creates an aberrant *QRS* morphology (Otten, 2005). Fig. 9.c and Fig. 9.d illustrates a sample ECG signal presenting LBBB and RBBB, respectively.

3.3. Premature Atrial Contraction (PAC)) and Premature Ventricular Contraction (PVC)

PAC and PVC occur when the heart's regular rhythm is interrupted by a premature or early beat. If the premature beat arises from the atria, it is called a PAC. If it arises from the ventricles, it is called PVC. Fig. 9.e and Fig. 9.f illustrates a sample ECG signal presenting PAC and PVC, respectively.

3.4. Ectopic Beats

Ectopic atrial rhythms happen when a site outside of the sinus node within the atria creates action potentials faster than the sinus node (with an atrial rate less than 100 beats/minute). Since this electrical activity does not originate from the sinus node, the *P* wave would not have its normal sinus appearance (Fig. 9.g). Ectopic beats are also frequent during periods of stress or exercise, and they may happen by consumption of some foods such as alcohol (TRIAL, 2011).

3.5. Myocardial Infarction (MI)

MI (aka heart attack) happens when blood flow decreases or stops in a part of the heart, causing permanent damage to the heart muscle or arteries. Fig. 9.h shows the ECG diagram of MI. some of the MI patterns include the two below groups:

1. Those with *ST* segment elevation or new *RBBB/LBBB*.
2. Those with *ST* segment depression or *T-wave* inversion.

3.6. Fusion Beat

A fusion beat happens when electrical impulses from different sources act upon the same region of the heart simultaneously. It is called a Ventricular Fusion Beats (VFB) if it acts upon the ventricular chambers, whereas colliding currents in the atrial chambers produce Atrial Fusion Beats (*AFB*) (Conover, 2002; Huff, 2006).

3.7. Sinus Bradycardia

Sinus bradycardia is a sinus rhythm with a lower than normal rate (fewer than 60 beats per minute). The decreased heart rate causes decreased cardiac output resulting in symptoms such as lightheadedness, dizziness, hypotension, vertigo, and syncope (Thornton & Hochachka, 2004) (Fig. 9.i).

3.8. Tachycardia

Tachycardia happens when the heart rate exceeds the normal resting rate (so-called tachyarrhythmia). Generally, a resting heart rate over 100 beats per minute in adults is accepted as tachycardia (Awtry et al., 2006). Fig. 9.j illustrates the ECG pattern of *Tachycardia*. Types of tachycardias are including:

1. Atrial or Supraventricular Tachycardia (SVT): is a fast heart rate staring in the upper heart chambers.
2. Sinus Tachycardia: happens when heart sends out electrical signals faster than usual leading to a normal increase in the heart rate.
3. Ventricular Tachycardia (VT): is a series of more than three abnormal consecutive *QRS* complex heart rhythm with a duration beyond 120 ms and the *ST-T* vector that points opposite the *QRS* deflection (Bonow et al., 2011).

3.9. Atrial Flutter (AFL)

AFL is a prevalent abnormal heart rhythm that starts in the atrial chambers of the heart (Sawhney et al., 2009; Link, 2012). When it first occurs, it is usually associated with a fast heart rate and is classified as a type of SVT (Fig. 9.k).

3.10. Ventricular Flutter (VF)

It is an unstable arrhythmia in which a tachycardia affecting the ventricles with a rate of over 150-300 beats per minute. VF is a possible transition stage between VT and fibrillation that can cause sudden cardiac death (Bonow et al., 2011). A sinusoidal waveform characterizes it without clear definition of the *T-waves* and *QRS*.

3.11. Ventricular Fibrillation (VFib)

VFib is a cardiac arrhythmia in which the heart quivers instead of pumping due to disorganized electrical activity in the ventricles characterized by showing irregular unformed *QRS* complexes without any clear *P-waves* (Baldzizhar et al., 2016; Weiler et al., 2014) (Fig. 9.l). VFib results in cardiac arrest with loss of consciousness followed by death in the absence of treatment (Baldzizhar et al., 2016; Weiler et al., 2014).

3.12. Idioventricular Rhythm

An idioventricular rhythm is highly similar to VT but with the ventricular rate less than 60 beats per minute. Therefore, the idioventricular rhythm is referred as a slow ventricular tachycardia.

3.13. Ventricular Bigeminy

Ventricular Bigeminy is an abnormal cardiac rhythm problem in which there are repeated rhythms heartbeats that each sinus beat is followed by an ectopic beat and pause frequently.

3.14. Pacemaker Rhythm

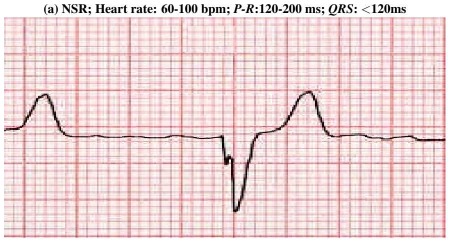
Pacemaker clinical syndrome representing the consequences of pacemaker implantation, regardless of the pacing mode, due to suboptimal atrioventricular synchrony or dyssynchrony (Chalvidan et al., 2000). It is an iatrogenic disease resulting from medical treatment (Frielingsdorf et al., 1994). Individuals with a low heart rate before pacemaker implantation are more at risk of developing pacemaker syndrome. Patients who develop pacemaker syndrome may require pacemaker adjustment or fitting of another lead for better coordinating the timing of atrial and ventricular contraction.

4. Databases

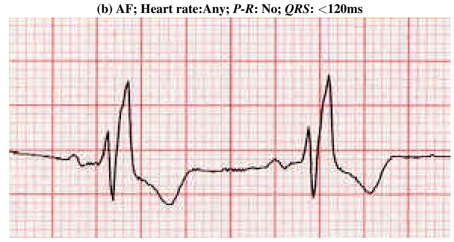
For adhering the ethical aspects, most of the papers use the existing ECG records provided as online databases. The most common ECG databases such as PhysioNet, MITDB, PTB, etc are labeled as normal and abnormal groups of rhythms to train CADS systems. Table 3 specifies popular existing ECG databases used for many years in community (Appendix C).

(a) NSR; Heart rate: 60-100 bpm; *P-R*:120-200 ms; *QRS*: <120ms

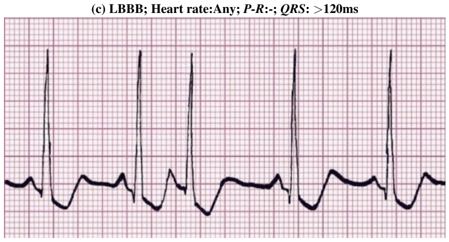
(b) AF; Heart rate:Any; *P-R*: No; *QRS*: <120ms



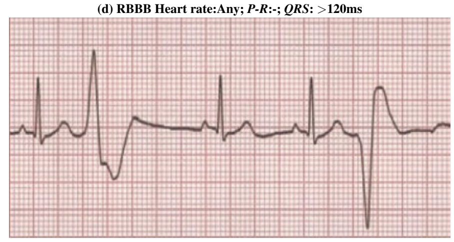
(c) LBBB; Heart rate:Any; *P-R*:-; *QRS*: >120ms



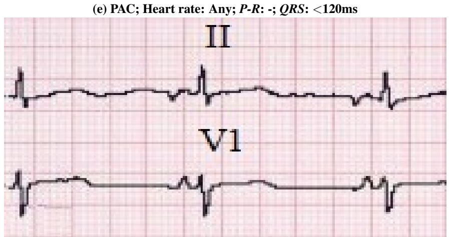
(d) RBBB Heart rate:Any; *P-R*:-; *QRS*: >120ms



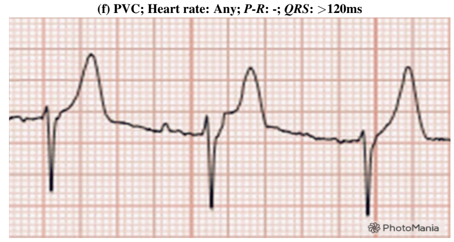
(e) PAC; Heart rate: Any; *P-R*: -; *QRS*: <120ms



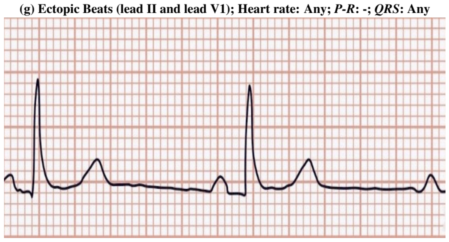
(f) PVC; Heart rate: Any; *P-R*: -; *QRS*: >120ms



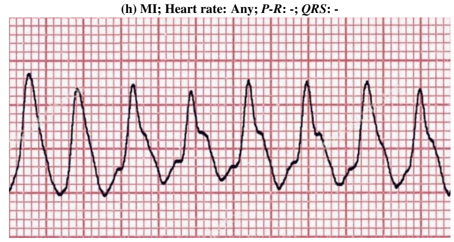
(g) Ectopic Beats (lead II and lead V1); Heart rate: Any; *P-R*: -; *QRS*: Any



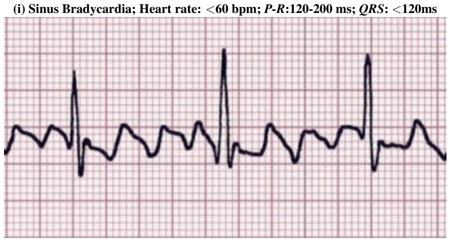
(h) MI; Heart rate: Any; *P-R*: -; *QRS*: -



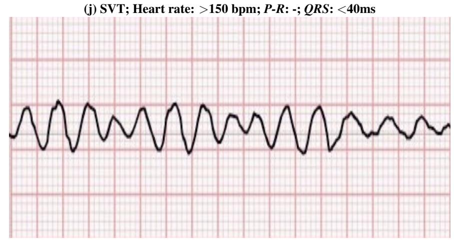
(i) Sinus Bradycardia; Heart rate: <60 bpm; *P-R*:120-200 ms; *QRS*: <120ms



(j) SVT; Heart rate: >150 bpm; *P-R*: -; *QRS*: <40ms



(k) AFL; Heart rate: Any; *P-R*: No; *QRS*: <120ms



(l) VFib; Heart rate: >250bpm; *P-R*: -; *QRS*: >120

Figure 9: Illustrating different arrhythmias including: (a) Normal Sinus Rhythm , (b) Atrial Fibrillation, (c) Left Bundle Branch Block, (d) Right Bundle Branch Block, (e) Premature Atrial Contraction, (f) Premature Ventricular Contraction, (g) Ectopic Beats (illustrating both lead II and lead V1), (h) Myocardial Infarction, (i) Sinus Bradycardia, (j) Atrial or Supraventricular Tachycardia, (k) Atrial Flutter, and (l) Ventricular Fibrillation.

5. Performance Measurements

This section presents common quantitative metrics used for evaluation of classifiers' performance. The classification resulted from a learning method, can be either abnormal case or normal, named as positive class or negative class, respectively. Result of the prediction can also be either true or false, implying on correct prediction or incorrect prediction, respectively. Thus, We can summarize classification into four possible states:

1. *True positive (TP)*: Correct prediction of positive class
2. *True negative (TN)*: Correct prediction of negative class
3. *False positive (FP)*: Incorrect prediction of positive class
4. *False negative (FN)*: Incorrect prediction of negative class

Based on the classifications predictions, the Accuracy, Specificity, Sensitivity, Precision, Recall, Positive Predictive Value (PPV), Negative Predictive Value (NPV) and Area under the Curve (AUC) are calculated in Equation 1 to Equation 8.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

$$Specificity = \frac{TN}{TN + FP} \tag{2}$$

$$Sensitivity = \frac{TP}{TP + FN} \tag{3}$$

$$Precision = \frac{TP}{TP + FP} \tag{4}$$

$$Recall = \frac{TP}{TP + FN} \tag{5}$$

$$PPV = \frac{TP}{TP + FP} \tag{6}$$

$$NPV = \frac{TN}{TN + FN} \tag{7}$$

$$\begin{gathered} AUC = \frac{1}{n_p} \sum_{j=1}^{n_p} f_j \\ f_j = \frac{1}{T} \sum_{t=1}^{T} w_t \; |1 \; if \; P_j \; and \; 0 \; otherwise \\ w_t = \frac{1}{2}(\mathrm{prec}_{t+1} - \mathrm{prec}_{t-1}) \\ \mathrm{prec}_t = \frac{\# \; of \; points \; i \; where \; p_i \; and \; c_i = 1}{\# \; of \; points \; i \; where \; p_i} \end{gathered} \tag{8}$$

The traditional F-measure (F_1 score) is the harmonic mean of precision and recall:

$$F_1 = \left(\frac{recall^{-1} + precision^{-1}}{2} \right)^{-1} = 2. \, \frac{precision \; . \; recall}{precision + recall} \tag{9}$$

6. Research Methodology

A topical survey is retrospectively performed on the reachable reports, published in the technical, interdisciplinary and medical journals within 2017 and 2018, when the was introduced to the medical context. *PubMed* and *Google Scholar* are employed as the main search engines, using the following keywords:

1. Deep Learning (DL) and Electrocardiogram (ECG)
2. Deep Neural Network (DNN) and Electrocardiogram (ECG)
3. Convolutional Neural Network (CNN) and Electrocardiogram (ECG)
4. Deep Belief Network (DBN) and Electrocardiogram (ECG)
5. Recurrent Neural Network (RNN) and Electrocardiogram (ECG)
6. Long Short-Term Memory (LSTM) and Electrocardiogram (ECG)
7. Gated Recurrent Unit (GRU) and Electrocardiogram (ECG)

Results of the survey are studied in light of two different manners: technical and applicative contents of the publications. In this perspective, the objective of the reviewed papers are basically categorized by technical and application, where the former includes classification and feature extraction, and the later contains classification of different kinds of arrhythmia. It is worth noting that the deep learning methods are sometimes employed for feature extraction to provide informative inputs to another classifier, i.e. conventional classifier, in contrary to other applications in which the deep learning methods serve as powerful classifiers. Superiority of different methods for specific research questions of symptom detection are investigated and a pervasive comparison is performed.

