A novel application of deep learning for single-lead ECG classification

Article *in* Computers in Biology and Medicine · June 2018

DOI: 10.1016/j.compbiomed.2018.05.013

3 authors, including:

Sherin Mary
University of Delaware
8 PUBLICATIONS **51** CITATIONS
SEE PROFILE

Kenneth Barner
University of Delaware
243 PUBLICATIONS **4,382** CITATIONS
SEE PROFILE

Computers in Biology and Medicine 99 (2018) 53–62

Contents lists available at ScienceDirect

Computers in Biology and Medicine

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

A novel application of deep learning for single-lead ECG classification

Sherin M. Mathews [a,b], Chandra Kambhamettu [c,*], Kenneth E. Barner [b]

[a] Intel, 2821 Mission College, Santa Clara, CA, 95051, USA
[b] Dept. of Electrical and Computer Engineering, University of Delaware, Newark, DE, 19716, USA
[c] Video/Image Modeling and Synthesis (VIMS) Lab, Department of Computer and Information Sciences, University of Delaware, USA

ARTICLE INFO

Keywords:
Heartbeat classification
Restricted Boltzmann machine
Deep belief networks (DBN)
Deep learning
MIT-BIH database
single-lead ECG recognition

ABSTRACT

Detecting and classifying cardiac arrhythmias is critical to the diagnosis of patients with cardiac abnormalities. In this paper, a novel approach based on deep learning methodology is proposed for the classification of single-lead electrocardiogram (ECG) signals. We demonstrate the application of the Restricted Boltzmann Machine (RBM) and deep belief networks (DBN) for ECG classification following detection of ventricular and supraventricular heartbeats using **single-lead ECG**. The effectiveness of this proposed algorithm is illustrated using real ECG signals from the widely-used MIT-BIH database. Simulation results demonstrate that with a suitable choice of parameters, RBM and DBN can achieve high average recognition accuracies of ventricular ectopic beats (93.63%) and of supraventricular ectopic beats (95.57%) at a low sampling rate of **114 Hz**. Experimental results indicate that classifiers built into this deep learning-based framework achieved state-of-the art performance models at **lower sampling rates and simple features** when compared to traditional methods. Further, employing features extracted at a sampling rate of 114 Hz when combined with deep learning provided enough discriminatory power for the classification task. This performance is comparable to that of traditional methods and uses a much lower sampling rate and simpler features. Thus, our proposed deep neural network algorithm demonstrates that deep learning-based methods offer accurate ECG classification and could potentially be extended to other physiological signal classifications, such as those in arterial blood pressure (ABP), nerve conduction (EMG), and heart rate variability (HRV) studies.

1. Introduction

1.1. Motivation of deep learning for ECG classification

Cardiac arrhythmias (abnormal heart rhythms) pose a serious threat to patients recovering from acute myocardial infarction [1]. Some types of arrhythmias are life-threatening, capable of triggering cardiac arrest and sudden death. Therefore, early automatic detection and classification of ECG patterns is critical to diagnosing and treating patients with life-threatening cardiac arrhythmias [2], [3]. An inexpensive and noninvasive technique to detect these disorders is by analyzing electrocardiograms (ECGs) that furnish valuable information on the electrophysiology and functional aspects of the cardiovascular system.

In past decades, computerized recognition of ECGs has become a well-established practice, assisting cardiologists in the task of classifying long-term ECG recordings. Feature extraction methods to discriminate heartbeats have included wave shape functions [4] [5], [6] [7], [8], Hermite functions [9], wavelet-based features [10] [11], frequency-based features [12], ECG morphology [13], hermite polynomials [14], higher order cumulant features [15], statistical features [16] [17], and Karhunen-Loeve expansion of ECG morphology [5]. Methodologies to classify these extracted features have included support vector machines [8] [18], [16], self-organizing maps with learning vector quantization [14], k-th nearest-neighbor rules [19], decision trees [18], artificial neural networks [20], linear discriminants [4] [5], [7], active learning framework [21] and back propagation neural networks [15]. However, these state of-the-art automatic ECG recognition systems often rely on a pattern-matching framework that represents an ECG signal as a sequence of stochastic patterns, so they require complex feature extraction process and high sampling rates and thus burdensome computational times to classify arrhythmias. Consequently, to enable implementation in real time and at reasonable cost, these systems must enlarge their classification criteria by using a set of simple features and a lower sampling rate.

Though several algorithms have focused on automatically classifying heartbeats in ECGs, another drawback is the scalability failure to handle large intra-class variations wherein the robustness of many existing ECG

* Corresponding author.
E-mail addresses: sherin_mathews@mcafee.com, sherinm@udel.edu (S.M. Mathews), chandrak@udel.edu (C. Kambhamettu), barner@udel.edu (K.E. Barner).

https://doi.org/10.1016/j.compbiomed.2018.05.013
Received 15 May 2017; Received in revised form 8 May 2018; Accepted 9 May 2018

classification techniques remains limited. A major limitation of above approaches is that they are highly dependent on the supervised trained dataset and perform poorly while dealing with large amount of unknown ECG records. Furthermore, extracting complex features in the transform domains when combined with dimensionality reduction algorithms significantly upsurge the computational complexity of the overall process. This limits the usage of such complicated feature-extraction frameworks in low power applications such as wearable health monitoring devices or mobile applications. Moreover, the classifier algorithms have not performed well in practice in case of inter-patient variations of the ECG signals, thereby demonstrating a common shortcoming of having an inconsistent performance while classifying a new patients ECG signal. This makes them unreliable to be widely used clinically or in practice, as they tend to have high variations in their accuracy and efficiency for larger databases.

Deep learning [22] [23], [24] [25], [26] [27], [28] [29] [30], (also known as unsupervised feature learning or representation learning) is a new technique that is becoming mainstream in machine learning and pattern recognition. It has been successfully used in object recognition [31] [32], image verification [33], classification [34], and speech recognition [35] [36], [37], [38]. In recent years, deep learning approaches have dramatically improved the accuracy of recognition tools, creating a deep, multi-stage architecture for unsupervised learning and recognition systems. Deep learning networks are implemented using stacked autoencoders and can represent a highly expressive abstraction. Such abstractions can compactly represent a much larger set of functions than shallow networks can. Thus, they offer tremendous representational power that can help reveal unknown feature coherences of input signals, an important capability for learning tasks that involve complicated models. Here, we propose a more accurate and robust approach using deep learning for single-lead ECG classification that generated fewer false alarms.

1.2. *Literature review of existing deep learning work for ECG classification*

The most successful type of deep learning models are restricted Boltzmann machines (RBM) [39], stacked autoencoder (SAE) [40], Convolutional Neural Networks (CNNs) [41] [42], [43] [44], and Deep Belief Networks(DBN) [45], [46]. CNNs as a framework contain many layers that transform their input with convolution filters and has been recently employed in the automated classification of ECG signals. RBMs belong to category of Markov Random Field (MRF) which constitutes an visible layer corresponding to the input layer and a hidden layer corresponding to the latent feature representation. The connections between the nodes are bidirectional, so given an input vector one can obtain the latent feature representation and also vice versa. Training of the individual layers is done in an unsupervised manner and final fine-tuning is performed by adding a linear classifier to the top layer of the DBN and performing a supervised optimization.

The ECG classification community has taken notice of these pivotal developments and a recent work [39] introduced a restricted Boltzmann machine learning algorithm for two-lead heart beat classification wherein the unsupervised learning algorithm of restricted Boltzmann machine helped in mining the large set of unlabeled ECG wave beats in the heart healthcare monitoring applications. Wang and Shang [45] used DBN to automatically extract features from raw unlabeled physiological two-lead ECG data. For the automatic classification ECG signals, one can find the solution proposed in Ref. [46] based on the combination DBN and SVM. In particular, DBN was used for feature learning and the obtained features are fed to SVM for training and classification. Rahlal et al. [40] while also performing ECG classification using deep neural network (DNN) didn't use complex features but used stacked denoising autoencoders (SDAEs) for suitable feature representation from the raw two -lead ECG data. Zubair et al. [43] used CNN with 44 recordings of ECG signals obtained from MIT-BIH database. They extracted R-peak ECG beat patterns for the training of the three-layer CNN. They achieved 92.70% accuracy in detecting the ECG beats into their respective classes (normal, fusion beat, supraventricular ectopic beat, unknown beat, and ventricular ectopic beat).

