Artificial Intelligence In Medicine 103 (2020) 101789

Contents lists available at ScienceDirect

Artificial Intelligence In Medicine

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

Comprehensive electrocardiographic diagnosis based on deep learning☆

Oh Shu Lih[a], V Jahmunah[a], Tan Ru San[b], Edward J Ciaccio[c], Toshitaka Yamakawa[d], Masayuki Tanabe[d,g], Makiko Kobayashi[d], Oliver Faust[e], U Rajendra Acharya[a,f,g,*]

[a] Department of Electronics and Computer Engineering, Ngee Ann Polytechnic, Singapore
[b] National Heart Centre, Singapore
[c] Department of Medicine, Cardiology, Columbia University, USA
[d] Department of Computer Science and Electrical Engineering, Kumamoto University, Japan
[e] Department of Engineering and Mathematics, Sheffield Hallam University, United Kingdom
[f] Department of Bioinformatics and Medical Engineering, Asia University, Taichung, Taiwan
[g] International Research Organization for Advanced Science and Technology (IROAST) Kumamoto University, Kumamoto, Japan

ARTICLE INFO

Keywords:
Cardiovascular diseases
Coronary artery disease
Myocardial infarction
Congestive heart failure
Deep learning
10-fold validation
Convolutional neural network
Long short-term memory

ABSTRACT

Cardiovascular disease (CVD) is the leading cause of death worldwide, and coronary artery disease (CAD) is a major contributor. Early-stage CAD can progress if undiagnosed and left untreated, leading to myocardial infarction (MI) that may induce irreversible heart muscle damage, resulting in heart chamber remodeling and eventual congestive heart failure (CHF). Electrocardiography (ECG) signals can be useful to detect established MI, and may also be helpful for early diagnosis of CAD. For the latter especially, the ECG perturbations can be subtle and potentially misclassified during manual interpretation and/or when analyzed by traditional algorithms found in ECG instrumentation. For automated diagnostic systems (ADS), deep learning techniques are favored over conventional machine learning techniques, due to the automatic feature extraction and selection processes involved. This paper highlights various deep learning algorithms exploited for the classification of ECG signals into CAD, MI, and CHF conditions. The Convolutional Neural Network (CNN), followed by combined CNN and Long Short-Term Memory (LSTM) models, appear to be the most useful architectures for classification. A 16-layer LSTM model was developed in our study and validated using 10-fold cross-validation. A classification accuracy of 98.5% was achieved. Our proposed model has the potential to be a useful diagnostic tool in hospitals for the classification of abnormal ECG signals.

1. Introduction

Cardiovascular disease (CVD) is the leading cause of death globally. In 2012, 17.5 million deaths attributable to CVD were reported worldwide, accounting for 31% of all deaths. Of these, approximately 7.4 million deaths were due to coronary artery disease (CAD) [1]. In 2013, it was reported that 1 in every 7 Americans died due to CAD [2]. CAD is primarily the result of atherosclerosis, in which fibrofatty plaques develop and thicken within the wall of the coronary arteries, leading to stenosis of the coronary lumen [3–5]. CAD that is undiagnosed and/or untreated may progress and lead to complications. Composed of lipids contained within a luminal surface fibrous cap, an advanced or "vulnerable" atherosclerotic plaque can rupture suddenly. The contents spill into the coronary lumen, precipitating acute thrombosis, luminal occlusion and interruption of myocardial blood flow, which results in acute myocardial infarction (MI). [8,9]. The resultant MI induces chronic adverse cardiac remodeling that potentially leads to the development of congestive heart failure (CHF). Hence, timely diagnosis of CAD and MI is imperative; otherwise left ventricular function may become impaired. The electrocardiographic (ECG) signal is typically altered in established MI. In contrast, ECG perturbations in early CAD may be subtle and are easily missed and/or misinterpreted [6]. Clinically, the ECG is the most commonly used diagnostic tool for CAD because it is non-invasive and inexpensive. As ECG signals possess small amplitudes and short durations, measured in millivolts and milliseconds respectively, interpretation of these signals may suffer from wide inter- and intra-observer variabilities [7]. Automated diagnostic systems utilizing machine learning techniques may overcome these limitations [36]. Traditional machine learning techniques involve manual extraction and selection of features, which are cumbersome. In