7. Taxonomy of the Review

This section categorized the studied papers in the four following groups including: ❶ Method-based categorization of feature extraction papers (Section 7.1). ❷ Method-based categorization of classification papers (Section 7.2). ❸ Method-based categorization of both feature extraction and classification papers (Section 7.3). ❹ Arrhythmia-based categorization of all the studied papers (Section 7.4).

7.1. DL Methods Applied as Feature Extraction

There are a few papers that used DL techniques just as feature extraction (Table 4 in Appendix D). Although feature selection by DL speeds up the process, our study indicates that the results are not excellent for finding abnormal heartbeat. For example Li et al. (Li et al., 2018a) proposed considerable results on obstructive sleep apnea (OSA) detection, they used two traditional classifiers including SVM and MLP. All in all, selecting features is a big challenge and sometimes is not possible due to noise and unsustainability, leading researchers to perform trial and error.

7.2. DL Methods Applied as Classification

Significant portion of the studied papers used DL techniques as the classification part. Xia (Xia et al., 2017) proposed the best classification performance for detecting AF. Table 4 presents studied papers that use DNNs for classification (Appendix D).

7.3. DL Methods Applied as Both Feature Extraction and Classification

The majority of studied papers in this review applied at least one type of DL technique for feature extraction and/or classification. According to the experimental results, DL are proven to be robust and efficient (Table 4). For instance, Zhang (Zhang et al., 2017) proposed excellent results on detecting VEB, and SVEB by applying LSTM (a network with two LSTM layers and two FCN layers). Although, DNNs can provide prosperous result, some cases show a violation due to the inherent uncertainty in the biological signals. For example Zhong et al. (Zhong et al., 2018) employed a CNN in both parts for fetal *QRS* complex detection and the result was not perfectly good.

7.4. Arrhythmia-Based Categorization

Table 5 lists arrhythmia-Based heart diseases considered in the studied papers (Appendix E). Table 5 is highly profitable for researchers to more efficiently fetch papers that contain a specified heart disease.

8. Results of the Review

In total, a number of 77 publications were found, from which 2 publications were excluded from the study, as their common focus was noise removal that is well beyond the objective of the review. From the rest of the 75 publications, We found 5 survey publications. Results of these 5 surveys are all compared and represented in Section 9. This section presents a technical overview of the outstanding studies regarding the highest reported accuracy on ECG-based arrhythmia diagnosis. Besides, the summary of other studied articles are presented in Table 6 to Table 11.

8.1. Variants of Multilayer Perceptron (MLP)

Table 6 lists the specifications of all papers that used a MLP model for arrhythmia diagnosis (Appendix F). In addition, the MLP techniques with highest accuracy are explained in below.

Sannino and De Pietro (Sannino & De Pietro, 2018) proposed a novel DL approach for classifying NSR, SVEB, VEB, and fusion of Ventricular and NSR. They found the best classification performance by proposing a MLP composed of seven hidden layers with the ReLU activation function, and 5, 10, 30, 50, 30, 10 and 5 neurons in each layer, respectively. The output layer leverages *Softmax* activation function, and the cost function was the cross-entropy. Signals are located on the *P*, *R*, and *T* peaks and proceeded to segment the ECG signal into single heartbeats. Accuracy of the results were 100% on the training set, 99.09% on the test set and 99.68% on the Whole data.

Li et al. (Li et al., 2018a) proposed a method to detect Obstructive Sleep Apnea (OSA) based on DNN and Hidden Markov model (HMM) using a single-lead ECG signal. They used the verified R-peaks position to compute the RR interval series and interpolate the *RR* interval series into 100 points. DNN extracted the features. Two types of classifiers (SVM and ANN) were used to classify the features.

8.2. Variants of Convolutional Neural Network (CNN)

CNN is widely used in various applications such as noise filtering, feature learning, and classifications. In general, classification using CNNs is in the supervised learning approach. Table 7 lists the specifications of other papers using CNN model for arrhythmia diagnosis (Appendix G). In addition, the CNN techniques with highest accuracy are explained in below.

Liu et al. (Liu et al., 2018) proposed a multiple-feature-branch Convolutional Neural Network (MFB-CNN) for automated myocardial (MI) detection and localization using ECG. Each independent feature branch of the MFB-CNN corresponded to a particular lead. The global fully-connected Softmax layer could have exploited the integrity, summarizing all the feature branches. Based on the DL framework, no hand-designed features were used for analysis. Furthermore, the patient-specific paradigm was adopted to manage the inter-patient variability, which was a significant challenge for automated diagnosis. For class-based MI detection and localization, the average accuracies are 99.95% and 99.81%, respectively. For patient-specific experiment, the average accuracies of MI detection and localization are 98.79% and 94.82%, respectively.

Andreotti et al. (Andreotti et al., 2017) classified short segments of ECG into four distinct classes as part of the PhysioNet database including NSR and AF. They compared a state-of-the-art feature-based classifier with a CNN approach. They increased the number of AF and noisy recordings by 2,000 10-s ECG segments with AF from PhysioBank, Circulation 2000. Each ECG segment was preprocessed using 10^{th} order band-pass Butterworth filters with 5Hz and 45Hz cut-off frequencies for narrow-band and 1Hz to 100Hz for wide-band filtering. They divided the preprocessed ECG signals into 10-second segments with 50% overlap. They computed the features based on each segment and then computed the summary statistics such as mean standard deviation and min/max. They used the 34 layers ResNet (see Table 2) and 16 convolutional filters per layer. The feature-based classifier obtained an F_1-score of 72.0% and 79% on the training set (5-fold cross-validation) and on the hidden test set, respectively. Similarly, CNN scored 72.1% on the augmented database and 83% on the test set. The latter method resulted in a final score of 79%.

Another best consequence is Al Rahhal et al. (Al Rahhal et al., 2018) proposed a CNN for VEB, and SVEB classification. They utilized a continuous wavelet transform (CWT) and an 11-layer CNN. The utilized MITDB, INCART, and SVDB databases. The maximum average accuracy on MITDB database for VEB and SVEB is 99.3% and 99.3%, respectively. Regarding the other databases, the obtained average accuracy by the method in for VEB is equal to 99.23% (INCART database), and 99.4% (SVDB database). For SVEB, the average accuracy is 99.82% for INCART database and 98.4% for SVDB database.

8.3. Variants of Deep Belief Network (DBN)

There are a few papers applied DBN in their work for arrhythmia classification, therefore, DBN is highly potential for further research. Table 8 lists the specifications of other papers using DBN model for arrhythmia diagnosis (Appendix H).

Sayantan et al. (Sayantan et al., 2018) proposed a feature representation using Gaussian-Bernoulli Deep Belief Network (GB-DBN), and a linear SVM classifier has been considered to train the models for the classification task. The visible layer is a Gaussian RBM since the input features are real valued and the rest of layers are Bernoulli RBMs. The method achieved an accuracy of 99.5% in for SVEB and 99.4% accuracy for VEB on MIT-BIH Arrhythmia Database. Also, it provides accuracy of 97.5% for SVEB and 98.6% for VEB on SVDB database.

Taji et al. (Taji et al., 2018) proposed a method to reduce the false alarm rate caused by poor-quality ECG measurements during AF detection. They designed a DBN with three layers of RBMs. The first two RBMs were generative RBMs which did not need labels, and the last layer included discriminative RBM which used data with their labels and classified the input data. Results show that for ECG with low Signal-Noise-Ratio (SNR), gating which is a remember data mechanism, significantly improved the performance of AF detection. Without gating, the precision, recall, accuracy, and specificity at 20 dB were 25.5%, 29.3%, 58.7%, and 70.5%, respectively. With gating, there was a significant improvement with these metrics becoming 65%, 68.1%, 81%, and 85%.

8.4. Variants of Recurrent Neural Network (RNN)

Table 9 lists the specifications of other papers using RNN model for arrhythmia diagnosis (Appendix I). In addition, the RNN techniques with highest accuracy are explained in below.

Wang et al. (Wang et al., 2019) proposed a global and updatable classification scheme named Global Recurrent Neural Network (GRNN). Their has three main innovations. First, relying on the large capacity and fitting ability of GRNN. Second, the GRNN improves generalization performance when training samples and test samples are from distinct databases. Finally, GRNN automatically learns the underlying differences among the samples from different classes. The GRNN has four layers in total. In the morphological part, LSTM blocks were applied instead of traditional RNN to memorize longer history. A 20-node fully-connected layer was added after the second LSTM layer. The GRNN showed great fitting ability and high performance on the training set, with a minimum accuracy of 99.8% in VEB and SVEB detection.

Zhang et al. (Zhang et al., 2017) proposed a patient-specific ECG classification to detect NSR, VEB, and SVEB. They use RNN to learn time correlation of ECG signal points. Morphology information of the ECG signal including the$T$ wave of former beat and present beat are fed into RNN to learn the deep features automatically. According to the experimental results, the classification accuracy for SVEB and VEB are 98.7% and 99.4%, respectively.

8.5. Variants of Long Short-Term Memory (LSTM)

Table 10 lists the specifications of other papers using LSTM model for arrhythmia diagnosis (Appendix J).

Yildirim. (Yildirim, 2018) proposed a new model named () for classifying ECG signals. Two filter banks consisted of high-pass and low-pass filters used for reducing noises. A new wavelet-based layer is used to generate ECG signal sequences. In this layer, the ECG signals were decomposed into frequency sub-bands at different scales. These sub-bands were used as sequences for the input of LSTM networks. They used the MIT-BIH arrhythmia database for considering five different types of heartbeats. These five types were NSR, PVC, Paced Beat, RBBB, and LBBB. The results showed that the model gave a high recognition performance of 99.39%. It had been observed that the wavelet-based layer proposed in the study significantly improved the recognition performance of CNN.

Faust et al. (Faust et al., 2018) proposed a DL model to detect AF beats. The data was partitioned with a sliding window of 100 beats. The resulting signal blocks were directly fed into an RNN with LSTM. The system was validated and tested with data from the MIT-BIH Atrial Fibrillation Database. It achieved 98.51% accuracy with 10-fold cross-validation (20 subjects) and 99.77% with blindfold validation (3 subjects). The proposed structure of system was straight forward because there was no need for information reduction through feature extraction.

8.6. Variants of Gated Recurrent Unit (GRU)

Table 11 lists the specifications of other papers using GRU model for arrhythmia diagnosis (Appendix K). In addition, the RNN techniques with highest accuracy are explained in below.

Singh et al. (Singh et al., 2018) proposed GRU, RNN and LSTM models for the effective detection of arrhythmia from ECG signals that consisted of sixteen types of heartbeats divided into two groups of normal and arrhythmia heartbeats. They evaluated three different neural networks. First, three layers of RNN had been used with 128, 256 and 100 neurons in each layer, respectively, with nine iterations. Second, a GRU with two gates, a reset gate, and an update gate. In this paper, three layers of RNN-GRU (Gated Recurrent Unit) have been used with 64, 128 and 100 number of neurons in each layer, respectively (with five iterations). Third, using LSTM to model temporal sequences and the long-range dependencies. The LSTM showed accuracy of 88.1%, sensitivity of 92.4% and specificity of 83.35%. There were 64, 256 and 100 neurons per hidden layer, respectively which showed better detection of arrhythmia than RNN and GRU as the accuracy of RNN was 85.4%, sensitivity was 80.6%, specificity was 85.7%, and GRU accuracy was 82.5%, sensitivity was 78.9%, and specificity was 81.5%.