In Ref. [43], a CNN based classification system was developed which automatically learns a suitable feature representation from two lead ECG data, thereby negating the need of hand-crafted features. However this framework did not involve any QRS wave detection resulting in a large network. In another work [42], segmented ECGs are processed by an eleven-layer convolutional neural network resulting in maximum accuracy of 93.18% using short duration ECG data. Kiranyaz et al. [44] studied the patient-specific ECG monitoring system using three-layer CNN with only R-peak wave. They also attained good accuracy in the detection of supraventricular ectopic beats and ventricular ectopic beats but incorporated only R-peak wave resulting in very high training time.

The drawback of direct usage of deep learning frameworks is the lack of usage of robust features which not only necessitates lot of training data but also made the physiological interpretation of the resultant network is sometimes cumbersome. Another line of work employing deep learning for ECG classification involved developing patient specific deep architecture models. Authors proposed a patient specific heartbeat classification framework using time frequency representation and a DL architectural model in Ref. [47]. Here, Modified Frequency Slice Wavelet Transform (MFSWT) produced the time-frequency image for heartbeat signal and a deep neural network (DNN) classifier was used for classification. Features were automatically abstracted by stacked denoising auto-encoder (SDA) from the transferred timefrequency image and DNN classifier was constructed by an encoder layer of SDA and a softmax layer.

Another paper [44] on similar ideology employed a 1-D convolutional 7 neural networks (CNNs) that fused the feature extraction and classification for two-lead ECG classification. For each patient, an individual and simple CNN was trained by using relatively small common and patient-specific training data, and such patient-specific feature extraction ability did provide an improvement in the classification performance. This did negate the necessity to extract hand-crafted manual features, but since a dedicated CNN is trained for a particular patient, it can solely be used to only classify a specific patients long ECG data stream in an accurate manner making it practically unscalable for large number of users.

Kutlu [48] performed a multi-stage classification system for two lead ECG using considerable number of features(106 ECG waveform features) and a Deep Belief Network (DBN) as classifier to overcome the disadvantage of having a cumbersome network with no physiology interpretation. Though use of features does make the physiological interpretation easier incase of false positives, the training time was extremely large due to large number of feature extracted and use of DBN network. This can be overcomed by introduction of few but robust features thereby providing data augmentation as as presented in our developed framework. Thus inspired by recent progress in the area of deep learning [49] [50], [51], we developed a deep learning framework that includes restricted Boltzmann Machine (RBM) and deep belief networks (DBN) for single lead ECG classification with **simpler features** and **low sampling rate**. This framework of simple features and a low sampling rate yielded competitive ECG classification performance at lower computational cost, making it a highly practical option in a clinical setting. Although deep learning algorithms are statistically motivated approaches, to the best of our knowledge a deep learning framework has not yet been used to perform **single lead ECG classification** tasks.

The remainder of this paper is organized as follows. Section 2 covers the proposed methodology. Here we present the data processing chain (which includes preprocessing, segmentation, and feature extraction). Section 3 briefly describes the proposed deep learning framework. Section 4 details the experiments, and provides, evaluates and discusses their results. Section 5 concludes the paper.

2. Methodology

A pattern recognition system provides a framework that

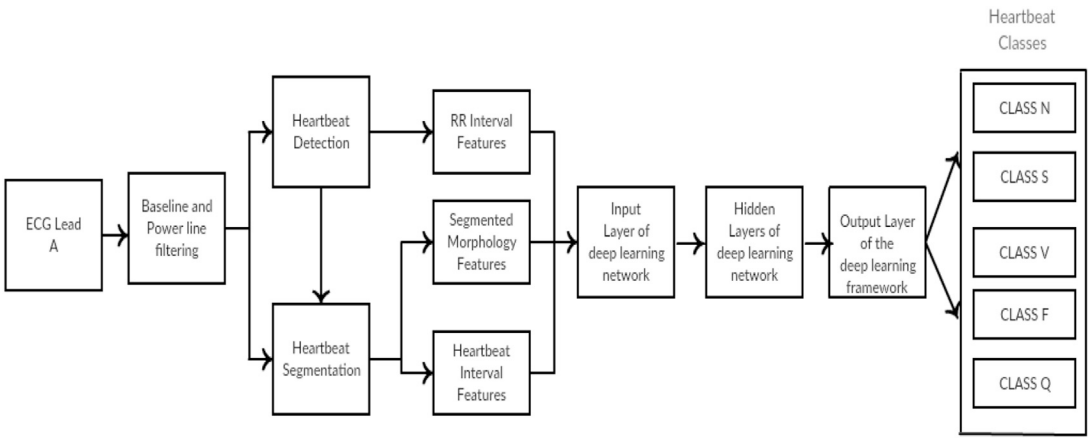

Fig. 1. Block diagram of the proposed methodology.

automatically maps an input signal to a class label by analyzing the features extracted from the signal. The two symbolic stages of this recognition system are feature extraction and classification. Before feature extraction, the data is pre-processed (i.e., filtered), detected and segmented. Then, feature extraction uses mathematical techniques on the input signal to build an association with known models and to obtain the best discriminative representation of the data by exploiting the underlying signal characteristics. Each stage is described below [Fig. 1].

2.1. Preprocesing

Each ECG signal is first bandpass filtered at 0.11 Hz and sampled at 360 Hz. It is preprocessed to remove artifacts, such as baseline wander, power-line interference, high-frequency noise, and motion artifacts [4]. Baseline wander is a low frequency artifact that may be caused by chest-lead ECG signals suffering from coughing or breathing with large chest movements, by poor electrode to skin contact, or by limb-lead ECG signals suffering from arm or leg movements [4]. To remove baseline wander, we pass the signal through median filters with window sizes of 200 ms and 600 ms, thus removing P-waves, QRS complexes, and T-waves. Power line interference is an interfering voltage with frequencies at integral multiples of 50 Hz that can completely obscure an ECG waveform. This strong interference can stem from improper grounding, loose contact of a patient's cable, or disconnected electrodes. Power-line interference and high-frequency noise are removed from a baseline-corrected ECG using a 12-tap low-pass filter, a finite impulse response filter that has 3 dB at 35 Hz and equal ripple in both pass and stop bands [13]. Motion artifacts represent transient baseline interference that is introduced by electrode skin impedance caused by electrode motion. Because the peak amplitude of a motion artifact is 500% of the peak-to-peak ECG amplitude, and its duration is about 100–500 ms, these artifacts can obscure ECG waveforms, making their interpretation quite difficult.

Motion artifacts are removed using an adaptive filter by utilizing an Recursive Least Square (RLS) algorithm. Spectrograms and convergence plots results in Ref. [52] concluded RLS algorithm to be more efficient in removing motion artifacts from ECG signals when compared to Least Mean Squares (LMS) algorithm. Hence, we adopted the motion artifacts removal approach in Ref. [52] wherein adaptive filters altered their filter coefficients with the continuous change of signal using adaptive algorithms, providing the optimum noise removal features for non-stationary signals like ECG.

2.2. Processing: heart beat detection and segmentation

2.2.1. HeartBeat detection

The heartbeat detection is accomplished at a lower sampling rate of 114 Hz as compared to the input sampling rate of 360 Hz. We followed the filterbank based approach enumerated in Ref. [53] for ECG detection algorithm wherein the Filterbanks subbands are downsampled and components of the one-channel detection block like the feature computation, MWI, and peak detector are operated at a lower rate than the input sampling rate of the ECG. The adoption of filter bank enables the analysis of multiple frequency bands very efficiently. This results in competent heartbeat detection at subband rate along with high computational efficiency. For R-peak detection at subband rate, it is crucial to have a deterministic relationship between fiducial points in the input ECG and the subband signal. This requires that each of the analysis and synthesis filters have a linear phase response as the linear phase requirement ensures that all frequencies in the input signal will have the same sample delay through the analysis filters. It is then possible to determine the exact location of the R wave in the input ECG signal, and other fiducial points from analysis of the subbands, thereby resulting in accurate R peak detection [53].