☆ This article belongs to the Special Issue: Deep learning Methods for Medical Applications

* Corresponding author at: Department of Electronics and Computer Engineering, Ngee Ann Polytechnic, 599489, Singapore.
E-mail address: aru@np.edu.sg (U.R. Acharya).

https://doi.org/10.1016/j.artmed.2019.101789
Received 12 September 2019; Received in revised form 6 November 2019; Accepted 31 December 2019

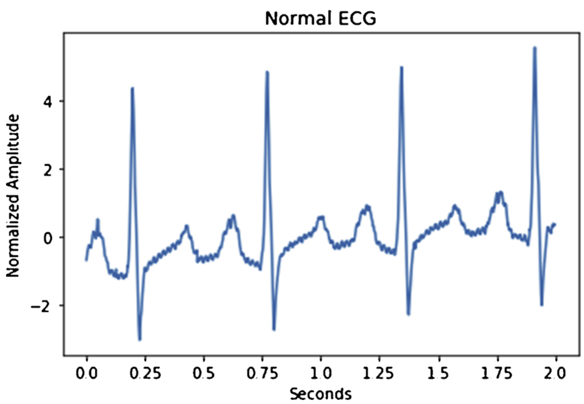

Fig. 1. Normal ECG signal.

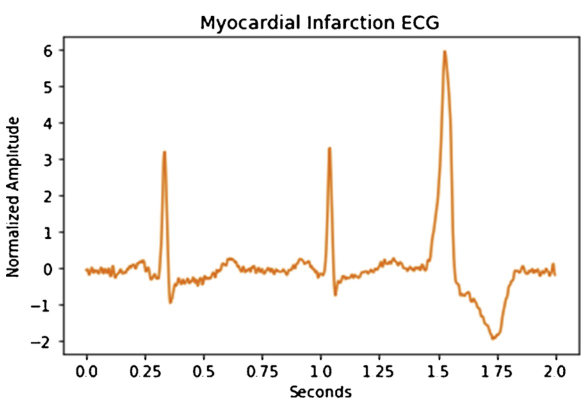

Fig. 2. Typical MI ECG signal.

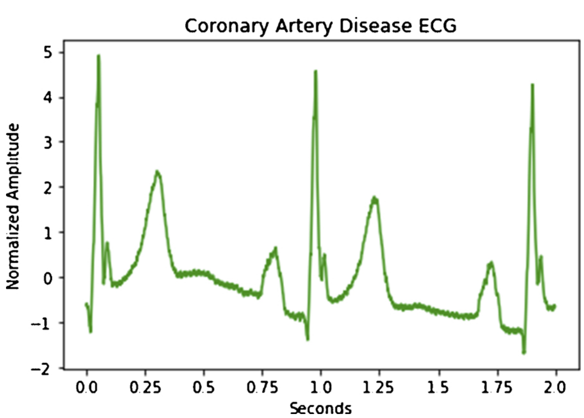

Fig. 3. Typical CAD ECG signal.

contrast, deep learning systems automatically extract and select significant features, and are the preferred method used in extant disease diagnosis applications [37–40]. In this paper, the characterization of three cardiac abnormalities (CAD, MI, CHF) using deep learning algorithms is discussed. Figs. 1–4 depict the typical ECG signals (the isolated ECGs may not show classical patterns) in normal, MI, CAD, and CHF subjects. Conventional machine learning techniques have been used for the detection of MI [43,44,51–54], CAD [55–58,61–63] and CHF [59]. These methods are laborious and require the extraction of best performing features manually to obtain the highest performance. Hence, the deep learning algorithms were used in this work.

2. Deep learning

Deep learning is a subset of machine learning in which a large dataset is often used to train the network. Significant features are created through each successive concealed layer of neurons as the network learns the input data. The Artificial Neural Network (ANN) is built upon several concealed layers, delivering a deep structure. The ANN is the most basic algorithm in deep learning, wherein synthetic neurons are the very essence of the neural network [10]. The neurons are connected with weights, in which the weighted sum is computed once data have been sent to the input layer. The bias from each neuron is subsequently added to the weighted sum of inputs. The activation function determines the activation of a neuron. Once a neuron is activated, it passes information to other neurons in the successive layers, until the penultimate layer. Once a neuron activated in the output layer tallies with the input digit, the weights and biases are continually adjusted to ensure that the network is well-trained [11].