Sujadevi (Sujadevi et al., 2017) employed different DL methods such as RNN, LSTM, and GRU to detect the AF faster in the given electrocardiogram traces. Their methodology did not require any de-noising, filtering, and preprocessing methods. The networks distinguished a signal as NSR and AF. They used the publicly available MIT-BIH PhysioNet database. The experimental results demonstrate that the achieved accuracy by RNN, LSTM, and GRU is 95.0%, 100%, and 100%, respectively. Results were encouraging enough to use clinical trials for the real-time AF classification.

9. Discussion

In the previous sections, we present the use of different DL methods in arrhythmia classification. In this section, we not only compare our achievements with other surveys, but also present the relevance of a method to specific arrhythmia pattern. In addition, we analyze the computational complexity of different DL methods and the distribution share of each arrhythmia and method statistically. Finally, the current DL limitations and future trends of DL-based arrhythmia classification will be discussed.

9.1. Distinction to the Other Survey Papers

There exist other survey papers that focus on ECG signal feature extraction and classification including (Jambukia et al., 2015), (Dewangan & Shukla, 2015), (Luz et al., 2016), (Bizopoulos & Koutsouris, 2018), and (Dinakarrao et al., 2019). Dewangan et al. (Dewangan & Shukla, 2015) discuss old-fashioned feature extraction techniques such as Hidden Markov Model (HMM), and independent component analysis. (Jambukia et al., 2015) is a short survey papers that mainly focus on machine learning techniques such as SVM and MLP. Luz et al. (Luz et al., 2016) review automatic ECG-based abnormalities classification papers that consider ECG signal preprocessing, heartbeat segmentation, feature description and learning algorithms. Bizopoulos et al. (Bizopoulos & Koutsouris, 2018) survey deep learning papers used imaging modalities and signal data from cardiology. Compared to these surveys, we only present state-of-the-art deep learning techniques that provide the highest accuracy results. In addition, our review cover wide topics including arrhythmias medical background, introduction on different deep learning methods, performance evaluation metrics, popular databases of ECG records, and discussion on computational complexity and limitations of deep learning methods used for ECG arrhythmia classification. Contemporary to our review, Diankaro et al. (Dinakarrao et al., 2019) presents a comprehensive survey on arrhythmia diagnosis. They analyzed a wide number of techniques for arrhythmia detection, plus, their present performance and involved complexities with these techniques. Compared to (Dinakarrao et al., 2019), we only focus on deep learning based techniques to consider more related papers. In addition, we consider a broader range of arrhythmia such as PAC, PVC, Ectopic Beat, and MI. To the best of our knowledge, this is the first review paper covering all the popular ECG arrhythmia and analyzed performance and characteristics of DL-based arrhythmia detection methods as well as the variants of these methods.

9.2. *The Methodological Comparison*

In this section, we present an overall comparison on the share of each method for arrhythmia classification, and the percentage of each arrhythmia regarding the total studied papers.

We study applicability of six major DL methods on ECG arrhythmia classification including CNN, MLP, RNN, LSTM, DBN, and GRU. The percentage of association of each model in the studied papers is illustrated in Fig. 10.a. Unequivocally, CNN is the most favorable method for feature extraction (with 52% contribution). Fig. 10.b shows the percentage of heart diseases which have been considered in studied papers. Classifying AF, and SVEB/VEB are the most considered arrhythmia with 48% and 21% contribution, respectively. Besides, Fig. 11 summarizes studied arrhythmia based on the most reported performance (accuracy) concerning the classification methods for all the studied arrhythmia.

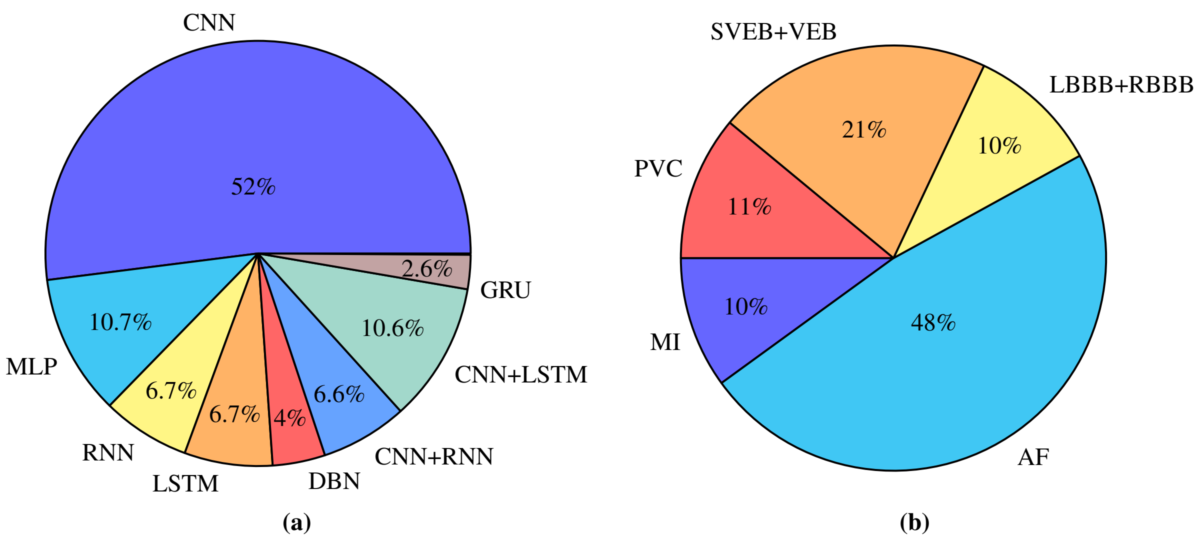


Figure 10: (a) The percentage of contributing each DL model in the studied articles. (b) The percentage of each heart diseases considered in the studied articles. The category of each pie in the graph is specified in Table 5.

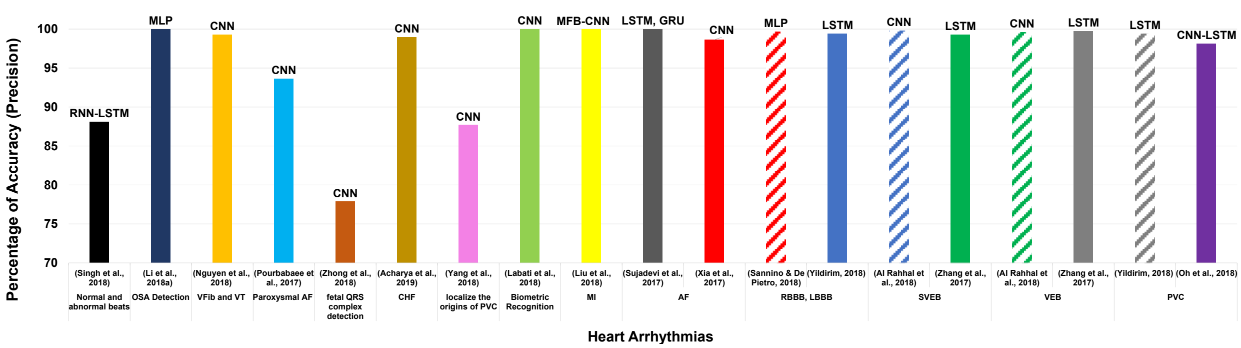


Figure 11: Reporting the best accuracy for each studied arrhythmia regarding the classification method.

9.3. *Relevance of the DL Methods*

This study presented the DL methods commonly employed for detecting deflection of a ECG waves from its normal range. A common feature of all the presented DL methods is their capabilities in preserving temporal variation of the signal, that is regarded as a necessity for arrhythmia classification. It is important to note that the variations can occur both within the beats and over the beats. This necessitates a capability to learn both short term and long term learning for an efficient classification. As can be seen in Fig. 9, the *QRS* complex is noticeably changed for cases with RBBB or LBBB, and a dynamic classifier like LSTM can learn and classify such the variations of *QRS* complex. For cases with premature atrial or ventricular contraction such the variations occur on certain beats, and therefore a dynamic classier which is capable to preserve long term memories can be of interest, as confirmed by the review results. Ectopic beats on the other side entail deflection of *P*-Wave form the sinus form, and hence a dynamic method with capability to preserve the short memory can provide a quick learning for the classification.

9.4. Computational Complexity of the DL Methods

In general, the processing complexity of DL methods is dependent on the number of required floating-point operations for processing the model. Such that there exist an strong correlation between floating-point operations of a CNN model and the the model inference time ($R^2 = 0.8888, p-value < 0.0015$) and the model energy consumption ($R^2 = 0.9641, p-value < 0.0001$) (Loni et al., 2019). The actual inference time of a DL method is dependent on various parameters including: hardware platform, compiler optimization, and the utilized APIs for implementing the model (e.g. TensorFlow (Abadi et al., 2016), PyTorch (Paszke et al., 2019), etc). Therefore, we present the computational overhead of various DL methods in an abstract way summarized in Table 1. they need huge computing resource for real-time processing (Loni et al., 2020). In general, DL methods are slower than other machine learning-based techniques such as SVM (Dinakarrao et al., 2019).

Table 1: The computational complexity of different DLs Methods.

Deep Learning Method	Computational Complexity
MLP	medium-complexity
CNN	high-complexity
DBN	low-complexity
RNN	medium-complexity
LSTM	medium-complexity
GRU	low-complexity

9.5. Limitations of the DL Methods

Despite the success of DL methods in improving the classification performance compared to traditional machine learning methods, thy have limitations. In this section, we list the major limitation of DL methods involved in arrhythmia classification.

1. For smaller amount of training data, DL methods face the overfitting problem since the model highly pay attention to training data and do not generalize well for the test data. Thus, shallow techniques provide better performance on small amount of data samples.
2. Most of the DL methods are disposed to learn the peculiarities such as the noise of ECG signal leading to inaccurate results. Th problem is pronounced with the size of dataset.
3. In general, DNNs are computational intensive processing methods with huge memory footprint (Loni et al., 2019) which their implementation is challenging on low-power embedded devices. Hence, DNN-based arrhythmia classification are primarily deployed on software on CPU and/or GPUs that is not a real-time solution. Therefore, existing hardware implementations of DNN are huge to be deployed on the energy-constraint wearable devices.
4. Gradient of the complex models hardly converge to the optimal loss function due to the vanishing gradient problem. Therefore, carelessly increasing DNN layers in order to achieve higher classification accuracy is not necessarily gain benefit.
5. According to the studied papers analysis, the proposed DL methods are effective for limited number of arrhythmia classes (e.g. roughly six classes). Generating a complex model for classifying all the ECG arrhythmia are not proven to be effective due to difficulty of training model and needed resources.
6. Most of the studied papers focused on ECG signal characteristics, however, other important characteristics such as patients' physical state (e.g. age, gender, physical conditions, lifestyle, etc) are still excluded in the community.

9.6. *Future Research Trend*

According to the best classification methods represented in Fig. 11, CNN-based have proven to be effective for arrhythmia classification. Recent trend of research in this scope shows that dynamic classification methods that are capable to learn both short and long term contents of the signal in an efficient way, would be employed for such applications. CNN has shown excellent performance in classifying different types of arrhythmia. This powerful method would be one of the most efficient learning tool for this application.

10. Conclusions

The study presented results of a review on different methods for classifying arrhythmia on ECG signals. The objective of the review method was to investigate the most powerful Deep Learning methods detecting various types of arrhythmia. Technical details of the common methods were discussed. The GRU/LSTM, CNN, and LSTM, showed outstanding results for correct classification of Atrial Fibrillation, Supraventricular Ectopic Beats, and Ventricular Ectopic Beats, respectively. It is also concluded that the use of a proper type of Deep Learning method can considerably improve the classification performance for the corresponding application.

11. Acknowledgment

The authors would like to thank M.D. Fateme Poursalim for providing helpful tips to realize the medical background. This work was supported by the Swedish Knowledge Foundation (KKS) within the DeepMaker project and also HCST-Develop.

References

A Gharehbaghi, A. S. A. K. M. L., T Dutoit (2015). A novel method for screening children with isolated bicuspid aortic valve. *Cardiovascular engineering and technology*, *6*, 546–556.

Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., Isard, M. et al. (2016). Tensorflow: A system for large-scale machine learning. In *12th {USENIX} Symposium on Operating Systems Design and Implementation ({OSDI} 16)* (pp. 265–283).

Acharya, U. R., Fujita, H., Lih, O. S., Hagiwara, Y., Tan, J. H., & Adam, M. (2017a). Automated detection of arrhythmias using different intervals of tachycardia ecg segments with convolutional neural network. *Information sciences*, *405*, 81–90.

Acharya, U. R., Fujita, H., Oh, S. L., Hagiwara, Y., Tan, J. H., & Adam, M. (2017b). Application of deep convolutional neural network for automated detection of myocardial infarction using ecg signals. *Information Sciences*, *415*, 190–198.