2.2.2. Segmentation

The segmentation module follow the ECG-detection module. For segmentation, we utilize the heartbeat segmentation program of Laguna et al. [4], since the accuracy of this system in determining heartbeat segmentation points has been validated on the MIT-BIH database and has proven to be commensurate with the inter-expert variation. The heartbeat segmentation stage provides QRS onset and T-wave offset times; a Boolean value indicates the presence/absence of a P-wave and, if present, gives the P-wave onset and offset time for each heartbeat fiducial point.

2.2.3. T-wave offset detection

T wave is the representation of repolarization of ventricles through which the myocardium is prepared for the next ECG cycle. In the automatic delineation process of the ECG i.e the process of locating the onset and offset of different waveforms, one of the most critical calculations is the detection of T wave ends, particularly T offset. The detection of the T offset is the most challenging among the ECG fiducial points, primarily due to the slow transition of the signal near the end of T wave and due to the presence of oscillatory patterns that vary from one person to another. We use the proposed method in Ref. [54], which is based on ECG signal filtering, value estimation of different fiducial points and application of backward and forward search windows with adaptive thresholds.

For an infallible detection of T waveform and its fiducial points (e.g. T onset, T peak, and T offset), the first step is to define the search window in the entitled signal. The boundaries of the T wave search window are set to be adaptive and relative to the position of QRS offset and RR interval wherein the search window encloses the T wave boundaries, extended from the QRS offset to two-third of previous RR interval. A small segment post the QRS offset is excluded from the search window

due to the fact that T wave does not exist immediately after QRS offset and occurs only after the ST segment.

The next stage is the adaptive threshold estimation wherein the position of T peak is registered by finding either the local maximum and the local minimum in the entitled ECG signal within the search window and the threshold is calculated using the previous T and R peak level values. The threshold is used to classify the morphology of the T wave present in each heartbeat. For eg: if the maximum point exceeds the threshold and the absolute value of the minimum point does not, then the T wave is identified as a positive monophasic and the maximum point is registered as T peak. For a negative monophasic T wave, the absolute minimum value must meet the defined requirement in Ref. [54]. In case of a biphasic T wave, both the local maximum and the absolute value of the local minimum should be greater than the threshold. For cases when T wave is not detected in the first attempt, the threshold is then decreased by half and the search method is done over again until a T peak is detected. Furthermore, the detection algorithm traces the onset and offset T values by finding the sample corresponding to the zero slope of the entitled ECG signal. The sample point which has a zero slope and former T peak is identified as T onset. Similarly, T offset is determined at a further distance of T peak as described in Ref. [54].

2.3. Feature extraction

After down-sampling the ECG signal recordings to 114 Hz, we employ two feature extraction methods. Feature Set 1 (FS1) yielded 26 features comprising RR intervals, heart-beat intervals, and segmented morphology. Feature Set 2 (FS2) produced 22 features consisting of RR intervals and fixed interval morphologies [4]. **We settled upon the single-lead feature extraction method as its lower sampling rate and smaller feature vector both translate to lessened power consumption and lower hardware complexity.**

2.3.1. Feature set 1

FS1 consisted of 26 features comprising of RR intervals, heartbeat intervals, and segmented morphologies.

2.3.2. RR intervals features

RR intervals also known as Heartbeat fiducial point intervals correspond to the interval between successive heartbeat fiducial points. The following four features were extracted from RR intervals:

- Pre-RR interval: the RR interval between a given heartbeat and the preceding heartbeat.
- Post-RR interval: the RR interval between a given heartbeat and the following heartbeat.
- Average RR interval: the mean of RR intervals for a recording. This value remains the same for all heartbeats in a recording.
- Local average RR interval: estimated by averaging ten RR intervals surrounding a heartbeat.

2.3.3. Heartbeat interval features

Three features were extracted from post-heartbeat interval segmentation.

- QRS duration: time interval between QRS onset and offset.
- T-wave duration: time interval between QRS offset and T-wave offset.
- Boolean variable: a third variable which indicates the presence or absence of a P-wave.

2.3.4. Segmented morphology interval features

Segmented morphology encompasses amplitude values of the ECG signal calculated by a sampling window between QRS onset and offset and a sampling window between QRS offset and T-wave offset points. Two sampling windows were used following the determination of the fiducial point (FP), the first of these bounded by the QRS onset and offset and the second bounded by the QRS offset and the T-wave offset. Ten evenly spaced sample features were derived by uniformly sampling the ECG amplitude in the first window (Fig. 2a) and nine more by uniformly sampling the second window, resulting in a total of 19 features.

2.3.5. Feature set 2

FS2's 22 features consisted of RR intervals and fixed interval morphologies [4].

2.3.5.1. RR intervals features. RR intervals (also known as Heartbeat fiducial point intervals) correspond to the interval between successive heartbeat fiducial points, and match the same four features extracted in Feature Set 1.

2.3.5.2. Fixed-interval morphology features. To determine fixed interval morphologies, sampling windows were first positioned at the heartbeat FP. Two sampling windows were formed based on FP. The first window approximately encompassed the QRS-complex and covered the portion of the ECG between FP-50 ms and 100 ms. Nine samples of the ECG between FP-50 ms and FP+100 ms were extracted from this window. The second window approximately covered the T-wave and started at 150 ms and finished at 500 ms. The next nine samples between FP+150 ms and FP+500 ms were extracted from the second window, for a total of 18 features used in FS2[Fig. 2b].

So in totality, the feature sets can be defined as the combination of the following:

- **Feature Set1 (FS1) (26)**: RR intervals (4), heartbeat intervals (3), segmented morphology (19)
- **Feature Set1 (FS2) (22)**: RR intervals (4), fixed interval morphology (18)

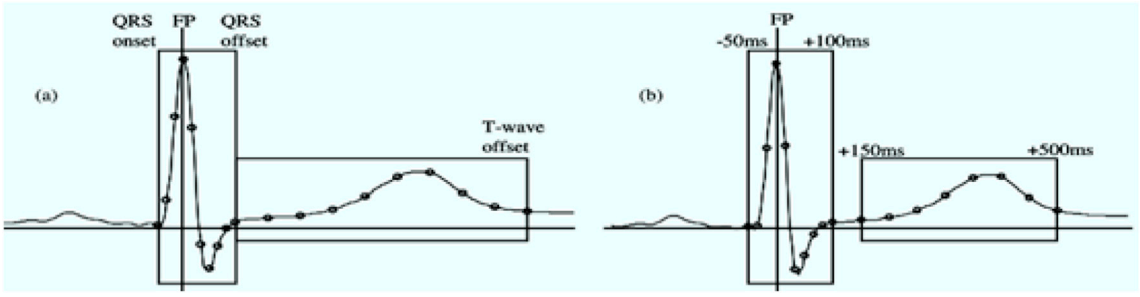

Fig. 2. Two time-sampling methods for extracting ECG morphology features. (a)Segmented Morphology Intervals Features: Post determination of the fudicial point (FP), the QRS onset and offset and T-wave offset points are found. Ten evenly spaced samples of the ECG between the QRS onset and offset and nine evenly spaced samples of the ECG between the QRS offset and T-wave offset are extracted. (b) Fixed-interval morphology features: Post determination of the FP, nine samples of the ECG between FP-50 ms and FP + 100 ms and nine samples between FP+150 ms and FP+500 ms are extracted [4].

An extensive statistical Hypothesis testing experimentation was performed to define these feature sets. The Hypothesis testing results indicated that segmented morphology features, heart-beat interval features and RR Interval features change significantly for **SVEB arrhythmia** case as compared to VEB and non-arrhythmia class. Due to the robustness of these features to distinguish between SVEB and remaining classes, these features constituted our **Feature Set 1**. The test also revealed that Fixed Interval Morphology and RR Intervals alter significantly for the **VEB arrhythmia** samples and hence constituted our **Feature Set 2**. In addition to Hypothesis testing, we also performed the multi-collinearity test which included a pair-wise correlation and found that multi-collinearity was not a problem in case of RBM- DBN model as the paired *t*-test revealed that the predictors are not correlated with other predictors in the model. The entire set of features can be summarized as shown in Table 1.