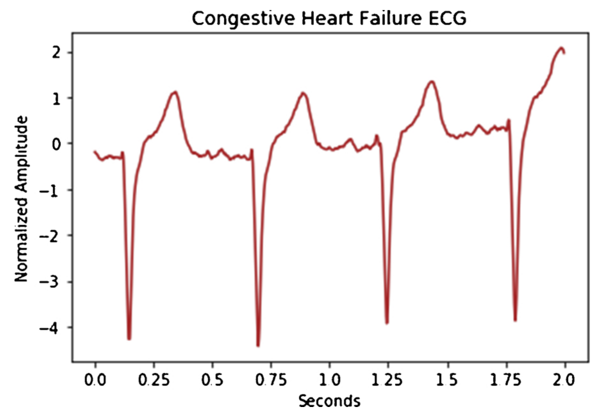

Fig. 4. Typical CHF ECG signal.

The Convolutional neural network (CNN) is a type of deep learning model that is commonly used in image and data analyses, as well as for classification of disease. It comprises three main layers: input, hidden, and output layers. Some models contain more layers, including non-convolutional layers. The hidden layers, known as the convolutional layers, form the heart of the CNN model. Different sized kernels are used in the convolutional layer to deduce the input, after which various feature maps are concatenated for analysis. The features that are created are used for classification in the successive layers [12]. The deeper the layers, the better the kernels become at detecting or classifying data. CNN is trained using the backpropagation algorithm [13], with the weights continually adjusted to reduce errors for optimum training performance.

Long Short-Term Memory (LSTM) is another model commonly utilized for the classification of physiological signals [14]. LSTM is a gated architecture that comprises blocks of memory cells, through which signals flow. It encompasses three gates: input, forget, and output gates. These gates control input activations into memory cells, reset the cells' memory, and control output flow into the network [15]. Mapping from an input x is calculated to form an output y by computing the unit activations of the network. In the network, the symbols t, $t-1$, $t+1$ denote the present, previous, and successive block values, respectively, while h and y represent the cell state and output values, respectively. The model works by retaining crucial information of previous states and building upon them. LSTM models are expedient for automatic feature extraction, as shown in earlier studies [16].

The autoencoder uses an unsupervised algorithm to train the network. The encoders are arranged together to form a deeper network. Three main steps are employed to train the model. First, a series of encoders are trained layer-by-layer using unsupervised data. Second, the last layer is trained with supervised data. Finally, the backpropagation algorithm is incorporated for refining the whole network [17]. In the first step, coding and decoding steps are applied. Unlabeled inputs are encoded and the inputs are reconstructed accurately. During the coding and decoding phases, identical weights are used to encode the feature and reconstruct the output. The loss function calculates the

Table 1a
Summarized studies involving deep learning for normal versus MI detection using ECG signals (2-class).