Acharya, U. R., Fujita, H., Oh, S. L., Hagiwara, Y., Tan, J. H., Adam, M., & San Tan, R. (2019). Deep convolutional neural network for the automated diagnosis of congestive heart failure using ecg signals. *Applied Intelligence*, *49*, 16–27.

Acharya, U. R., Fujita, H., Oh, S. L., Raghavendra, U., Tan, J. H., Adam, M., Gertych, A., & Hagiwara, Y. (2018). Automated identification of shockable and non-shockable life-threatening ventricular arrhythmias using convolutional neural network. *Future Generation Computer Systems*, *79*, 952–959.

Acharya, U. R., Oh, S. L., Hagiwara, Y., Tan, J. H., Adam, M., Gertych, A., & San Tan, R. (2017c). A deep convolutional neural network model to classify heartbeats. *Computers in biology and medicine*, *89*, 389–396.

Adams, E. R., & Choi, A. (2012). Using neural networks to predict cardiac arrhythmias. In *2012 IEEE International Conference on Systems, Man, and Cybernetics (SMC)* (pp. 402–407). IEEE.

Al Rahhal, M. M., Bazi, Y., Al Zuair, M., Othman, E., & BenJdira, B. (2018). Convolutional neural networks for electrocardiogram classification. *Journal of Medical and Biological Engineering*, *38*, 1014–1025.

Andersen, R. S., Peimankar, A., & Puthusserypady, S. (2019). A deep learning approach for real-time detection of atrial fibrillation. *Expert Systems with Applications*, *115*, 465–473.

Andreotti, F., Carr, O., Pimentel, M. A., Mahdi, A., & De Vos, M. (2017). Comparing feature-based classifiers and convolutional neural networks to detect arrhythmia from short segments of ecg. In *2017 Computing in Cardiology (CinC)* (pp. 1–4). IEEE.

Awtry, E., Jeon, C., & Ware, M. G. (2006). *Blueprints cardiology*. Lippincott Williams & Wilkins.

Baldzizhar, A., Manuylova, E., Marchenko, R., Kryvalap, Y., & Carey, M. G. (2016). Ventricular Tachycardias: Characteristics and Management. doi:`10.1016/j.cnc.2016.04.004`.

Benjamin, E. J., Virani, S. S., Callaway, C. W., Chamberlain, A. M., Chang, A. R., Cheng, S., Chiuve, S. E., Cushman, M., Delling, F. N., Deo, R. et al. (2018). Heart disease and stroke statistics-2018 update: a report from the american heart association. *Circulation*, *137*, e67.

Bizopoulos, P., & Koutsouris, D. (2018). Deep learning in cardiology. *IEEE reviews in biomedical engineering*, *12*, 168–193.

Bonow, R. O., Mann, D. L., Zipes, D. P., & Libby, P. (2011). *Braunwald's heart disease e-book: a textbook of cardiovascular medicine*. Elsevier Health Sciences.

Ceylan, R., & Özbay, Y. (2007). Comparison of fcm, pca and wt techniques for classification ecg arrhythmias using artificial neural network. *Expert Systems with Applications*, *33*, 286–295.

Chalvidan, T., Deharo, J., & Djiane, P. (2000). Pacemaker syndromes. In *Annales de cardiologie et d'angeiologie* (pp. 224–229). volume 49.

Chamatidis, I., Katsika, A., & Spathoulas, G. (2017). Using deep learning neural networks for ecg based authentication. In *2017 International Carnahan Conference on Security Technology (ICCST)* (pp. 1–6). IEEE.

Chandra, B., Sastry, C. S., Jana, S., & Patidar, S. (2017). Atrial fibrillation detection using convolutional neural networks. In *2017 Computing in Cardiology (CinC)* (pp. 1–4). IEEE.

Chen, M., Wang, G., Xie, P., Sang, Z., Lv, T., Zhang, P., & Yang, H. (2018). Region aggregation network: Improving convolutional neural network for ecg characteristic detection. In *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (pp. 2559–2562). IEEE.

Chung, J., Gulcehre, C., Cho, K., & Bengio, Y. (2014). Empirical evaluation of gated recurrent neural networks on sequence modeling. *arXiv preprint arXiv:1412.3555*, .

Ciresan, D., Meier, U., & Schmidhuber, J. (2012). Multi-column Deep Neural Networks for Image Classification Supplementary Online Material. *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*, . doi:10.1109/CVPR.2012.6248110. arXiv:1202.2745.

Conover, M. B. (2002). *Understanding electrocardiography*. Elsevier Health Sciences.

Dewangan, N. K., & Shukla, S. (2015). A survey on ecg signal feature extraction and analysis techniques. *International journal of innovative research in electrical, electronics, instrumentation and control engineering*, *3*.

Dinakarrao, S. M. P., Jantsch, A., & Shafique, M. (2019). Computer-aided arrhythmia diagnosis with bio-signal processing: A survey of trends and techniques. *ACM Computing Surveys (CSUR)*, *52*, 23.

Dua, D., & Graff, C. (2017). UCI machine learning repository. URL: http://archive.ics.uci.edu/ml.

Fan, X., Yao, Q., Cai, Y., Miao, F., Sun, F., & Li, Y. (2018). Multiscaled fusion of deep convolutional neural networks for screening atrial fibrillation from single lead short ecg recordings. *IEEE journal of biomedical and health informatics*, *22*, 1744–1753.

Faust, O., Shenfield, A., Kareem, M., San, T. R., Fujita, H., & Acharya, U. R. (2018). Automated detection of atrial fibrillation using long short-term memory network with rr interval signals. *Computers in biology and medicine*, *102*, 327–335.

Frielingsdorf, J., Gerber, A., & Hess, O. (1994). Importance of maintained atrio-ventricular synchrony in patients with pacemarkers. *European heart journal*, *15*, 1431–1440.

Gharehbaghi, A., & Babic, A. (2018). Structural risk evaluation of a deep neural network and a markov model in extracting medical information from phonocardiography. *Studies in health technology and informatics*, *251*, 157—160. URL: http://europepmc.org/abstract/MED/29968626.

Gharehbaghi, A., Babic, A., & Sepehri, A. A. (2019a). Extraction of diagnostic information from phonocardiographic signal using time-growing neural network. In L. Lhotska, L. Sukupova, I. Lacković, & G. S. Ibbott (Eds.), *World Congress on Medical Physics and Biomedical Engineering 2018* (pp. 849–853). Singapore: Springer Singapore.

Gharehbaghi, A., Dutoit, T., Ask, P., & Sörnmo, L. (2014). Detection of systolic ejection click using time growing neural network. *Medical Engineering and Physics*, *36*, 477–483.

Gharehbaghi, A., Lindén, M., & Babic, A. (2019b). An artificial intelligent-based model for detecting systolic pathological patterns of phonocardiogram based on time-growing neural network. *Applied Soft Computing*, *83*, 105615. URL: http://www.sciencedirect.com/science/article/pii/S1568494619303953. doi:https://doi.org/10.1016/j.asoc.2019.105615.

Gharehbaghi, A., & Lindén, M. (2018). A deep machine learning method for classifying cyclic time series of biological signals using time-growing neural network. *IEEE Transactions on Neural Networks and Learning Systems*, *29*, 4102–4115. doi:10.1109/TNNLS.2017.2754294.

Ghiasi, S., Abdollahpur, M., Madani, N., Kiani, K., & Ghaffari, A. (2017). Atrial fibrillation detection using feature based algorithm and deep convolutional neural network. In *2017 Computing in Cardiology (CinC)* (pp. 1–4). IEEE.

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. (2000). Physiobank, physiotoolkit, and physionet: components of a new research resource for complex physiologic signals. *Circulation*, *101*, e215–e220.

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT press.

He, K., Zhang, X., Ren, S., & Sun, J. (2015). Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, . doi:10.1109/TPAMI.2015.2389824.

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 770–778).

Hinton, G. E., Osindero, S., & Teh, Y.-W. (2006). A fast learning algorithm for deep belief nets. *Neural computation*, *18*, 1527–1554.

Hinton, G. E., Sejnowski, T. J. et al. (1986). Learning and relearning in boltzmann machines. *Parallel distributed processing: Explorations in the microstructure of cognition*, *1*, 2.

Huff, J. (2006). *ECG workout: Exercises in arrhythmia interpretation*. Lippincott Williams & Wilkins.

Isin, A., & Ozdalili, S. (2017). Cardiac arrhythmia detection using deep learning. *Procedia computer science*, *120*, 268–275.

Jambukia, S. H., Dabhi, V. K., & Prajapati, H. B. (2015). Classification of ecg signals using machine learning techniques: A survey. In *2015 International Conference on Advances in Computer Engineering and Applications* (pp. 714–721). IEEE.

Ji, J., Chen, X., Luo, C., & Li, P. (2018). A deep multi-task learning approach for ecg data analysis. In *2018 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI)* (pp. 124–127). IEEE.

Joshi, A. J., Chandran, S., Jayaraman, V. K., & Kulkarni, B. D. (2009). Hybrid svm for multiclass arrhythmia classification. In *2009 IEEE International Conference on Bioinformatics and Biomedicine* (pp. 287–290). IEEE.

Jun, T. J., Nguyen, H. M., Kang, D., Kim, D., Kim, D., & Kim, Y.-H. (2018). Ecg arrhythmia classification using a 2-d convolutional neural network. *arXiv preprint arXiv:1804.06812*, .

Kamaleswaran, R., Mahajan, R., & Akbilgic, O. (2018). A robust deep convolutional neural network for the classification of abnormal cardiac rhythm using single lead electrocardiograms of variable length. *Physiological measurement*, *39*, 035006.

Kasper, D. L., Fauci, A. S., Hauser, S. L., Longo, D. L., Jameson, J. L., & Loscalzo, J. (2018). *Harrison's Principles of Internal Medicine, (Vol. 1 & Vol. 2)*. McGraw Hill Professional.

Keyvanrad, M. A., & Homayounpour, M. M. (2015). Deep belief network training improvement using elite samples minimizing free energy. *International Journal of Pattern Recognition and Artificial Intelligence*, *29*, 1551006.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. In *Advances in Neural Information Processing Systems*.

Labati, R. D., Muñoz, E., Piuri, V., Sassi, R., & Scotti, F. (2018). Deep-ecg: Convolutional neural networks for ecg biometric recognition. *Pattern Recognition Letters*, .

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *nature*, *521*, 436.

LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, . doi:10.1109/5.726791.

Li, D., Zhang, J., Zhang, Q., & Wei, X. (2017). Classification of ecg signals based on 1d convolution neural network. In *2017 IEEE 19th International Conference on e-Health Networking, Applications and Services (Healthcom)* (pp. 1–6). IEEE.

Li, K., Pan, W., Li, Y., Jiang, Q., & Liu, G. (2018a). A method to detect sleep apnea based on deep neural network and hidden markov model using single-lead ecg signal. *Neurocomputing*, *294*, 94–101.

Li, Y., Pang, Y., Wang, J., & Li, X. (2018b). Patient-specific ecg classification by deeper cnn from generic to dedicated. *Neurocomputing*, *314*, 336–346.

Limam, M., & Precioso, F. (2017). Atrial fibrillation detection and ecg classification based on convolutional recurrent neural network. In *2017 Computing in Cardiology (CinC)* (pp. 1–4). IEEE.

Link, M. S. (2012). Evaluation and initial treatment of supraventricular tachycardia. *New England Journal of Medicine*, . doi:10.1056/NEJMcp1111259.

Liu, M., & Kim, Y. (2018). Classification of heart diseases based on ecg signals using long short-term memory. In *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (pp. 2707–2710). IEEE.

Liu, W., Huang, Q., Chang, S., Wang, H., & He, J. (2018). Multiple-feature-branch convolutional neural network for myocardial infarction diagnosis using electrocardiogram. *Biomedical Signal Processing and Control*, *45*, 22–32.

Liu, W., Zhang, M., Zhang, Y., Liao, Y., Huang, Q., Chang, S., Wang, H., & He, J. (2017). Real-time multilead convolutional neural network for myocardial infarction detection. *IEEE journal of biomedical and health informatics*, *22*, 1434–1444.

Loni, M., Sinaei, S., Zoljodi, A., Daneshtalab, M., & Sjödin, M. (2020). Deepmaker: A multi-objective optimization framework for deep neural networks in embedded systems. *Microprocessors and Microsystems*, (p. 102989).

Loni, M., Zoljodi, A., Sinaei, S., Daneshtalab, M., & Sjödin, M. (2019). Neuropower: Designing energy efficient convolutional neural network architecture for embedded systems. In *International Conference on Artificial Neural Networks* (pp. 208–222). Springer.