3. Deep learning framework

Deep Learning, inspired by the human brain's deep hierarchical architecture, is a technique focused on learning deep hierarchical models of data [55], [56]. This system learns an empirical set of features at multiple levels of abstraction, thereby allowing it to learn complex functions from input data without using human-engineered features.

Deep learning networks, implemented using stacked autoencoders, are capable of representing highly expressive abstractions, thereby compactly yielding much larger sets of functions than shallow networks can [57]. Through the tremendous representational power of hierarchical feature learning, these networks can help discover unknown feature coherences of input signals, a characteristic that is crucial for learning tasks involving complicated models.

As suggested in Ref. [25], the main concept of a DBN training algorithm is to first initialize greedily the weights of each layer in an unsupervised manner by treating each pair of layers as a Restricted Boltzmann Machine (RBM), and to later jointly refine these weights to further improve the likelihood. The resulting DBN can be considered a hierarchy of nonlinear feature detectors that can capture complex statistical patterns in the data.

3.1. Restricted Boltzmann Machine

Derived from a Boltzmann Machine, the RBM is a bi-directionally connected network of stochastic processing units that learns significant features of an unknown probability distribution based on samples from that distribution. An RBM can be described as a bipartite graph having a visible layer and a hidden layer (Fig. 3). Units in the visible layer are typically characterized by Bernoulli or Gaussian distributions and those in the hidden layer are typically characterized by Bernoulli distributions.

Table 1
Description of features included in feature group label.

Feature Group Label	Feature Included within the Group Label
RR Intervals	• Pre-RR Interval • Post-RR Interval • Average-RR Interval • Local avg -RR Interval
Heartbeat Intervals	• QRS Duration (QRS offset-QRS onset) • T-wave duration(T-wave offset-QRS offset) • P wave flag
Segmented Morphology Intervals	• ECG morphology (10 samples) between QRS onset and QRS offset of lead A • ECG morphology (9 samples) between QRS offset and T-wave offset of lead A
Fixed Interval Morphology	• ECG morphology (10 samples) between FP-50 ms to FP + 100 ms of lead A • ECG morphology (9 samples) between FP + 150 ms to FP + 500 ms of lead A

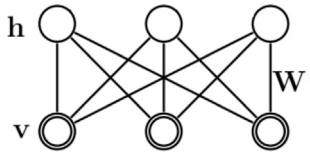

Fig. 3. Schematic of a restricted Boltzmann machine.

Stochastic units in the visible layer associate with stochastic units in the hidden layer by means of a weight matrix. No connections exist between units in the same layer. In schematic representation, each edge in the bipartite graph is attached to a weight, denoted as a symmetric matrix W, that is associated with the visible layer (v) and the hidden layer (h).

A given RBM defines an energy function for every configuration of visible and hidden state vectors. If both v and h are binary states (i.e., the Bernoulli-Bernoulli RBM) the energy function is given by

$$\mathrm{E}(v,h) = -v^T Wh - b_v^T v - b_h^T h \tag{1}$$

Thus, an RBM represents the joint distribution $p(v;h)$ between visible unit v and hidden random unit h. The joint probability is defined as

$$p(v,h) = \frac{\exp(-\mathrm{E}(v,h))}{Z}, \tag{2}$$

where $Z = \sum_v \sum_h \exp(-\mathrm{E}(v,h))$ is the partition function.

The probability assigned by the network model to a visible unit v is

$$p(v) = \frac{1}{Z}\sum_h \exp(-\mathrm{E}(v,h)), \tag{3}$$

The lack of connections within a given layer of an RBM results in the visible layer variables being conditionally independent, given the hidden layer variables, and vice versa. Thus the conditional probabilities can be rewritten as:

$$p\left(v_j = 1/h\right) = \sigma\left(a_i + \sum h_j w_i, j\right) \tag{4}$$

$$p\left(h_j = 1/v\right) = \sigma\left(b_j + \sum v_i w_i, j\right) \tag{5}$$

where σ is the sigmoid fnuction defined by $\sigma = \frac{1}{1+\exp^{-x}}$

Signal propagation manifests in two ways: recognition, where visible activations propagate to the hidden units; and reconstruction, where hidden activations propagate to the visible units. Both recognition and reconstruction use the same weight matrix (simply transposed). The Contrastive Divergence (CD) algorithm finds the parameters W, a, and b and performs Gibbs sampling. We use CD to minimize the reconstruction error so the weights can be trained to generate input patterns that are presented to the RBM with high probability. (A guide to training an RBM is given in Ref. [58]).

3.2. Deep belief networks

DBNs are a type of multi-layer generative neural network that is recognized for its capability to model and visualize high-level learned features [59], [55]. It is composed of stacked, logistic RBMs wherein the lowest-level RBM learns a shallow model of the data and the next-level RBM learns to model first-layer hidden units, thereby representing high-level abstraction through hierarchical architecture (Fig. 4).

When a DBN is used for classification purposes, the RBM pre-training procedure can be used to initialize the weights of the deep neural network, and then these weights can be discriminatively fine-tuned by back-propagating error derivatives. The recognition weights of the DBN become the weights of a standard neural network. This unsupervised pre-training establishes the platform for a final training phase: a fine-tuning

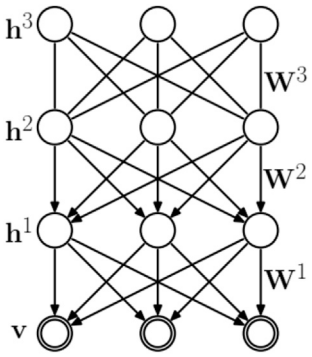

Fig. 4. Schematic of a deep belief network of three layers.

process with respect to a supervised training criterion based on gradient descent optimization. (A detailed description of DBN varieties and their training is available in Refs. [31] [60]).

4. Evaluation of proposed methodology with results and discussion

4.1. The MIT-BIH database

For our evaluation experiments, we used the acclaimed MIT/Beth Israel Hospital (BIH) Arrhythmia Database available at MIT medical data storage Physionet [61]. Briefly, this database incorporates 48 half-hour ECG recordings, each containing two ECG lead signals digitized at 360 samples per second with 11-bit resolution over a 10 mV range [61]. Twenty-three of the recordings were randomly selected from a set of 4000 ambulatory 24-h ECGs that were collected from a mixed population of inpatients and outpatients at the medical center. The remaining 25 recordings were selected from the same set but included less common but clinically symbolic arrhythmias. All recordings have been annotated by two or more cardiologists and contain modified limb lead II. In our experiments, we focused on using lead A only. In 45 recordings, lead A is modified lead II, and in the other three recordings, lead A is lead V5 [6]. According to the Association for the Advancement of Medical Instrumentation (AAMI) recommended practice, the 4 paced beats are excluded from this experimental evaluation process because these beats possess insufficient signal quality for reliable processing [4], [5].

4.2. AAMI standard

The AAMI standard emphasizes the problem of distinguishing ventricular ectopic beats (VEBs) from non-ventricular ectopic beats [4], and hence normal and arrhythmic beats are remapped to the five AAMI heartbeat classes [62] [63], using the mapping in Ref. [4] with each class including heartbeats of one or more types [Table 2].

In this work, we used AAMI recommended practice to combine the MIT-BIH heartbeat types into the following five heartbeat classes that we used in all subsequent processing:

1. Class N corresponding to beats originating in the sinus node (normal and bundle branch block beat types)
2. Class S corresponding to supraventricular ectopic beats (SVEBs)
3. Class V corresponding to ventricular ectopic beats (VEBs)
4. Class F corresponding to beats that result from fusing normal and VEBs
5. Class Q corresponding to unknown beats including paced beats

4.3. Evaluation metrics

The MIT-BIH database contained a series of manually verified QRS detection points that we utilized in this study. The 48 records from MIT/BIH ECG arrhythmia database [17] [19], are used for the development and evaluation of the classifier. Following the AAMI recommendation, 4 paced beats are removed from the database. After the four recordings containing paced beats were removed as in Ref. [62], the remaining 44 recordings were divided into two equal-sized training and testing datasets, each containing ECG data from 22 recordings [Fig. 5]. The first dataset (DS1) (training set) was used to train the classifier and to set parameter values that optimized performance of the classifier. The second dataset (DS2) (testing set)was employed to carry out an independent and unbiased performance evaluation of the heartbeat classification system.