Authors	Techniques	Number of participants	Results
Safdarian et al. [27], 2014.	ANN with RBF function T-wave and total integral features K-NN PNN MLP Naïve Bayes	MI: 290 patients	Classification of MI Naïve Bayes; Accuracy: 94.74% Classification and localisation of MI PNN; Accuracy: 76.67%
Kora et al. [26], 2015	Improved Bat algorithm SVM classifier k-NN classifier LMNN SCGNN	Normal: 52 healthy MI: 148 patients	IBA + k-NN Accuracy: 65.1% IBA + SVM Accuracy: 76.74% IBA + SCGNN Accuracy: 87.90% IBA + LMNN Accuracy: 98.90%
Acharya et al. [24], 2017	1-dimensional CNN network Daubechies wavelet (6 mother wavelet) 10-fold cross-validation	Normal: 52 healthy MI: 148 patients	Accuracy with noise: 93.53% Accuracy without noise: 95.22%
Diker et al. [22], 2017	RFE, k-NN, ANN classifiers 10-fold cross-validation	Normal: 52 healthy MI: 148 patients	Accuracy: 80.60% Sensitivity: 86.58% Specificity: 64.71%
Reasat et al. [20], 2018	CNN architecture Geometric separability index Euclidian distance	Normal: 52 healthy MI: 148 patients	Accuracy: 84.54% Sensitivity: 85.33% Specificity: 84.09%
Lui et al. [25], 2018	CNN + LSTM Network 10-fold cross-validation	Normal: 52 healthy MI: 148 patients	Positive predictive value: 97.20%% Sensitivity: 92.40% Specificity: 97.70% F1 score: 94.60%
Strodthoff et al. [19], 2019	CNN network 10-fold cross-validation	Normal: 52 healthy MI: 127 patients	Sensitivity: 93% Specificity: 89.7%
Feng et al. [21], 2019	16-layer CNN coupled with LSTM network Wavelet transform 10-fold cross-validation	Normal: 52 healthy MI: 148 patients	Accuracy: 95.54% Sensitivity: 98.2% Specificity: 86.5% F1 score: 96.8%
Baloglu et al. [36], 2019	10-layer CNN model 12 leads	Normal: 52 healthy MI: 148 patients	Accuracy: 99.78%
Han et al. [49], 2019	Multi-lead residual neural network Fusion of features 5-fold validation	Normal: 52 healthy(80 recordings) MI: 112 patients(113 recordings)	Accuracy: 99.92% F1 score: 99.94%
Liu et al. [50], 2019	CNN combined with BLSTM Class-based 5-fold cross-validation	Normal: 52 healthy MI: 148 patients	Intra patient strategy Accuracy: 99.90% Sensitivity: 99.97% Specificity: 99.54%

information lost during input construction. A reconstruction with minimal loss value is almost identical to the original input [18].

Table 1a summarizes studies that involve deep learning for the detection of MI (2-class). Strodthoff et al. [19] built a CNN network for the classification of MI. Ten-fold cross-validation was employed to evaluate the system performance. High specificity and sensitivity values of 89.7% and 93.3% were obtained, respectively. Reasat et al. [20] developed a CNN network and extracted 84 features from the filters. The developed model was tested on one patient and trained on 81 patients. Metric scores were used to evaluate the performance of the model, yielding an accuracy of 84.54%. Application of a geometric separability index and Euclidean distance revealed that the extracted features showed good discriminating power. Feng et al. [21] developed an architecture comprising CNN and LSTM models. After the signals were pre-processed, oversampling was used to balance the healthy data. Ten-fold validation was employed during training of the model to evaluate its robustness. An accuracy of 95.5% was achieved with an F1 score of 96.8%. Diker et al. [22] combined ANN, Recursive Feature Eliminator (RFE), and kNN (k-nearest neighbor) classifiers after extracting a total of eleven statistical and structural features. Ten-fold validation was used to evaluate the proposed system, yielding an accuracy of 80.6%. Acharya et al. [24] developed an eleven-layer deep learning model and used two datasets to train and validate it. One dataset was denoised, while noise was retained in the other dataset. The signals were segmented and normalized before being input to the network. Ten-fold validation was used to assess the system performance, wherein relatively high accuracies of 93.5% and 95.2% were obtained for signals with and without noise, respectively. Lui et al. [25] combined CNN and LSTM models, and developed a classifier to distinguish MI from normal ECG signals. One layer of the CNN model was replaced by a LSTM layer, causing the classification sensitivity to improve by 28% compared with using only the CNN model. The developed system was evaluated using 10-fold validation, achieving high sensitivity and specificity values of 92.4% and 97.7%, respectively. Kora et al. [26] employed the improved Bat algorithm (IBA) for feature extraction after the signals were pre-processed. The significant features were then input to the Scalar Conjugate Gradient Neural Network (SCG NN), k-NN, and SVM classifiers. The results were compared with that of the Levenberg-Marquardt Neural Network (LMNN). The proposed technique of using the Bat algorithm coupled with LMNN outperformed the other classifiers, achieving the highest accuracy of 98.9%. Safdaraian et al. [27] compared the performance of the Naïve Bayes, Probabilistic Neural Network (PNN), and k-NN classifiers and ANN and Multilayer Perceptron (MLP) Neural Networks with T-wave and total integral features. The Naïve Bayes classifier outperformed the other classifiers with an accuracy of 94.7% for the classification of MI, while PNN demonstrated to classify and localize MI most accurately, yielding an accuracy of 76.7%. Baloglu et al. [36] generated a 10-layer CNN model for the classification of 12-lead ECG signals. 70% of the data was used for training, 15% for validation, and another 15% for testing of the model. High classification results of 99.8% were obtained for both leads V4 and V5. Han et al. [49] explored using the multi-lead residual neural

Table 1b
Summarized studies involving deep learning for normal versus CAD detection using ECG signals (2 class).