Lui, H. W., & Chow, K. L. (2018). Multiclass classification of myocardial infarction with convolutional and recurrent neural networks for portable ecg devices. *Informatics in Medicine Unlocked*, *13*, 26–33.

Luo, K., Li, J., Wang, Z., & Cuschieri, A. (2017). Patient-specific deep architectural model for ecg classification. *Journal of healthcare engineering*, *2017*.

Luz, E. J. d. S., Schwartz, W. R., Cámara-Chávez, G., & Menotti, D. (2016). Ecg-based heartbeat classification for arrhythmia detection: A survey. *Computer methods and programs in biomedicine*, *127*, 144–164.

Majumdar, A., & Ward, R. (2017). Robust greedy deep dictionary learning for ecg arrhythmia classification. In *2017 International Joint Conference on Neural Networks (IJCNN)* (pp. 4400–4407). IEEE.

Maknickas, V., & Maknickas, A. (2017). Atrial fibrillation classification using qrs complex features and lstm. In *2017 Computing in Cardiology (CinC)* (pp. 1–4). IEEE.

Mathews, S. M., Kambhamettu, C., & Barner, K. E. (2018). A novel application of deep learning for single-lead ecg classification. *Computers in biology and medicine*, *99*, 53–62.

Nguyen, M. T., Van Nguyen, B., & Kim, K. (2018). Deep feature learning for sudden cardiac arrest detection in automated external defibrillators. *Scientific reports*, *8*, 1–12.

Oh, S. L., Ng, E. Y., San Tan, R., & Acharya, U. R. (2018). Automated diagnosis of arrhythmia using combination of cnn and lstm techniques with variable length heart beats. *Computers in biology and medicine*, *102*, 278–287.

Otten, E. J. (2005). Cecil textbook of medicine, 22nd edition. *The Journal of Emergency Medicine*, . doi:10.1016/j.jemermed.2004.10.006.

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L. et al. (2019). Pytorch: An imperative style, high-performance deep learning library. In *Advances in Neural Information Processing Systems* (pp. 8024–8035).

Pater, C. (2005). Methodological considerations in the design of trials for safety assessment of new drugs and chemical entities. *Current controlled trials in cardiovascular medicine*, *6*, 1.

Plesinger, F., Nejedly, P., Viscor, I., Halamek, J., & Jurak, P. (2017). Automatic detection of atrial fibrillation and other arrhythmias in holter ecg recordings using rhythm features and neural networks. In *2017 Computing in Cardiology (CinC)* (pp. 1–4). IEEE.

Poh, M.-Z., Poh, Y. C., Chan, P.-H., Wong, C.-K., Pun, L., Leung, W. W.-C., Wong, Y.-F., Wong, M. M.-Y., Chu, D. W.-S., & Siu, C.-W. (2018). Diagnostic assessment of a deep learning system for detecting atrial fibrillation in pulse waveforms. *Heart*, *104*, 1921–1928.

Pourbabaee, B., Roshtkhari, M. J., & Khorasani, K. (2017). Deep convolutional neural networks and learning ecg features for screening paroxysmal atrial fibrillation patients. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, *48*, 2095–2104.

Qiao, J., Wang, G., Li, W., & Li, X. (2018). A deep belief network with plsr for nonlinear system modeling. *Neural Networks*, *104*, 68–79.

University of Rocher Medical Center, T., & Warehouse, H. E. (). Intercity digital electrocardiogram alliance—ideal database (e-hol-03-0202-003). URL: http://thew-project.org/Database/E-HOL-03-0202-003.html.

Rubin, J., Parvaneh, S., Rahman, A., Conroy, B., & Babaeizadeh, S. (2018). Densely connected convolutional networks for detection of atrial fibrillation from short single-lead ecg recordings. *Journal of electrocardiology*, *51*, S18–S21.

Sadrawi, M., Lin, C.-H., Lin, Y.-T., Hsieh, Y., Kuo, C.-C., Chien, J. C., Haraikawa, K., Abbod, M. F., & Shieh, J.-S. (2017). Arrhythmia evaluation in wearable ecg devices. *Sensors*, *17*, 2445.

Sannino, G., & De Pietro, G. (2018). A deep learning approach for ecg-based heartbeat classification for arrhythmia detection. *Future Generation Computer Systems*, *86*, 446–455.

Savalia, S., & Emamian, V. (2018). Cardiac arrhythmia classification by multi-layer perceptron and convolution neural networks. *Bioengineering*, *5*, 35.

Sawhney, N. S., Anousheh, R., Chen, W. C., & Feld, G. K. (2009). Diagnosis and Management of Typical Atrial Flutter. doi:10.1016/j.ccl.2008.09.010.

Sayantan, G., Kien, P., & Kadambari, K. (2018). Classification of ecg beats using deep belief network and active learning. *Medical & Biological Engineering & Computing*, *56*, 1887–1898.

Schuster, M., & Paliwal, K. K. (1997). Bidirectional recurrent neural networks. *IEEE Transactions on Signal Processing*, *45*, 2673–2681.

Schwab, P., Scebba, G. C., Zhang, J., Delai, M., & Karlen, W. (2017). Beat by beat: Classifying cardiac arrhythmias with recurrent neural networks. In *2017 Computing in Cardiology (CinC)* (pp. 1–4). IEEE.

Shashikumar, S. P., Shah, A. J., Clifford, G. D., & Nemati, S. (2018). 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* (pp. 715–723). ACM.

Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, .

Singh, S., Pandey, S. K., Pawar, U., & Janghel, R. R. (2018). Classification of ecg arrhythmia using recurrent neural networks. *Procedia computer science*, *132*, 1290–1297.

Sodmann, P., Vollmer, M., Nath, N., & Kaderali, L. (2018). A convolutional neural network for ecg annotation as the basis for classification of cardiac rhythms. *Physiological measurement*, *39*, 104005.

Sugimoto, K., Lee, S., & Okada, Y. (2018). Deep learning-based detection of periodic abnormal waves in ecg data. In *Proceedings of the International MultiConference of Engineers and Computer Scientists*. volume 1.

Sujadevi, V., Soman, K., & Vinayakumar, R. (2017). Real-time detection of atrial fibrillation from short time single lead ecg traces using recurrent neural networks. In *The International Symposium on Intelligent Systems Technologies and Applications* (pp. 212–221). Springer.

Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., & Rabinovich, A. (2015). Going deeper with convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1–9).

Taherisadr, M., Asnani, P., Galster, S., & Dehzangi, O. (2018). Ecg-based driver inattention identification during naturalistic driving using mel-frequency cepstrum 2-d transform and convolutional neural networks. *Smart Health*, *9*, 50–61.

Taji, B., Chan, A. D., & Shirmohammadi, S. (2018). False Alarm Reduction in Atrial Fibrillation Detection Using Deep Belief Networks. *IEEE Transactions on Instrumentation and Measurement*, . doi:`10.1109/TIM.2017.2769198`.

Takalo-Mattila, J., Kiljander, J., & Soininen, J.-P. (2018). Inter-patient ecg classification using deep convolutional neural networks. In *2018 21st Euromicro Conference on Digital System Design (DSD)* (pp. 421–425). IEEE.

Tan, J. H., Hagiwara, Y., Pang, W., Lim, I., Oh, S. L., Adam, M., San Tan, R., Chen, M., & Acharya, U. R. (2018). Application of stacked convolutional and long short-term memory network for accurate identification of cad ecg signals. *Computers in biology and medicine*, *94*, 19–26.

Tang, H., Wang, T., Li, M., & Yang, X. (2018). The design and implementation of cardiotocography signals classification algorithm based on neural network. *Computational and mathematical methods in medicine*, *2018*.

Thornton, S., & Hochachka, P. (2004). Oxygen and the diving seal., .

TRIAL, S.-H. (2011). Stitch trial. *N Engl J Med*, *364*, 1607–16.

Wang, G., Zhang, C., Liu, Y., Yang, H., Fu, D., Wang, H., & Zhang, P. (2019). A global and updatable ecg beat classification system based on recurrent neural networks and active learning. *Information Sciences*, *501*, 523–542.

Wang, L., & Zhou, X. (2019). Detection of congestive heart failure based on lstm-based deep network via short-term rr intervals. *Sensors*, *19*, 1502.

Warrick, P., & Homsi, M. N. (2017). Cardiac arrhythmia detection from ecg combining convolutional and long short-term memory networks. In *2017 Computing in Cardiology (CinC)* (pp. 1–4). IEEE.

Weiler, N. C., Collman, F., Vogelstein, J. T., Burns, R., & Smith, S. J. (2014). Types of Arrhythmia - NHLBI, NIH. doi:10.1038/sdata.2014.46.

Wilkins, E., Wilson, L., Wickramasinghe, K., Bhatnagar, P., Leal, J., Luengo-Fernandez, R., Burns, R., Rayner, M., & Townsend, N. (2017). European cardiovascular disease statistics 2017, .

Xia, Y., Wulan, N., Wang, K., & Zhang, H. (2017). Atrial fibrillation detection using stationary wavelet transform and deep learning. In *2017 Computing in Cardiology (CinC)* (pp. 1–4). IEEE.

Xia, Y., Wulan, N., Wang, K., & Zhang, H. (2018). Detecting atrial fibrillation by deep convolutional neural networks. *Computers in biology and medicine*, *93*, 84–92.

Xiang, Y., Lin, Z., & Meng, J. (2018). Automatic qrs complex detection using two-level convolutional neural network. *Biomedical engineering online*, *17*, 13.

Xie, P., Wang, G., Zhang, C., Chen, M., Yang, H., Lv, T., Sang, Z., & Zhang, P. (2018). Bidirectional recurrent neural network and convolutional neural network (bircnn) for ecg beat classification. In *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (pp. 2555–2558). IEEE.

Xiong, Z., Stiles, M. K., & Zhao, J. (2017). Robust ecg signal classification for detection of atrial fibrillation using a novel neural network. In *2017 Computing in Cardiology (CinC)* (pp. 1–4). IEEE.

shensheng Xu, S., Mak, M.-W., & Cheung, C.-C. (2017). Deep neural networks versus support vector machines for ecg arrhythmia classification. In *2017 IEEE International Conference on Multimedia & Expo Workshops (ICMEW)* (pp. 127–132). IEEE.

Xu, X., Wei, S., Ma, C., Luo, K., Zhang, L., & Liu, C. (2018). Atrial fibrillation beat identification using the combination of modified frequency slice wavelet transform and convolutional neural networks. *Journal of healthcare engineering*, *2018*.

Yang, T., Yu, L., Jin, Q., Wu, L., & He, B. (2018). Localization of Origins of Premature Ventricular Contraction by Means of Convolutional Neural Network from 12-Lead ECG. *IEEE Transactions on Biomedical Engineering*, . doi:10.1109/TBME.2017.2756869.

Yao, Q., Fan, X., Cai, Y., Wang, R., Yin, L., & Li, Y. (2018). Time-incremental convolutional neural network for arrhythmia detection in varied-length electrocardiogram. In *2018 IEEE 16th Intl Conf on Dependable, Autonomic and Secure Computing, 16th Intl Conf on Pervasive Intelligence and Computing, 4th Intl Conf on Big Data Intelligence and Computing and Cyber Science and Technology Congress (DASC/PiCom/DataCom/CyberSciTech)* (pp. 754–761). IEEE.

Yildirim, Ö. (2018). A novel wavelet sequence based on deep bidirectional lstm network model for ecg signal classification. *Computers in biology and medicine*, *96*, 189–202.

Yildirim, Ö., Özal, Pławiak, P., Tan, R.-S., & Acharya, U. R. (2018). Arrhythmia detection using deep convolutional neural network with long duration ecg signals. *Computers in biology and medicine*, *102*, 411–420.

Yosinski, J., Clune, J., Bengio, Y., & Lipson, H. (2014). How transferable are features in deep neural networks? In *Advances in Neural Information Processing Systems*.

Zhai, X., & Tin, C. (2018). Automated ecg classification using dual heartbeat coupling based on convolutional neural network. *IEEE Access*, *6*, 27465–27472.

Zhang, C., Wang, G., Zhao, J., Gao, P., Lin, J., & Yang, H. (2017). Patient-specific ecg classification based on recurrent neural networks and clustering technique. In *2017 13th IASTED International Conference on Biomedical Engineering (BioMed)* (pp. 63–67). IEEE.

Zhao, Z., Zhang, Y., Deng, Y., & Zhang, X. (2018). Ecg authentication system design incorporating a convolutional neural network and generalized s-transformation. *Computers in biology and medicine*, *102*, 168–179.

Zhong, W., Liao, L., Guo, X., & Wang, G. (2018). A deep learning approach for fetal QRS complex detection. *Physiological Measurement*, . doi:10.1088/1361-6579/aab297.