The training set consisted of record numbers 101, 106, 108, 109, 112, 114, 115, 116, 118, 119, 122, 124, 201, 203, 205, 207, 208, 209, 215, 220, 223, and 230; and the testing set consisted of record numbers 100, 103, 105, 111, 113, 117, 121, 123, 200, 202, 210, 212, 213, 214, 219, 221, 222, 228, 231, 232, 233, and 234. The objective of the proposed framework is to effectively classify the two critically abnormal

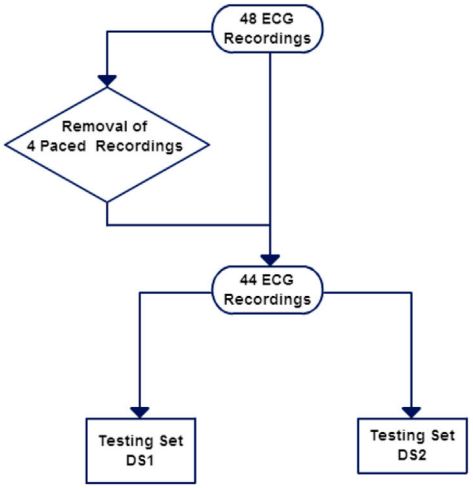

Fig. 5. MIT-BIH database division into training and testing sets.

Table 2
MIT-BIH arrhythmia database heartbeat mapped to AAMI heartbeat classes.

Remapped Classes	Classes Mapped in accordance to AAMI Standard				
Class N	Normal Beat (NOR)	Left Bundle Branch Block (LBBB)	Right Bundle Branch Block (RBBB)	Atrial Escape Beat (AE)	Nodal Escape Beat (NE)
Class S	Atrial Premature beat (AP)	Aberrated atrial premature beat (aAP)	supra ventricular premature beat (SP)	Nodal premature beat (NP)	
Class V	Premature Ventricular Contraction (PVC)	Ventricular escape beat (VE)			
Class F	Fusion of normal & ventricular beat (Fvn)				
Class Q	paced beat (P)	Fusion of paced & normal beat (fPN)	unclassified beat (U)		

heartbeats: supraventricular ectopic beats (SVEB) and ventricular ectopic beats (VEB).

To validate the algorithms on the MIT-BIH database, we used the following performance metrics: accuracy (Acc), sensitivity (Se), positive predictive value (PPV), and false positive rate (FPR).

4.4. Experimental results and discussion

We performed our classification on the MIT-BIH arrhythmia database [61] to detect two types of heartbeat arrhythmias: **VEBs and SVEBs**. In agreement with AAMI recommended practice, four recordings containing paced beats were removed from the 48 recordings. The data from the remaining 44 recordings were divided into two sets: training (DS1) and test (DS2). We trained the classifier using DS1 and assessed classifier performance using test set DS2.

For the RBM and DBN algorithms, we used the toolbox developed by Drausin Wulsin [64]. To determine the best configurations and parameters for the RBM, we performed a large number of experiments where we used varying combinations of batch sizes (i.e., number of training vectors used in each pass of each epoch for the contrastive divergence algorithm), numbers of hidden units, learning rates, and numbers of stacked RBMs. The final classification layer had five output units, one for each class, and the unit with the highest activation level was considered the most probable class.

We have reported our ECG classification results at sampling rates of 360 Hz and 114 Hz in Table 3 and Table 4, respectively. Column 1 indicates the methodology and column 2 indicates the sampling rate; column 3–4 represent the CPU Cycles and time consumed; columns 5–12 indicate the gross classifier performance in terms of Acc (Accuracy), Se (Sensitivity), PPV (Positive predictive value) and FPR (False positive rate) for the two arrhythmia types. Rows 1 and 2 in both tables report the overall performance of our classification using Feature Set 1 and Deep Learning and using Feature Set 2 and Deep Learning, respectively. The independent performance assessment of the configuration of **FS1 and Deep learning** (FS1-DL)framework resulted in an accuracy of 93.78%, a sensitivity of 88.39%, a positive predictivity of 33.63%, and an FPR of 6.68% for the **SVEB class**. For the **VEB class**, **FS2 and Deep learning framework** (FS1-DL) performed better with an accuracy of 96.94%, sensitivity of 85.22%, positive predictivity of 56.63%, and False positive ratio of 4.11%.

Chazel [4] reported a classification accuracy of 94.6% and 97.4% for SVEB & VEB Class but consumed higher CPU cyles as they used complex feature extraction calculation at a higher sampling rate (360 Hz). A comparable accuracy of 93.6% at 360 Hz was obtained using a simpler feature sets in Ref. [6] along with reduced CPU cycle and time consumption. Our proposed Deep learning framework at a sampling rate of 360 Hz provided a high accuracy of 93.78% for the SVEB class using the feature set 1 and of 96.94% for the VEB class using feature set 2.

Our classifier also provided high levels of accuracy, 93.63% and 95.87% for SVEB and VEB classes, respectively when compared to LCKVD [65], LDA [66], QDA [66] and ANN [66] at a lower sampling rate of 114 Hz. This provided advantages of not only reduced computation time, reduced CPU Cycles but also a higher accuracy achieved at lower sampling rate of 114 Hz. Thus at both sampling rates, our algorithm provides competitive accuracy performance when compared to previously reported results (rows 3–5 in Table 3 and rows 3–6 in Table 4) for automated heartbeat classification systems in Refs. [4] and [6].

We also evaluated the recognition accuracy for different sample rates from 72 Hz to 360 Hz for VEB & SVEB Classes in Fig. 6 & Fig. 7. As per the analysis, the classification accuracy increases with higher sampling rates, and with the proposed simpler set of features and deep learning framework, the accuracy stabilizes between 114 Hz and 180 Hz, and is only improved marginally with higher sampling rates. The accuracy with 20% of samples (72 Hz) is comparitively lower for all algorithms and it stabilizes for our proposed deep learning framework beginning 114 Hz (which is 30% of the sampling rate). At the 60% sampling rate (i.e 216 Hz), the accuracies of all methods are nearly the same, although the FS1-DL yielded higher accuracies than the other methods for SVEB Arrhythmia and FS2-DL yielded somewhat higher accuracies than the other methods for VEB Arrthymia. The results obtained by the Deep Learning and Dictionary learning (LCKSVD) [65] methods are very similar over all sampling rates starting 114 Hz, with other algorithms yielding comparitively lower accuracies at 114 Hz. The boosting methods algorithms in Chazel [4] [6], and LDA, QDA, ANN methods in Ref. [66] uniformly produce inferior average accuracies as the sampling rate is reduced starting from 180 Hz(50%) to 72 Hz(20%).

The reduced accuracy of other algorithms at low sampling rates, relative to the deep learning methods, can be attributed to the value of the inherent exploitation of class affinities by the deep learning approaches. Though the boosting algorithms perform inferior at low sampling rates, they seem to benefit from bagging of the training sample occurring at the higher sampling rates. Thus, the boosting and other algorithms accuracies increase only after the sampling rate of the training data is increased substantially. Overall, the average accuracy of FS1-DL and FS2-DL is consistently higher for SVEB arthymia and VEB arrthymia clasification respectively and improves relative to other boosting methods used in Chazel [4] [6] methods and LDA, QDA, ANN methods in Ref. [66], as the fraction of training data is reduced. The proposed deep

Table 3
Comparison of classification results on the test dataset at sampling rate of 360 Hz.