Authors	Techniques	Number of participants	Results
Acharya et al. [28], 2017	11-layer CNN Net 1 Net 2 10-fold validation	Normal: 40 healthy CAD: 7 patients	Net 1 Accuracy: 94.95% Sensitivity: 93.72% Specificity: 95.18% Net 2 Accuracy: 95.11% Sensitivity: 91.13% Specificity: 95.88%
Altan et al. [31], 2017	Deep Belief Network Hilbert-Huang transform 10-fold validation	Normal: 25 healthy CAD: 60 patients	Accuracy: 98.05% Sensitivity: 98.88% Specificity: 96.02%
Caliskan et al. [32], 2017	Autoencoders + Softmax classifier 4 datasets 10-fold validation	CAD: 303 patients	Switzerland dataset Accuracy: 92.20%
Tan et al. [29], 2018	8-layer LSTM + CNN network Blindfold technique Non-subject specific, subject-specific	Normal: 40 healthy CAD: 7 patients	Accuracy: 99.85%
Miao et al. [30], 2018	DNN based on deeper multilayer perceptron Diagnostic accuracy	Coronary heart disease: 303 patients	Accuracy: 83.67% Sensitivity: 93.51% Specificity: 72.86%
Oh et al. [45], 2018	CNN-LSTM 10-fold cross-validation	48 recordings from 47 subjects (arrhythmia)	Accuracy: 98.10% Sensitivity: 97.50% Specificity: 98.70%
Acharya et al. [60], 2018	11 layered CNN model 10-fold validation	105 signals (arrhythmia)	Accuracy: 93.18% Sensitivity: 95.32% Specificity: 91.04%
Gao et al [46], 2019	LSTM with focal loss	93371 ECG beats (arrhythmia)	Highest accuracy of 99.26% for denoised data.
Yildirim et al [47], 2019	LSTM Deep coded features	100 022 signals(5 beat type, arrhythmia)	Accuracy: > 99%
Pawiak et al. [48], 2019	Deep genetic ensemble of classifiers Spectral power density 10-fold cross-validation	29 subjects(744 segments, arrhythmia)	Accuracy: 99.37% Sensitivity: 94.62% Specificity: 99.66%

network coupled with a feature fusion technique with 12 leadECG recordings. Five-fold cross validation was used to validate the system. A high classification accuracy of 99.92% and F1 score of 99.94% were achieved with the intra-patient scheme. Liu et al. [50] developed a hybrid network comprising CNN coupled with bidirectional long short-term memory models (BLSTM). Class-based 5-fold validation was used to evaluate the proposed system, which achieved a high accuracy of 99.90% with the intra-patient strategy.