Zihlmann, M., Perekrestenko, D., & Tschannen, M. (2017). Convolutional recurrent neural networks for electrocardiogram classification. In *2017 Computing in Cardiology (CinC)* (pp. 1–4). IEEE.

Appendices

Appendix A

Glossary

AF Atrial Fibrillation. 1, 6, 9, 11–15, 17, 29–37

AFL Atrial Flutter. 8, 9, 30, 33

ANN Artificial Neural Network. 4, 12, 30

APB Atrial Premature Beats. 30, 33

BLSTM Bidirectional LSTM. 37

BRNN Bidirectional Recurrent Neural Network. 5, 6

CADS Computer-Aided Diagnosis. 1, 2, 8

CHF Congestive Heart Failure. 30, 34, 35

CNN Convolutional Neural Network. 1, 3, 4, 11–13, 15–17, 28, 32–35

CVD Cardiovascular Disease. 1

DBN Deep Belief Network. 1, 3–5, 11, 13, 15, 16, 36

DL Deep Learning. 1–3, 11–17

DNN Deep Neural Network. 1–3, 11, 12, 16, 29–31

DWT Discrete Wavelet Transform. 37

ECG Electrocardiogram. 1–3, 6–8, 11–17, 29–37

FCN Fully Connected Neural Network. 11, 32–36

GB-DBN Gaussian-Bernoulli Deep Belief Network. 13, 36

GCNN Generic CNN. 34

GRNN Global Recurrent Neural Network. 13, 36

GRU Gated Recurrent Unit. 1, 6, 11, 14–17, 30, 36, 37

LBBB Left Bundle Branch Block. 6, 7, 9, 13, 16, 30, 33–35, 37

LSTM Long Short-Term Memory. 1, 3, 5, 6, 11, 13–17, 30, 35–37

LV Left Ventricle. 6, 7

MI Myocardial Infarction. 7, 9, 12, 14, 29, 30, 32–34

MLP Multilayer Perceptron. 3, 11, 12, 14–16, 31

NSR Normal Sinus Rhythm. 7, 9, 12–14, 29–37

PAC Premature Atrial Contraction. 7, 9, 14

PAF Paroxysmal Atrial Fibrillation. 29, 30, 34

PVC Premature Ventricular Contraction. 7, 9, 13–15, 30, 33, 35, 37

RBBB Right Bundle Branch Block. 6, 7, 9, 13, 16, 29, 30, 33, 37

RBM Restricted Boltzmann Machine. 4, 13, 36

RNN Recurrent Neural Network. 1, 3–5, 11, 13–16, 30, 36, 37

RV Right Ventricle. 6, 7

SNR Signal-Noise-Ratio. 13

SVEB Supraventricular Ectopic Beats. 1, 11–13, 15, 17, 29–34, 36

SVM Support Vector Machine. 3, 11–14, 16, 31, 34, 36

SVT Atrial or Supraventricular Tachycardia. 8, 9, 30, 33

TDCNN Tuned Dedicated Convolutional Neural Network. 34

VEB Ventricular Ectopic Beats. 1, 11–13, 15, 17, 29–34, 36

VF Ventricular Flutter. 8, 29, 30, 33

VFB Ventricular Fusion Beats. 7

VFib Ventricular Fibrillation. 8, 9, 29–31, 33–35

VT Ventricular Tachycardia. 8, 29, 30, 33–35

WPW Pre-Excitation. 29, 30, 33

Appendix B

Table 2: The architecture of different popular CNNs.

CNN Model	Publish Year	CNN Structure	Achievement
AlexNet (Krizhevsky et al., 2012)	2012	5 convolutional layers + 3 fully-connected layers	An important architecture that attracted many researchers in the field of computer vision.
Clarifai (Yosinski et al., 2014)	2013	5 convolutional layers + 3 fully-connected layers	It was committed to see what's happening inside the network.
SPP (He et al., 2015)	2014	5 convolutional layers + 3 fully-connected layers	By providing a spatial pyramid pooling, the size of the images is eliminated.
VGG (Simonyan & Zisserman, 2014)	2014	13-15 convolutional layers + 3 fully-connected layers	Complete evaluation of the network with incremental depth.
GoogLeNet (Szegedy et al., 2015)	2014	21 convolutional layers + 3 fully-connected layers	Increase network depth and width without increasing computational requirements.
ResNet (He et al., 2016)	2015	152 convolutional layers + 3 fully-connected layers	Increase network depth and provide a method to prevent gradient saturation.
Efficient DenseNet (Loni et al., 2020)	2020	121 convolutional layers + 1 fully-connected layers	An inference efficient CNN by optimizing DenseNet architecture.

Appendix C

Appendix D

Appendix E

Appendix F

Appendix G

Table 3: The popular ECG databases.

Database Name	Number of Recordings	Data Sampling Information	Included Disease
PhysioNet/Computing in Cardiology Challenge (Goldberger et al., 2000)	Length: between 30 s and 60 s, total of 12,186 ECGs were used: 8,528 in the public training set and 3,658 in the private hidden test set	Digitized in real-time at 44.1 kHz and 24-bit resolution	NSR AF
The DNNIH Arrhythmia Database (MITDB) (Obtained between 1975 and 1979) (Goldberger et al., 2000)	48 half-hour excerpts of two-channel ambulatory ECG recordings, obtained from 47 subjects. The subjects were 25 men aged 32 to 89 years, and 22 women aged 23 to 89 years, Twenty-three recordings were chosen at random from a set of 4000 24-hour ambulatory ECG recordings collected from a mixed population of inpatients (about 60%) and outpatients (about 40%)	Digitized at 360 samples per second per channel with 11-bit resolution over a 10 mV range	Complex Ventricular, Supraventricular Arrhythmias Conduction Abnormalities
Physikalisch-Technische Bundesanstalt (PTB) (Goldberger et al., 2000)	549 records from 290 subjects (aged 17 to 87, mean 57.2; 209 men, mean age 55.5, and 81 women, mean age 61.6)	Digitized at 1000 samples per second, with 16 bit resolution over a range of ± 16.384 mV. Resolution: 16 bit with 0.5 V/LSB (2000 A/D units per mV)	MI Cardiomyopathy/Heart Failure Bundle Branch Block Dysrhythmia Myocardial hypertrophy Valvular Heart Disease Myocarditis
MIT-BIH Supraventricular Arrhythmia Database (SVDB) (Goldberger et al., 2000)	78 two-lead recordings of approximately 30 minutes	Digitized at 128 Hz	VEB SVEB
PhysioNet, The ECG-ID Database (Goldberger et al., 2000)	310 ECG recordings, obtcitained from 90 persons	20 seconds, digitized at 500 Hz with12-bit resolution over a nominal ±10 mV range	NSR AF
The MIT-BIH Atrial Fibrillation Database (MIT-BIHAF) (Goldberger et al., 2000)	25 long-term ECG recordings of human subjects with atrial fibrillation (mostly paroxysmal)	ECG signals each sampled at 250 samples per second with 12-bit resolution over a range of ±10 millivolts.	NSR AF
Creighton University VT Database (CUDB) (Goldberger et al., 2000)	35 eight-minute ECG recordings of human subjects	Digitized at 250 Hz with 12-bit resolution over a 10 V range (10 mV nominal relative to the unamplified signals). Each record contains 127,232 samples (slightly less than 8.5 minutes).	Sustained VT VF VFib
The MIT-BIH Malignant Ventricular Arrhythmia Database (VFDB) (Goldberger et al., 2000)	22 half-hour ECG recordings	Digitized at 250 Hz	Sustained VT VF VFib
The UCI cardiac arrhythmia (Dua & Graff, 2017)	Number of Instances: 452 Number of Attributes:279	-	NSR Old Inferior MI Sinus Bradycardia RBBB
Long Term ST Database (LTSTDB). (Goldberger et al., 2000)	Contains 86 lengthy ECG recordings of 80 human subjects	Digitized at 250 samples per second with 12-bit resolution over a range of ±10 millivolts.	NSR SVEB VEB
CinC Challenge 2000 Datasets. (Goldberger et al., 2000)	70 records	16 bits per sample, least significant byte first in each pair, 100 samples per second, nominally 200 A/D units per mV	Sleep Apnea NSR
E-HOL-03-0202-003 (Intercity Digital Electrocardiogram Alliance—IDEAL) Database (University of Rocher Medical Center & Warehouse)	202 healthy subjects 24 hour Holter recordings	Sampling Frequency : 200Hz Amplitude Resolution: 10 microV	Healthy ECG signal
The PAF Prediction Challenge Database (Goldberger et al., 2000)	50 record sets come from 48 different subjects	Digitized ECGs (16 bits per sample, 128 samples per signal per second, samples from each channel alternating, nominally 200 A/D units per mV).	PAF
St.-Petersburg Institute of Cardiological Technics 12-lead Arrhythmia Database (NCART) (Goldberger et al., 2000)	75 annotated recordings extracted from 32 Holter records. Each record is 30 minutes long and contains 12 standard leads	Each sampled at 257 Hz, with gains varying from 250 to 1100 analog-to-digital converter units per mV.	Acute MI Prior MI Coronary Artery Disease with Hypertension Sinus Node Dysfunction Supraventricular ectopy WPW AF Bundle Branch Block
Fantasia Database- PhysioBank (Goldberger et al., 2000)	20 young (21 - 34 years old) and 20 elderly (68 - 85 years old) rigorously-screened healthy subjects underwent 120 minutes of continuous supine resting	Digitized at 250 Hz. Each heartbeat was annotated using an automated arrhythmia detection algorithm	Normal Sinus Rhythm while watching a Fantasia movie
The MIT-BIH Normal Sinus Rhythm (NSR) Database (Goldberger et al., 2000)	18 long-term ECG recordings of subjects, 5 men, aged 26 to 45, and 13 women, aged 20 to 50	-	Normal Sinus Rhythm (NSR)
BIDMC PPG and Respiration Dataset (Goldberger et al., 2000)	The 53 recordings within the dataset, each of 8-minute duration	Sampled at 125 Hz	-

Table 4: Categorizing DNNs based on the studied papers.

ANN	Feature Extraction	Classification	Feature Extraction + Classification
CNN	(Tang et al., 2018), (Liu et al., 2018), (Labati et al., 2018), (Takalo-Mattila et al., 2018), (Chen et al., 2018), (Plesinger et al., 2017), (Sodmann et al., 2018)	(Rubin et al., 2018), (Xia et al., 2018), (Zhao et al., 2018), (Taherisadr et al., 2018), (Kamaleswaran et al., 2018), (Zhai & Tin, 2018), (Jun et al., 2018), (Acharya et al., 2017b), (Andreotti et al., 2017), (Acharya et al., 2017a), (Yang et al., 2018), (Xu et al., 2018)	(Al Rahhal et al., 2018), (Yildirim et al., 2018), (Acharya et al., 2019), (Fan et al., 2018), (Zhong et al., 2018), (Savalia & Emamian, 2018), (Li et al., 2018b), (Li et al., 2017), (Chandra et al., 2017), (Acharya et al., 2017c), (Xiang et al., 2018), (Nguyen et al., 2018) (Pourbabaee et al., 2017), (Liu et al., 2017), (Acharya et al., 2018), (Xia et al., 2017), (Xiong et al., 2017), (Isin & Ozdalili, 2017), (Poh et al., 2018)
MLP	(Li et al., 2018a)	(Sannino & De Pietro, 2018), (Ghiasi et al., 2017), (Chamatidis et al., 2017), (Majumdar & Ward, 2017), (Sadrawi et al., 2017), (shensheng Xu et al., 2017), (Luo et al., 2017)	-
RNN		(Wang et al., 2019), (Maknickas & Maknickas, 2017)	(Singh et al., 2018), (Zhang et al., 2017), (Sujadevi et al., 2017)
LSTM		(Yildirim, 2018)	(Singh et al., 2018), (Faust et al., 2018), (Liu & Kim, 2018), (Sujadevi et al., 2017)
DBN	(Sayantan et al., 2018)	(Mathews et al., 2018)	(Taji et al., 2018)
GRU		(Schwab et al., 2017)	(Sujadevi et al., 2017)
CNN & RNN		(Zihlmann et al., 2017), (Andersen et al., 2019), (Ji et al., 2018), (Xie et al., 2018), (Shashikumar et al., 2018)	
CNN & LSTM			(Oh et al., 2018), (Lui & Chow, 2018), (Sugimoto et al., 2018), (Warrick & Homsi, 2017), (Yao et al., 2018), (Limam & Precioso, 2017), (Wang & Zhou, 2019), (Tan et al., 2018)

Table 5: Categorizing studied papers according their focus on diagnosing various heart arrhythmias detected by analyzing ECG.