Method	Rate (Hz)	CPU Cycles	Time in μ s	SVEB				VEB			
				Acc	Se	PPV	FPR	Acc	Se	PPV	FPR
FS1+DL	360	18917	328.1	93.78	88.39	33.63	6.68	96.63	77.74	69.20	2.17
FS2+DL	360	19563	317.8	93.47	70.99	32.44	5.66	96.94	85.22	56.63	4.11
Chazel et al. [4]	360	57341	964.2	94.6	75.9	38.5	4.7	97.4	77.7	81.9	1.2
Chazel et al. [6]	360	22731	357.4	93.6	61.2	31.2	5.2	95.4	72.4	62.3	3.0
Chazel et al. [6]	360	23545	392.7	94.4	73.5	37.0	4.8	97.8	87.6	80.3	1.5

Table 4
Comparison of classification results on the test dataset at sampling rate of 114 Hz.

Method	Rate (Hz)	CPU Cycles	Time in μ s	SVEB				VEB			
				Acc	Se	PPV	FPR	Acc	Se	PPV	FPR
FS1+DL	114	12884	225.8	93.63	88.62	35.49	6.17	95.57	78.49	59.65	3.34
FS2+DL	114	13664	227.4	93.42	59.16	30.10	5.26	95.87	85.54	60.83	3.47
LCKSVD [65]	114	14820	252.3	93.4	75.12	32.84	5.89	93.51	76	49.97	5.27
LDA_Basil [66]	114	15785	257.3	–	–	–	–	93.4	75.8	61.9	4.8
QDA_Basil [66]	114	15911	258.6	–	–	–	–	83.1	97	35.2	18.4
ANN_Basil [66]	114	16089	294.5	–	–	–	–	96.9	79.7	74.6	1.9

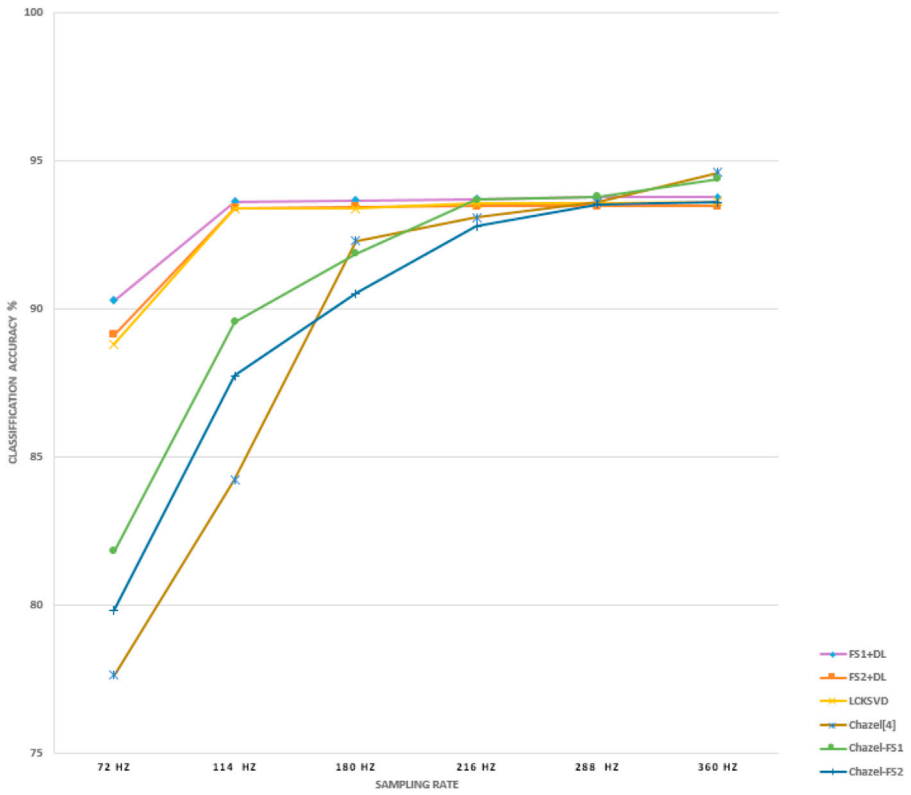

Fig. 6. Recognition Accuracy for VEB arrhythmia at different sampling rate.

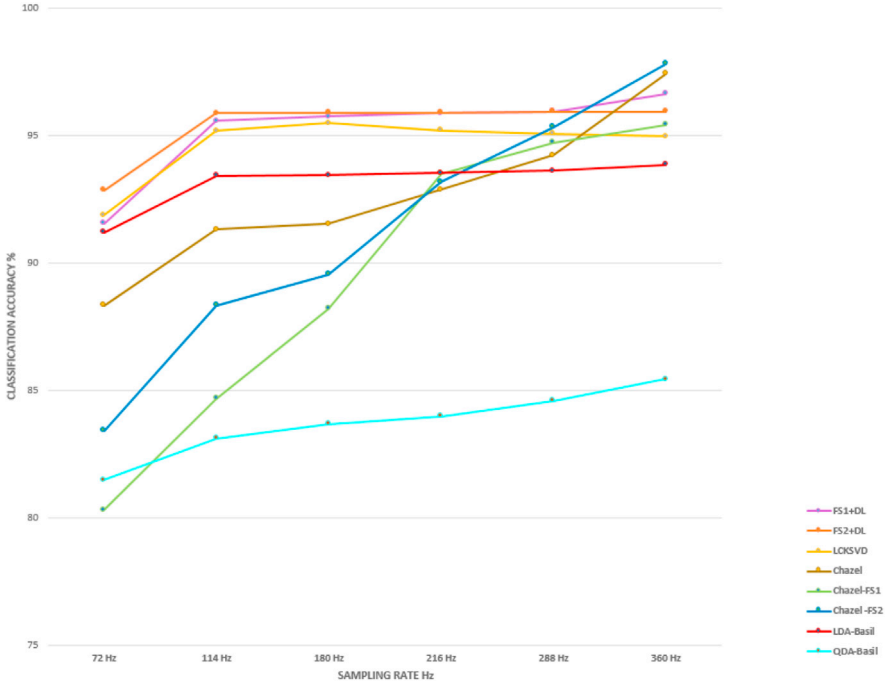

Fig. 7. Recognition Accuracy for VEB arrhythmia at different sampling rate.

learning framework thus proves to be the most stable method over all training fractions as observed in both the cases, as measured by the deviation of the accuracies.

Comparing Table 3 against Table 4, our results demonstrate that a deep learning algorithm framework is better suited to detect VEB and SVEB types of arrhythmia at the lower sampling rate of 114 Hz. Since increasing the sampling rate to 360 Hz did not provide any significant gain in performance, it follows that a 114 Hz sampling rate can provide robust discriminatory power for this classification task.

In addition to evaluating the performance of VEB and SVEB classification, we also assessed the per-class classification across all five classes at sampling rates of 360 Hz and 114 Hz (Tables 5–8). Our per-class classification results are competitive when compared to results from methods in Refs. [6] [8], [13], [40]. **Since varying the sampling rate**

Table 5
Confusion Matrix obtained with test set using FS1 + Deep Learning Framework at Sampling Rate of 360 Hz.

Method	N	S	V	F	Q
N	31228	2758	841	7011	422
S	38	1539	126	36	2
V	54	275	2173	293	0
F	16	3	5	360	0
Q	0	1	1	2	2

Table 6
Confusion Matrix obtained with test set using FS2 + Deep Learning Framework at Sampling Rate of 360 Hz.

Method	N	S	V	F	Q
N	36386	2347	1470	2045	12
S	148	1236	354	3	0
V	96	225	2382	91	1
F	22	1	103	258	0
Q	1	1	1	0	1

Table 7
Confusion Matrix obtained with test set using FS1 + Deep Learning Framework at Sampling Rate of 114 Hz.

Method	N	S	V	F	Q
N	31238	2565	1366	6451	640
S	41	1543	118	39	0
V	60	238	2194	303	0
F	17	1	22	344	0
Q	1	0	2	0	1

Table 8
Confusion Matrix obtained with test set using FS2 + Deep Learning Framework at Sampling Rate of 114 Hz.

Method	N	S	V	F	Q
N	36213	2189	1181	2648	29
S	347	1030	358	6	0
V	100	201	2391	100	3
F	19	1	113	251	0
Q	1	1	1	0	1

had minimal impact on performance, we conclude that our approach emulates the performance of state-of-the-art models at a lower sampling rate and with a set of simple features.