Table 1b summarizes studies that involve deep learning for the detection of CAD (2-class system). Acharya et al. [28] developed two 11-layer CNN networks for the classification of normal and CAD ECG signals. Net 1 and Net 2 were used to classify ECG signals of 2- and 5-second duration, respectively. Ten-fold cross-validation was employed to assess the performance of both architectures. High accuracies of 95.0% and 95.1% were obtained for Nets 1 and 2, respectively. Tan et al. [29] built an 8-layer deep learning model comprising CNN and LSTM networks. Non-subject and subject specific validations were employed to evaluate the proposed technique, wherein 10% of the data was used for training, 90% for testing, and 15 subjects' data were used for training and the rest for testing. A high classification accuracy of 99.9% was achieved with the blindfold technique. A deep neural network based on the MLP architecture was developed by Miao et al. [30], wherein the ECG signals were input. The system's performance was evaluated using a diagnostic accuracy value computed based on positive, false positive, true negative, and false negative values during training. An accuracy of 83.7% was yielded. Altan et al. [31] exploited the Deep Belief Network, which was used to classify the signals after features were extracted from short-term ECG signals using the Hilbert transform. Ten-fold validation was used to evaluate the performance of the model, achieving a high accuracy of 98.1%. Caliskan et al. [32] developed a classification system comprising two autoencoders coupled with the Softmax classifier. The developed system was used to classify signals in four different datasets, and its performance was compared against other classifiers including k-NN, SVM, Naïve Bayes, and Random forest. Ten-fold validation was used for evaluation. Compared with the other methods, the proposed system achieved the highest accuracy in each dataset, with the highest accuracy of 92.2% in the Switzerland dataset. Similar deep learning methods have also been used for arrhythmia detection. Oh et al. [45] developed a CNN-LSTM model and fed the acquired signals to it. Ten foldcross-validation was used to assess the performance of the model. A high accuracy of 98.10% was achieved for the classification of ECG signals to detect arrhythmia. Gao et al. [46] employed the LSTM model with focal loss, after pre-processing, for arrhythmia classification. 10% of the data was used for testing, while 90% and 10% of the remaining data were used for training and validation, respectively. The highest accuracy of 99.26% was achieved with the denoised data, as compared to data with noise. Yildirim et al. [47] investigated the LSTM network coupled with deep coded features, which yielded a high classification accuracy of over 99% for the detection of arrhythmia. Pawiak et al. [48] computed the power spectral density to intensify ECG signal features. These features were then input to the developed deep genetic ensemble of classifiers. Ten-fold validation was used to evaluate the performance of the system, which achieved a high accuracy of 99.37%. Acharya et al. [62] fed the acquired signals to the developed 11-layered CNN model. The performance of the model was evaluated using 10-fold cross-validation yielding an accuracy of 93.18%.

Table 1c summarizes studies that involve deep learning for the detection of CHF (2-class). Masetic et al. [23], extracted the Autoregressive (AR) Burg parameters from the signals and classified them using five classifiers: k-NN, SVM, random forest, ANN, and C4.5 decision tree classifiers. Ten-fold cross validation was used to evaluate the performances of the different classifiers. Random forest was reported to achieve the highest accuracy of 100%. Acharya et al. [33] developed an 11-layer CNN network and tested its performance using four different datasets. Standard 10-fold validation was utilized to evaluate the performance of the architecture. A highest accuracy of 99.0% was obtained for dataset B, which was the largest. Kwon et al. [34] employed a deep-learning algorithm for ECG-based diagnosis of heart failure (DEHF) and compared its performance with the logistic regression (LR) and random forest (RF) classifiers. The area under the receiver operating characteristic (AUROC) was used to gauge performance, yielding a higher value of 0.843 for the detection of heart failure with reduced ejection fraction, compared with AUROC values of 0.800 and 0.807 for LR and RF classifiers, respectively. Khade et al. [35] developed a system combining SVM and CNN. ECG signals were subjected to the SVM classifier coupled with the Boosted Decision Tree for classification of the type of heart failure. Ten-fold validation was used to validate the system, achieving an accuracy of 84.0%. CNN was employed to determine the severity of heart failure, achieving an accuracy of 88.3%.

3. Methodology

3.1. Data used

Lead II ECG signals employed in this study were acquired from healthy subjects and patients from four databases. Signals from 92 normal subjects, 7 CAD patients, 148 MI patients, and 15 CHF patients were obtained from the PTB Diagnostic ECG and Fantasia Databases, St.

Table 1c
Summarized studies involving deep learning for normal versus CHF detection using ECG signals (2 class).

Authors	Techniques	Number of participants	Results
Masetic et al. [23], 2016	Random forest, SVM, ANN, k-NN, C4.5 decision tree classifiers AR Burg features	Normal: 13 healthy CHF: 15 patients	Random forest Accuracy: 100%
Acharya et al. [33], 2019	11-layer CNN 4 datasets 10-fold validation	Dataset A Normal: 70 308 data CHF: 30 000 data Dataset B Normal: 110 000 data CHF: 30 000 data Dataset C Normal: 30 000 data CHF: 30 000 data Dataset D Normal: 30 000 data CHF: 30 000 data	Dataset B Accuracy: 98.97% Sensitivity: 99.01% Specificity: 98.87%
Kwon et al. [34], 2019	DEHF algorithm Logistic regression classifier Random Forest classifier	Normal: 19 836 patients CHF: 1391 HFrEF, 1538 HFmrEF patients	AUROC of DEHF: 0.843
Khade et al. [35], 2019	CNN Boosted Decision Tree SVM classifier 10-fold validation	CHF: 10 801 patients	Classification accuracy using SVM: 84% Heart failure severity measurement accuracy using CNN: 88.30%
Present study (4 class)	Detection of normal, MI, CAD and CHF CNN coupled with LSTM 10-fold validation	Normal: 92 CAD: 7 MI: 148 CHF: 15	Accuracy: 98.51% Sensitivity: 99.30% Specificity: 97.89% Positive predictive value: 97.33%