#	Heart Arrhythmias	Papers
a	Normal Sinus Rhythm (NSR) Left Bundle Branch Block (LBBB) Right Bundle Branch Block (RBBB) Atrial Premature Beats (APB) Premature Ventricular Contraction (PVC)	(Oh et al., 2018), (Yildirim, 2018), (Mathews et al., 2018), (Li et al., 2017), (Isin & Ozdalili, 2017)
b	Supraventricular Ectopic Beats (SVEB) Ventricular Ectopic Beats (VEB)	(Wang et al., 2019), (Sayantan et al., 2018), (Al Rahhal et al., 2018), (Ji et al., 2018), (Xie et al., 2018), (Zhai & Tin, 2018), (Takalo-Mattila et al., 2018), (Li et al., 2018b), (Acharya et al., 2017c), (Majumdar & Ward, 2017), (Sadrawi et al., 2017), (Zhang et al., 2017), (Luo et al., 2017)
c	Atrial Fibrillation (AF)	(Savalia & Emamian, 2018), (Andersen et al., 2019), (Rubin et al., 2018), (Xia et al., 2018), (Zhao et al., 2018), (Faust et al., 2018), (Shashikumar et al., 2018), (Kamaleswaran et al., 2018), (Fan et al., 2018), (Chen et al., 2018), (Andreotti et al., 2017), (Maknickas & Maknickas, 2017), (Limam & Precioso, 2017), (Chandra et al., 2017), (Schwab et al., 2017), (Acharya et al., 2017a), (Pourbabaee et al., 2017), (Taji et al., 2018), (Plesinger et al., 2017), (Poh et al., 2018) (Xia et al., 2017), (Xiong et al., 2017), (Warrick & Homsi, 2017), (Zihlmann et al., 2017), (Sujadevi et al., 2017), (Xu et al., 2018), (Ghiasi et al., 2017), (Sodmann et al., 2018)
d	Myocardial Infarction (MI)	(Liu et al., 2018), (Lui & Chow, 2018), (Sugimoto et al., 2018), (Acharya et al., 2017b), (Liu et al., 2017), (shensheng Xu et al., 2017)
e	Biometric Recognition	(Labati et al., 2018), (Chamatidis et al., 2017)
f	Detecting Distracted and Non-Distracted Drivers	(Taherisadr et al., 2018)
g	Recognition of 8 pattern images (signal pictures)	(Jun et al., 2018)
h	Localize the Origins of Premature Ventricular Contraction (PVC)	(Yang et al., 2018)
i	Normal Sinus Rhythm (NSR) Atrial Premature Beats (APB) Atrial Flutter (AFL) Atrial Fibrillation (AF) Atrial or Supraventricular Tachycardia (SVT) Pre-Excitation (WPW) Premature Ventricular Contraction (PVC) Ventricular Bigeminy Ventricular Flutter (VF) Idioventricular Rhythm Ventricular Tachycardia (VT) Fusion of Ventricular and NSR Left Bundle Branch Block (LBBB) Right Bundle Branch Block (RBBB) Second-Degree Heart block Pacemaker Rhythm	(Yildirim et al., 2018)
j	Congestive Heart Failure (CHF)	(Acharya et al., 2019), (Wang & Zhou, 2019)
k	Fetal QRS complex detection	(Zhong et al., 2018), (Xiang et al., 2018)
l	Paroxysmal Atrial Fibrillation (PAF)	(Pourbabaee et al., 2017)
m	Normal Sinus Rhythm (NSR) Ventricular Fibrillation (VFib) Ventricular Tachycardia (VT)	(Acharya et al., 2018), (Nguyen et al., 2018)
n	Normal Sinus Rhythm (NSR) Supraventricular Ectopic Beats (SVEB) Ventricular Ectopic Beats (VEB) Fusion of Ventricular and and NSR	(Sannino & De Pietro, 2018)
o	OSA Detection	(Li et al., 2018a)
p	Normal and Abnormal Beats (Separation of Regular and Irregular Beats)	(Singh et al., 2018)
q	Sleep Apnea	(Liu & Kim, 2018)

Table 6: Properties of some notable MLP-based ECG arrhythmia classification.

Application	Ref.	Method		Database	Performance
		Feature Extraction	Classification		
Normal Sinus Rhythm (NSR) Supraventricular Ectopic Beats (SVEB) Ventricular Ectopic Beats (VEB) Fusion of Ventricular and and NSR Heartbeats That Cannot be Classified	(Sannino & De Pietro, 2018)	Pre-RR interval Post-RR interval Local average RR interval Global average RR interval	MLP	MITDB	Accuracy: 99,68%
OSA Detection	(Li et al., 2018a)	MLP	SVM MLP	PhysioNet 2000	Accuracy: 100%, Sensitivity: 100%, Specificity: 100%
Abnormality of Heart Rhythm AF	(Ghiasi et al., 2017)	Morphological ECG Characteristics: 1. *RR* intervals histogram 2. Geometric 3. Fractal Dimension 4. Correlation coefficient 5. Variance of *R* peaks	MLP (Softmax Activation)	PhysioNet 2017	Accuracy: Training: 80% Test: 71%
User Authentication	(Chamatidis et al., 2017)	FT DCT DWT	1. KNN 2. MLP 3. Radial Basis Function Network 4. Random Forest 5. DNN	PTB	Accuracy: 1. 81.616% - 86.974% 2. 81.409% - 85.753% 3. 0.233% - 85.873% 4. 83.993% - 88.447% 5. Average accuracy: 80% (Small Database)
Fusion Beat Supraventricular Ectopic Beats (SVEB) Ventricular Ectopic Beats (VEB)	(Majumdar & Ward, 2017)	*QRS Duration* *RR* Interval Amplitude of *P, Q, R, S, T* Points Robust Deep Dictionary Learning - (RDDL is their new approach)	MLP	MITDB	Overall Accuracy: 97.0% 1. Fusion Beat: Sensitivity: 100%, Specificity: 67.2% 2. SVEB: Sensitivity: 16.9%, Specificity: 100% 3. VEB Sensitivity: 90.1%, Specificity:100%
Supraventricular Ectopic Beats (SVEB) Ventricular Ectopic Beats (VEB)	(Luo et al., 2017)	Stacked Denoising Auto-Encoder (SDA)	MLP	MITDB	VEB: Accuracy: 99.1%, Sensitivity: 93.3%, Specificity: 99.5%, Positive Predictive: 93.3% SVEB: Accuracy: 98.8%, Sensitivity: 71.4%, Specificity: 99.8%, Positive predictive: 94.4%
Normal Sinus Rhythm (NSR) Atrial Fibrillation (AF) Supraventricular Ectopic Beats (SVEB) Ventricular Ectopic Beats (VEB) Ventricular Fibrillation (VFib)	(Sadrawi et al., 2017)	FFT	MLP	PhysioNet CUDB MITDB	VEB: Sensitivity: 93.1% False Positive Rate: 0.321% Positive Predictive: 95.65% SVEB: Sensitivity: 79.87 % False Positive Rate: 1.323 % Positive Predictive: 67.14%

Table 7: Properties of some notable CNN-based ECG arrhythmia classification.

Application	Ref.	Method		Database	Performance
		Feature Extraction	Classification		
MI	(Liu et al., 2018)	MFB-CNN	FCN	PTB	Class-based MI detection: Accuracy: 99.95%, Localization: 99.81% Patient-Specific Experiment: Accuracy: 98.79%, Localization: 94.82%
Biometric Recognition	(Labati et al., 2018)	CNN		PTB	Accuracy: 100%
NSR VEB SVEB	(Takalo-Mattila et al., 2018)	CNN		MITDB	Sensitivity, Positive Predictivity, and False Positive Rate: Class (NSR): 92%, 97%, and 23% Class (SVEB): 62%, 56%, and 2% Class (VEB): 89%, 51%, and 6%
NSR AF Alternative Rhythm	(Chen et al., 2018)	CNN+		PysioBank edb	F_1-score: 0.84
NSR AF Other Arrhythmia (OA)	(Plesinger et al., 2017)	Statistical Description of *RR* Intervals and CNN		PhysioNet Challenge 2017	F_1-score: 0.81, NSR F_1: 0.91 AF F_1: 0.80, F_1: 0.74
NSR AF	(Rubin et al., 2018)	FFT	CNN + FCN	PhysioNet challenge2017	Overall F_1:0.82 F_1 for NSR: 0.91 AF: 0.83 Other: 0.72
AF	(Xia et al., 2018)	Short Time Fourier Transform (STFT) Stationary Wavelet Transform (SWT)	CNNs: DeepNet1 (STFT+CNNs) DeepNet2 (SWT+CNNs)	MITDB AFIB	DeepNet1: Sensitivity:98.34% Specificity:98.24% Accuracy: 98.29% DeepNet2: Sensitivity:98.79% Specificity:97.87% Accuracy: 98.63%
NSR AF	(Zhao et al., 2018)	GST and getframe technology (trajectory image)	CNN	Physiobank ECG-IDdb 2017 PhysioNet challenge 2017	NSR (50): Accuracy=96.63%, EER=5.68% AF (50): Accuracy=96.23%, EER=5.96% Noisy (50): Accuracy=96.18%, EER=6.45%
Detecting Distracted/Non-Distracted Drivers	(Taherisadr et al., 2018)	2-D representation of raw ECG: Mel-frequency Cepstrum	CNN	Using customized dataset (10 subjects drive a car)	Accuracy: 95.51%
NSR AF Other Abnormal Rhythms	(Kamaleswaran et al., 2018)	1. 62 features from a combination of descriptive, linear, nonlinear, temporal and spectral statistical methods 2. raw signal	1. MLP 2. CNN	PhysioNet challenge2017	1. Accuracy: 76.79% f_1-score: NSR: 0.84, AF: 0.63, Other Abnormal Rhythm: 0.63 2. Accuracy: 74.84% f_1-score: NSR: 0.84, AF: 0.69, Other Abnormal Rhythm: 0.60
VEB SVEB	(Zhai & Tin, 2018)	2-D Matrix Capturing Morphology of Single Heartbeat	CNN	MITDB	(24 records) VEB: Accuracy: 98.6% Sensitivity: 93.8% Specificity: 99.2% SVEB: Accuracy: 97.6% Sensitivity: 76.8% Specificity: 98.7% 11 records for VEB and 14 records for SVEB: VEB: Accuracy: 99.1% Sensitivity: 96.4%, Specificity: 99.5% SVEB: Accuracy: 97.3% Sensitivity: 85.3% Specificity: 98.0%
Recognition of 8 pattern images (signal pictures)	(Jun et al., 2018)	Transform into 128*128 Gray-scale image	AlexNet VGGNet	MITDB	Accuracy: 99.05%, Sensitivity: 97.85%

Table 7: Properties of some notable CNN-based ECG arrhythmia classification.

Application	Ref.	Method		Database	Performance
		Feature Extraction	Classification		
MI	(Acharya et al., 2017b)	R-peaks detection	CNN	PTB NSR: 10,546 MI: 40,182	With noise: Accuracy: 93.53% Sensitivity:93.71% Specificity:92.83% Without noise: Accuracy: 95.22%, Sensitivity:95.49%, Specificity:94.19%
NSR AF Other Arrhythmia	(Andreotti et al., 2017)	Time Domain Frequency Domain Non-linear HRV metrics Range of signal-quality including the bSQI Morphological features: P-wave power and QT-interval.	Feature-Based Classifier ResNet	PhysioNet challenge 2017	F_1-score for Feature-Based: training set:72.0%, test set :79% F_1-score for CNN: training set: 72.1%, test set: 83% The Method Result: F_1-score: 79%
NSR AF AFL VFib	(Acharya et al., 2017a)	Statistical Description of RR Intervals as RR Deviation (total features =277)	CNN	MITDB AFDB CUDB Signal Length: Two seconds (21709) Five seconds (8683)	2 seconds: Accuracy: 92.50% Sensitivity: 98.09% Specificity: 93.13% 5 seconds: Accuracy: 94.90% Sensitivity: 99.13% Specificity: 81.44%
SVEB VEB	(Al Rahhal et al., 2018)	CWT With CNN	CNN	INCART SVDB MITDB 1. Scenario 1: 11 common testing records for VEB, and 14 testing records for SVEB 2. Scenario 2: 24 common testing records from 200 to 234 3. Scenario 3: all 48 testing records	Evaluation Metrics: {Sensitivity, Positive Predictive, Specificity, Accuracy} MITDB 1. Scenario 1: VEB: (99.9, 99.1, 100.0, 99.3)% SVEB: (99.9, 97.3,100.0, 99.3)% 2. Scenario 2: VEB: (99.7, 98.7, 99.9, 98.8)% SVEB: (99.3, 98.3, 99.4, 98.1)% 3. Scenario 3: VEB: (99.9, 99.0, 100.0, 99.6)% SVEB: (99.8, 95.9, 100.0, 98.3%)% INCART VEB: (99.23,96.5, 99.7, 98.0)% SVEB: (99.82, 89.30, 99.93, 97.5)% SVDB VEB: (99.4, 91.7, 99.8, 96.7)% SVEB:(98.4, 80.2, 99.7, 94.9)%
Normal Sinus Rhythm (NSR) Atrial Premature Beats (APB) Atrial Flutter (AFL) Atrial Fibrillation (AF) Atrial or Supraventricular Tachycardia (SVT) Pre-Excitation (WPW) Premature Ventricular Contraction (PVC) Ventricular Bigeminy Ventricular Flutter (VF) Idioventricular Rhythm Ventricular Tachycardia (VT) Fusion of Ventricular and NSR Left Bundle Branch Block (LBBB) Right Bundle Branch Block (RBBB) Second-Degree Heart block Pacemaker Rhythm	(Yildirim et al., 2018)	CNN + FCN		MITDB	Accuracy: 91.33%

Table 7: Properties of some notable CNN-based ECG arrhythmia classification.