Notably, the combination of parameters that yielded the best result was 112 hidden units, a batch size of 42, and a learning rate of 0.00001. With this combination, the DBN achieved a very low error rate of 4.7%. We used this DBN configuration while varying the number of stacked RBMs and found that outputs of three layered trained RBMs achieved the best performance results. Moreover, this performance level was comparable to that of state-of-the-art ECG classification algorithms.

In summary, we demonstrate that our approach is able to emulate state-of-the-art classification results while using a significantly lower sampling rate. In the case of an average sampling rate of 114 samples versus 360, we are actually achieving a **gain factor of three**. In fact the smaller feature-set representation and the deep learning framework together contain sufficient discriminative information for accurate ECG classification. Thus, with a suitable choice of parameters, the classifiers built using this deep learning framework provide competitive performance. In addition, our proposed framework opens a new window for future research, highlighting the huge potential of deep learning based methods for accurate classification of other physiological signals, such as arterial blood pressure (ABP), electromyograms (EMG), and heart rate variability (HRV).

5. Conclusion

In this work we considered the application of Restricted Boltzmann Machines and deep belief networks to the automatic classification of single-lead ECG signals. Experimental results indicate that our deep learning framework demonstrates a classification accuracy on the MIT-BIH database of 93.78% for SVEB class signals and 96.94% for VEB class signals at a sampling rate of 360 Hz. Thus our framework provides performance competitive with that of state-of-the-art methods. Experimental results also demonstrate that this framework provides similar classification accuracy (93.63% for SVEB and 95.87% for VEB) when sampling at only 114 Hz. Thus a lower sampling rate of 114 Hz is sufficient to provide good discriminatory power for the ECG classification task.

In conclusion, our approach emulated the performance of state-of-the-art models using a lower sampling rate and a set of simple features. As future work, we will investigate other types of embedding that represent ECG recordings as a feature vector and then use hierarchical deep learning algorithms for robust performance. We would also like to integrate Restricted Boltzmann Machines with ensemble based/Bagging approachto construct multiple individual classifiers. Considering the fact that both deep learning and ensemble learning have leverage in constructing complicated nonlinear functions, the combination of the two frameworks can better handle hard artificial intelligence tasks. We will also extend our framework to the classification of sensor-based cognitive assessment data and the recognition of daily life activities, areas important in healthcare for ubiquitous health computing and medical informatics.

References

[1] A. Guyton, J. Hall, 11th Textbook of Medical Physiology.
[2] L. Glass, Cardiac oscillations and arrhythmia analysis, in: Complex Systems Science in Biomedicine, Springer, 2006, pp. 409–422.
[3] ACC/AHA clinical competence statement on electrocardiography and ambulatory electrocardiography, J. Am. Coll. Cardiol. 38 (7) (2001) 2091–2100.
[4] P. De Chazal, M. O'Dwyer, R.B. Reilly, Automatic classification of heartbeats using ECG morphology and heartbeat interval features, IEEE (Inst. Electr. Electron. Eng.) Trans. Biomed. Eng. 51 (7) (2004) 1196–1206.
[5] P. de Chazal, R.B. Reilly, A patient-adapting heartbeat classifier using ECG morphology and heartbeat interval features, IEEE (Inst. Electr. Electron. Eng.) Trans. Biomed. Eng. 53 (12) (2006) 2535–2543.
[6] P. de Chazal, Detection of supraventricular and ventricular ectopic beats using a single lead ECG, in: 35th Annual International Conference of IEEE Engineering in Medicine and Biology Society (EMBC), 2013, pp. 45–48.
[7] M. Llamedo, J.P. Martínez, Heartbeat classification using feature selection driven by database generalization criteria, IEEE (Inst. Electr. Electron. Eng.) Trans. Biomed. Eng. 58 (3) (2011) 616–625.
[8] C. Ye, B.V. Kumar, M.T. Coimbra, Heartbeat classification using morphological and dynamic features of ECG signals, IEEE (Inst. Electr. Electron. Eng.) Trans. Biomed. Eng. 59 (10) (2012) 2930–2941.
[9] M. Lagerholm, C. Peterson, G. Braccini, L. Edenbrandt, L. Sornmo, Clustering ECG complexes using hermite functions and self-organizing maps, IEEE (Inst. Electr. Electron. Eng.) Trans. Biomed. Eng. 47 (7) (2000) 838–848.
[10] T. Ince, S. Kiranyaz, M. Gabbouj, A generic and robust system for automated patient-specific classification of ECG signals, IEEE (Inst. Electr. Electron. Eng.) Trans. Biomed. Eng. 56 (5) (2009) 1415–1426.
[11] X. Jiang, L. Zhang, Q. Zhao, S. Albayrak, ECG arrhythmias recognition system based on independent component analysis feature extraction, in: IEEE TENCON Region Conference, 2006, pp. 1–4.
[12] L. Senhadji, G. Carrault, J. Bellanger, G. Passariello, Comparing wavelet transforms for recognizing cardiac patterns, In IEEE Engineering in Medicine and Biology Magazine 14 (2) (1995) 167–173.
[13] Y.H. Hu, S. Palreddy, W.J. Tompkins, A patient-adaptable ECG beat classifier using a mixture of experts approach, IEEE (Inst. Electr. Electron. Eng.) Trans. Biomed. Eng. 44 (9) (1997) 891–900.
[14] M. Lagerholm, C. Peterson, G. Braccini, L. Edenbrandt, L. Sornmo, Clustering ECG complexes using hermite functions and self-organizing maps, IEEE (Inst. Electr. Electron. Eng.) Trans. Biomed. Eng. 47 (7) (2000) 838–848.
[15] S. Osowski, T.H. Linh, ECG beat recognition using fuzzy hybrid neural network, IEEE (Inst. Electr. Electron. Eng.) Trans. Biomed. Eng. 48 (11) (2001) 1265–1271.
[16] S. Osowski, L.T. Hoai, T. Markiewicz, Support vector machine-based expert system for reliable heartbeat recognition, IEEE (Inst. Electr. Electron. Eng.) Trans. Biomed. Eng. 51 (4) (2004) 582–589.
[17] G. de Lannoy, D. François, J. Delbeke, M. Verleysen, Weighted conditional random fields for supervised interpatient heartbeat classification, IEEE (Inst. Electr. Electron. Eng.) Trans. Biomed. Eng. 59 (1) (2012) 241–247.

[18] J. Rodriguez, A. Goni, A. Illarramendi, Real-time classification of ECGs on a PDA, IEEE Trans. Inf. Technol. Biomed. 9 (1) (2005) 23–34.
[19] I. Christov, I. Jekova, G. Bortolan, Premature ventricular contraction classification by the kth nearest-neighbours rule, Physiol. Meas. 26 (1) (2005) 123.
[20] W. Jiang, S.G. Kong, Block-based neural networks for personalized ECG signal classification, IEEE Trans. Neural Network. 18 (6) (2007) 1750–1761.
[21] J. Wiens, J.V. Guttag, Active learning applied to patient-adaptive heartbeat classification, in: Advances in Neural Information Processing Systems, 2010, pp. 2442–2450.
[22] M.D. Zeiler, R. Fergus, Visualizing and understanding convolutional networks, in: ECCV Computer Vision, Springer, 2014, pp. 818–833.
[23] A.S. Razavian, H. Azizpour, J. Sullivan, S. Carlsson, CNN features off-the-shelf: an astounding baseline for recognition, in: IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2014, pp. 512–519.
[24] Y. Taigman, M. Yang, M. Ranzato, L. Wolf, Deepface: closing the gap to human-level performance in face verification, in: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2014, pp. 1701–1708.
[25] L. Deng, O. Abdel-Hamid, D. Yu, A deep convolutional neural network using heterogeneous pooling for trading acoustic invariance with phonetic confusion, in: IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2013, pp. 6669–6673.
[26] P. Sermanet, K. Kavukcuoglu, S. Chintala, Y. LeCun, Pedestrian detection with unsupervised multi-stage feature learning, in: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2013, pp. 3626–3633.
[27] P. Weinzaepfel, J. Revaud, Z. Harchaoui, C. Schmid, Deepflow: large displacement optical flow with deep matching, in: IEEE International Conference on Computer Vision (ICCV), 2013, pp. 1385–1392.
[28] A. Toshev, C. Szegedy, Deeppose: human pose estimation via deep neural networks, in: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2014, pp. 1653–1660.
[29] A. Jain, J. Tompson, M. Andriluka, G. W. Taylor, C. Bregler, Learning Human Pose Estimation Features with Convolutional Networks, arXiv preprint arXiv:1312.7302.
[30] J. Tompson, M. Stein, Y. Lecun, K. Perlin, Real-time continuous pose recovery of human hands using convolutional networks, ACM Trans. Graph. 33 (5) (2014) 169.
[31] G.E. Hinton, S. Osindero, Y.-W. Teh, A fast learning algorithm for deep belief nets, Neural Computation 18 (7) (2006) 1527–1554.
[32] V. Nair, G.E. Hinton, 3D object recognition with deep belief nets, in: Advances in Neural Information Processing Systems, 2009, pp. 1339–1347.
[33] Y. Sun, Y. Chen, X. Wang, X. Tang, Deep learning face representation by joint identification-verification, in: Advances in Neural Information Processing Systems, 2014, pp. 1988–1996.
[34] A. Krizhevsky, I. Sutskever, G.E. Hinton, Imagenet classification with deep convolutional neural networks, in: Advances in Neural Information Processing Systems, 2012, pp. 1097–1105.
[35] G.E. Dahl, D. Yu, L. Deng, A. Acero, Context-dependent pre-trained deep neural networks for large-vocabulary speech recognition, IEEE Trans. Audio Speech Lang. Process. 20 (1) (2012) 30–42.
[36] G. Hinton, L. Deng, D. Yu, G.E. Dahl, A.-r. Mohamed, N. Jaitly, A. Senior, V. Vanhoucke, P. Nguyen, T.N. Sainath, et al., Deep neural networks for acoustic modeling in speech recognition: the shared views of four research groups, IEEE Signal Process. Mag. 29 (6) (2012) 82–97.
[37] F. Seide, G. Li, D. Yu, Conversational speech transcription using context-dependent deep neural networks, in: Interspeech, 2011, pp. 437–440.
[38] L. Deng, G. Hinton, B. Kingsbury, New types of deep neural network learning for speech recognition and related applications: an overview, in: IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2013, pp. 8599–8603.
[39] Y. Yan, X. Qin, Y. Wu, N. Zhang, J. Fan, L. Wang, A restricted Boltzmann machine based two-lead electrocardiography classification, in: Wearable and Implantable Body Sensor Networks (BSN), 2015 IEEE 12th International Conference on, IEEE, 2015, pp. 1–9.
[40] M.M. Al Rahhal, Y. Bazi, H. AlHichri, N. Alajlan, F. Melgani, R.R. Yager, Deep learning approach for active classification of electrocardiogram signals, Inf. Sci. 345 (2016) 340–354.
[41] J. Schmidhuber, Deep learning in neural networks: an overview, Neural Networks 61 (2015) 85–117.
[42] U.R. Acharya, H. Fujita, S.L. Oh, U. Raghavendra, J.H. Tan, M. Adam, A. Gertych, Y. Hagiwara, Automated identification of shockable and non-shockable life-threatening ventricular arrhythmias using convolutional neural network, Future Generat. Comput. Syst. 79 (2018) 952–959.
[43] M. Zubair, J. Kim, C. Yoon, An automated ecg beat classification system using convolutional neural networks, in: IT Convergence and Security (ICITCS), 2016 6th International Conference on, IEEE, 2016, pp. 1–5.
[44] S. Kiranyaz, T. Ince, M. Gabbouj, Real-time patient-specific ecg classification by 1-d convolutional neural networks, IEEE (Inst. Electr. Electron. Eng.) Trans. Biomed. Eng. 63 (3) (2016) 664–675.
[45] D. Wang, Y. Shang, Modeling physiological data with deep belief networks, International journal of information and education technology (IJIET) 3 (5) (2013) 505.
[46] M. Huanhuan, Z. Yue, Classification of electrocardiogram signals with deep belief networks, in: Computational Science and Engineering (CSE), 2014 IEEE 17th International Conference on, IEEE, 2014, pp. 7–12.
[47] K. Luo, J. Li, Z. Wang, A. Cuschieri, Patient-specific deep architectural model for ecg classification, J. Healthc. Eng. 2017 (2017), 4108720.
[48] Y. Kutlu, G. Altan, N. Allahverdi, Arrhythmia classification using waveform ecg signals, in: International Conference on Advanced Technology & Sciences Google Scholar, 2016.
[49] S. Chopra, R. Hadsell, Y. LeCun, Learning a similarity metric discriminatively, with application to face verification, in: IEEE Computer Society Conference on Computer Vision and Pattern Recognition(CVPR), vol. 1, 2005, pp. 539–546.
[50] J. Hu, J. Lu, Y.-P. Tan, Discriminative deep metric learning for face verification in the wild, in: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2014, pp. 1875–1882.
[51] Y. Sun, X. Wang, X. Tang, Hybrid deep learning for face verification, in: IEEE International Conference on Computer Vision (ICCV), 2013, pp. 1489–1496.
[52] A.C. Mugdha, F.S. Rawnaque, M.U. Ahmed, A study of recursive least squares (rls) adaptive filter algorithm in noise removal from ecg signals, in: Informatics, Electronics & Vision (ICIEV), 2015 International Conference on, IEEE, 2015, pp. 1–6.
[53] V.X. Afonso, W.J. Tompkins, T.Q. Nguyen, S. Luo, Ecg beat detection using filter banks, IEEE Transactions on Biomedical Engineering 46 (2) (1999) 192–202.
[54] N. Bayasi, T. Tekeste, H. Saleh, A. Khandoker, B. Mohammad, M. Ismail, Adaptive technique for p and t wave delineation in electrocardiogram signals, in: Engineering in Medicine and Biology Society (EMBC), 2014 36th Annual International Conference of the IEEE, IEEE, 2014, pp. 90–93.
[55] Q.V. Le, Building high-level features using large scale unsupervised learning, in: Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on, IEEE, 2013, pp. 8595–8598.
[56] Y. Bengio, Learning deep architectures for Ai, foundations and trends® in machine learning 2 (1) (2009) 1–127.
[57] P. Vincent, H. Larochelle, I. Lajoie, Y. Bengio, P.-A. Manzagol, Stacked denoising autoencoders: learning useful representations in a deep network with a local denoising criterion, J. Mach. Learn. Res. 11 (2010) 3371–3408.
[58] G.E. Hinton, A practical guide to training Restricted Boltzmann Machines, in: Neural Networks: Tricks of the Trade, Springer, 2012, pp. 599–619.
[59] H. Lee, R. Grosse, R. Ranganath, A.Y. Ng, Convolutional deep belief networks for scalable unsupervised learning of hierarchical representations, in: Proceedings of the 26th Annual International Conference on Machine Learning, ACM, 2009, pp. 609–616.
[60] G.E. Hinton, Training products of experts by minimizing contrastive divergence, Neural Computation 14 (8) (2002) 1771–1800.
[61] R. Mark, G. Moody, MIT-bih Arrhythmia Database, third ed..
[62] Testing and Reporting Performance Results of Cardiac Rhythm and ST-segment Measurement Algorithms, 1999.
[63] Recommended Practice for Testing and Reporting Performance Results of Ventricular Arrhythmia Detection Algorithms, 1986.
[64] D. Wulsin, DBN Toolbox, Department of Bioengineering, University of Pennsylvania, 2010.
[65] S.M. Mathews, L.F. Polania, K.E. Barner, Leveraging a discriminative dictionary learning algorithm for single-lead ecg classification, in: 41st Annual Northeast Biomedical Engineering Conference (NEBEC), IEEE, 2015, pp. 1–2.
[66] T. Basil, B.S. Chandra, C. Lakshminarayan, A comparison of statistical machine learning methods in heartbeat detection and classification, in: Big Data Analytics, Springer, 2012, pp. 16–25.