Application	Ref.	Method: Feature Extraction	Method: Classification	Database	Performance
CHF	(Acharya et al., 2019)	CNN + FCN		PhysioBank (BIDMC, Fantasia) MITDB (NSRDB) Set A: CHF: 30,000 NSR: 70,308 (NSRDB, BIDC) Set B: CHF: 30,000 NSR: 110,000 (Fantasia, BIDMC) Set C: CHF: 30,000 NSR: 30,000 (NSRDB, BIDC) Set D: CHF: 30,000 NSR: 30,000 (Fantasia, BIDMC)	Set A: Accuracy: 95.98% Sensitivity: 96.52% Specificity:95.75% Set B: Accuracy: 98.97% Sensitivity: 98.87% Specificity:99.01% Set C: Accuracy: 94.40% Sensitivity: 94.68% Specificity:94.12% Set D: Accuracy: 98.33% Sensitivity: 98.50% Specificity:98.16%
NSR AF Other Abnormal Rhythms	(Fan et al., 2018)	CNN + FCN		PhysioNet Challenge 2017	Recordings 5 Sec: Accuracy: 96.99% Recordings 20 Sec: Accuracy: 98.13%
Fetal QRS complex detection	(Zhong et al., 2018)	CNN + FCN		PhysioNet Challenge 2013	F_1-score: 77.85% Precision: 75.33% Recall: 80.54%
SVEB VEB	(Li et al., 2018b)	GCNN TDCNN		MITDB: D1: 14 records D2: 24 records	D1: Accuracy: 98.6% Sensitivity: 77.2% Specificity: 99.8% Positive Predictivity: 96.6% D2: Accuracy: 98.3 Sensitivity: 68.7% Specificity: 99.8% Positive Predictivity: 94.7%
NSR LBBB LBBB AF	(Li et al., 2017)	5-layer 1-D CNN	FCN	MITDB	Accuracy: 97.5%
NSR AF Others and noise Classes	(Chandra et al., 2017)	CNN		PhysioNet 2000 MITDB	f_1-score: 71% NSR: 86% AF: 73%, Other classes:56%
Namely Non-Ectopic SVEB VEB Fusion Beats Unknown beats	(Acharya et al., 2017c)	CNN + FCN		MITDB	Accuracy: 94.03% Sensitivity:96.71% Specificity:91.54%
PAF	(Pourbabaee et al., 2017)	CNN	1. FCN 1. KNN 3. SVM (linear) 4. SVM (Gaussian) 5. MLP	PAF	1. Precision: 0.936, Recall: 0.764 2. Precision: 0.907, Recall: 0.902 3, Precision: 0.875, Recall: 0.875 4. Precision: 0.929, Recall: 0.862 5. Precision: 0.906, Recall: 0.825
MI	(Liu et al., 2017)	Multilead-CNN (ML-CNN) Sub 2-D convolutional layers and LAP		PTB	Accuracy: 96.0% Sensitivity: 95.40% Specificity:97.37%
NSR VFib VT	(Acharya et al., 2018)	CNN		MITDB MITDB (VFDB) CUDB	Accuracy: 93.18% Sensitivity: 95.32% Specificity: 91.04%
AF	(Xia et al., 2017)	CNN		MITDB	Accuracy: 98.63%, Sensitivity: 98.79%, Specificity: 97.87%

Table 7: Properties of some notable CNN-based ECG arrhythmia classification.

Application	Ref.	Method		Database	Performance
		Feature Extraction	Classification		
NSR AF Other rhythms	(Xiong et al., 2017)	CNN		PhysioNet challenge 2017	Accuracy: NSR: 90%, AF: 82% Other rhythms: 75%
NSR LBBB Paced Beats	(Isin & Ozdalili, 2017)	AlexNet: 1. N-Fc6 output of the 6th FCN Layer 2. N-Fc7 Output of the 7th FCN Layer 3. N-tst 200 Samples of the R-T Intervals	CNN + FCN	MITDB	N-Fc7: Recognition Rate: 98.51%, Accuracy: 92.4% N-Fc6: Recognition Rate: 97.53%, Accuracy: 91.2% N-tst: Recognition Rate: 91.58%, Accuracy: 85%
AF	(Xu et al., 2018)	Wavelet Transform	CNN	MITDB	Train: Accuracy: 81.07%, Sensitivity: 74.96%, Specificity: 86.41%, AUC: 0.88 Test: Accuracy: 84.85%, Sensitivity: 79.05%, Specificity: 89.99%, AUC: 0.92
QRS Detection	(Xiang et al., 2018)	1-D CNN	1-D CNN	MITDB	Sensitivity: 99.77% Positive Predictivity: 99.91% Error Rate: 0.32%
NSR VT AF Ventricular Bigeminy	(Savalia & Emamian, 2018)	CNN	FCN	PhysioBank	Arrhythmia Accuracy: 88% NSR Accuracy: 87%
VFib VT	(Nguyen et al., 2018)	CNN	FCN	CUDB MIT-BIH (VFDB)	Accuracy: 99.26% Sensitivity: 97.07% Specificity: 99.44%
CHF	(Wang & Zhou, 2019)	CNN	LSTM	Database-1: BIDMC-CHF MIT-BIH NSR Fantasi (RR segment length=500)	Accuracy: 99.22% Sensitivity: 99.22% Specificity: 99.72%
AF	(Poh et al., 2018)	CNN	Linear Classifier	Customized Database	Specificity: 99.0%, Sensitivity: 95.2%, Negative Predictive: 99.9%, AUC: 0.997, Positive Predictive: 72.7%
PVC	(Yang et al., 2018)	Epi-Endo CNN Segment CNN	FCN	90 PVC beats from 9 patients with PVCs	Segment CNN Accuracy: 78% Epi-Endo CNN Accuracy: 90%
NSR AF	(Sodmann et al., 2018)	CNN	FCN	PhysioNet MIT-BIH PhysioNet/CinC Challenge 2017	Average F_1 -score for Rhythm Classes: Training Data: 99% Test Data: 89%
Fatal Heart Monitoring	(Tang et al., 2018)	CNN	FCN	Customized Database	Precision:94.71% Recall: 94.68% Accuracy: 94.7%

Appendix H

Table 8: Properties of some notable DBN-based ECG arrhythmia classification.

Application	Ref.	Method		Database	Performance
		Feature Extraction	Classification		
SVEB VEB	(Sayantan et al., 2018)	GB-DBN	SVM	MITDB SVDB	MITDB: SVEB Accuracy: 99.5%, VEB Accuracy: 99.4% SVDB: SVEB Accuracy: 97.5%, VEB: Accuracy: 98.6%
AF	(Taji et al., 2018)	DBN+ RBM		MITDB AFDB	Without Gating (at -20 dB): Precision: 25.5%, Recall: 29.3%, Accuracy: 58.7%, Specificity: 70.5% With Gating: Precision: 65%, Recall: 68.1%, Accuracy: 81%, Specificity: 85%

Appendix I

Table 9: Properties of some notable RNN-based ECG arrhythmia classification.

Application	Ref.	Method		Database	Performance
		Feature Extraction	Classification		
SVEB VEB	(Wang et al., 2019)	Morphological and Premature-or-Escape-Flag (PEF)	GRNN	MITDB SVDB LTSTDB-I (40 Records)	MITDB: Accuracy: 97.4%, Sensitivity:85.7%, Specificity: 98.3% SVDB: Accuracy: 97.2%, Sensitivity :77.2%, Specificity :99.2%
NSR AF	(Maknickas & Maknickas, 2017)	*RR, QQ, SS, PP, TT* Intervals *SQ, PR, QT, ST* Intervals		PhysioNet challenge2017	F_1-score: 0.78
NSR and Abnormal Beats (separation of regular and irregular beats)	(Singh et al., 2018)	1. 3-layer RNN 2. 3-layer RNN-GRU 3. 3-layer RNN 4. 3-layer RNN-GRU 5. 3-layer RNN-LSTM		MITDB	1. Accuracy: 85.4%, Sensitivity: 80.6%, Specificity: 85.7% 2. Accuracy:82.5%, Sensitivity: 78.9%, Specificity: 81.5% 3. Accuracy:85.4%, Sensitivity: 80.6%, Specificity: 85.7% 4. Accuracy:82.5%, Sensitivity: 78.9%, Specificity: 81.5% 5. Accuracy:88.1%, Sensitivity: 92.4%, Specificity: 83.35%
VEB SVEB	(Zhang et al., 2017)	RNN (2 LSTM Layers + 2 FCN)		MITDB: 1. DS1=11 records 2. DS2=24 records 3. DS3=44 records	VEB: 1. DS1: Accuracy: 99.4%, Sensitivity: 97.6%, Specificity: 99.7%, Positive Predictivity: 97.6% 2. DS2: Accuracy: 99.6%, Sensitivity: 97.5%, Specificity: 99.8%, Positive Predictivity: 97.9% 3. DS3: Accuracy: 99.7%, Sensitivity: 97.1%, Specificity: 99.9%, Positive Predictivity: 98.1% SVEB: 1. DS1: Accuracy: 98.7%, Sensitivity: 87.4%, Specificity: 99.4%, Positive Predictivity: 89.4% 2. DS2: Accuracy: 98.9%, Sensitivity: 86.7%, Specificity: 99.5%, Positive Predictivity: 89.0% 3. DS3: Accuracy: 99.3%, Sensitivity: 85.9%, Specificity: 99.7%, Positive Predictivity: 88.7%
NSR AF	(Sujadevi et al., 2017)	RNN		MITDB	Accuracy: 0.95, Precision: 1.00, Recall: 0.889, F-score: 0.941

Appendix J

Appendix K

Table 10: Properties of some notable LSTM-based ECG arrhythmia classification.

Application	Ref.	Method		Database	Performance
		Feature Extraction	Classification		
NSR PVC Paced Beat (PB) LBBB RBBB	(Yildirim, 2018)	DWT	Bidirectional LSTM	MITDB (total records=7376)	Accuracy: 99.39%
NSR and Abnormal Beats (separation of regular and irregular beats)	(Singh et al., 2018)	1. 3-layer RNN 2. 3-layer RNN-GRU 3. 3-layer RNN 4. 3-layer RNN-GRU 5. 3-layer RNN-LSTM		MITDB	1. Accuracy: 85.4%, Sensitivity: 80.6%, Specificity: 85.7% 2. Accuracy: 82.5%, Sensitivity: 78.9%, Specificity: 81.5% 3. Accuracy: 85.4%, Sensitivity: 80.6%, Specificity: 85.7% 4. Accuracy: 82.5%, Sensitivity: 78.9%, Specificity: 81.5% 5. Accuracy:88.1%, Sensitivity: 92.4%, Specificity: 83.35%
AF	(Faust et al., 2018)	Bidirectional LSTM		MITDB	Accuracy: 98.51%, Sensitivity: 98.32%, Specificity: 98.67%, Positive Predictive Accuracy: 98.39%
Sleep Apnea	(Liu & Kim, 2018)	LSTM		CinC Challenge (Apnea)	Accuracy: 98.4%

Table 11: Properties of some notable GRU-based ECG arrhythmia classification.

Application	Ref.	Method		Database	Performance
		Feature Extraction	Classification		
AF	(Schwab et al., 2017)	Time since the last heartbeat (δRR) Relative Wavelet Energy (RWE) Over 5 Frequency Bands Total Wavelet Energy (TWE) R amplitude Q amplitude QRS Duration	GRU and BLSTM	PhysioNet Challenge 2017	Average F_1-score: 0.79 Class-wise F_1 of the NSR: 0.90 AF: 0.79 Other Arrhythmias: 0.68
AF NSR	(Sujadevi et al., 2017)	GRU		MITDB	Accuracy: 1.00 Precision: 1.00, Recall: 1.00 F_1-score: 1.00