Petersburg Institute of Cardiological Technics 12-lead Arrhythmia Database, PTB Diagnostic ECG Database, and BIDMC Congestive Heart Failure Databases, respectively. Signals from the St. Petersburg and Fantasia databases were up-sampled to 1000 Hz to match the sampling frequency. The signals were segmented so that each segment consisted of a 2-second (2000 sample) window length. A total of 150,268 segments were used in this study. Table 2 details the breakdown of segments for each class.

3.2. CNN-LSTM deep learning architecture

Batch sizes of 10 and 60 epochs were used to develop the 16-layer CNN-LSTM model. Adam optimization parameters [42] exhibited a learning rate of 0.001. To improve generalization, dropout was applied to layers 14 and 16, with a dropout rate of 0.5. Bias was not introduced at the convolution layers, and weighted loss was employed for countering the class imbalance. The parameter details of different layers used to build the model are shown in Table 3. After the signals were input to the network, max pooling was employed after the convolution layers in every instance, in order to extract the optimal features for classification. Ten-foldcross-validation [41] was then incorporated to evaluate the performance of the developed model, whereby 80% of the training data was employed for training and 20% for validation. Fig. 5 presents the CNN-LSTM architecture, and its details are provided in Table 3. Each convolution layer was used to extract features from input signals to form feature maps for the subsequent layer. Max pooling layers were added each time after the convolution layers to sieve out the top features. The dropout layer was added to improve the generalization of features.

Table 2
Number of segments for each class.

Class	Number of segments
Normal	4 703(PTB) + 80 000(Fantasia)
MI	20 265
CAD	15 300
CHF	30 000

Table 3
Details of our developed model.

Layers	Type of layer	No. of neurons (output layer)	Kernel size	No. of filters	Stride
1	1d-convolution	1981 × 3	20 × 1	3	1
2	max-pooling	990 × 3	–	–	2
3	1d-convolution	981 × 6	10 × 1	6	1
4	max-pooling	490 × 6	–	–	2
5	1d-convolution	486 × 6	5 × 1	6	1
6	max-pooling	243 × 6	–	–	2
7	1d-convolution	239 × 5	5 × 1	6	1
8	max-pooling	119 × 6	–	–	2
9	1d-convolution	110 × 6	10 × 1	6	1
10	max-pooling	55 × 6	–	–	–
11	LSTM	10	–	–	–
12	dense	8	–	–	–
13	dropout	8	–	–	–
14	dense	8	–	–	–
15	dropout	8	–	–	–
16	dense	4	–	–	–

4. Results and discussion

High classification accuracy, specificity, sensitivity, and positive predictive values of 98.51%, 97.89%, 99.30%, 97.33%, respectively, were obtained with the proposed deep learning model. Fig. 6 depicts the accuracy of the network in terms of the accuracy against epoch plots. It is notable that the accuracy of the training set converges with that of the validation set, which implies that the developed model is robust. Fig. 7 presents the confusion matrix result based on classification. It is also apparent that the CNN-LSTM model is highly accurate in classifying the signals, as indicated by the low miscalculation rates of 0.02%, 0.02%, 0.03%, and 0.01%, respectively, for categorizing normal, MI, CAD, and CHF signals.

In an earlier study, Acharya et al. [43] investigated the use of the Discrete Wavelet Transform, Empirical Mode Decomposition, and Discrete Cosine Transform methods for the automated detection of CAD and MI on ECGs. In another study, Acharya et al. [44] studied the contourlet and shearlet transforms of ECGs for the classification of CAD, MI, and CHF. These two prior studies were foundational for the current

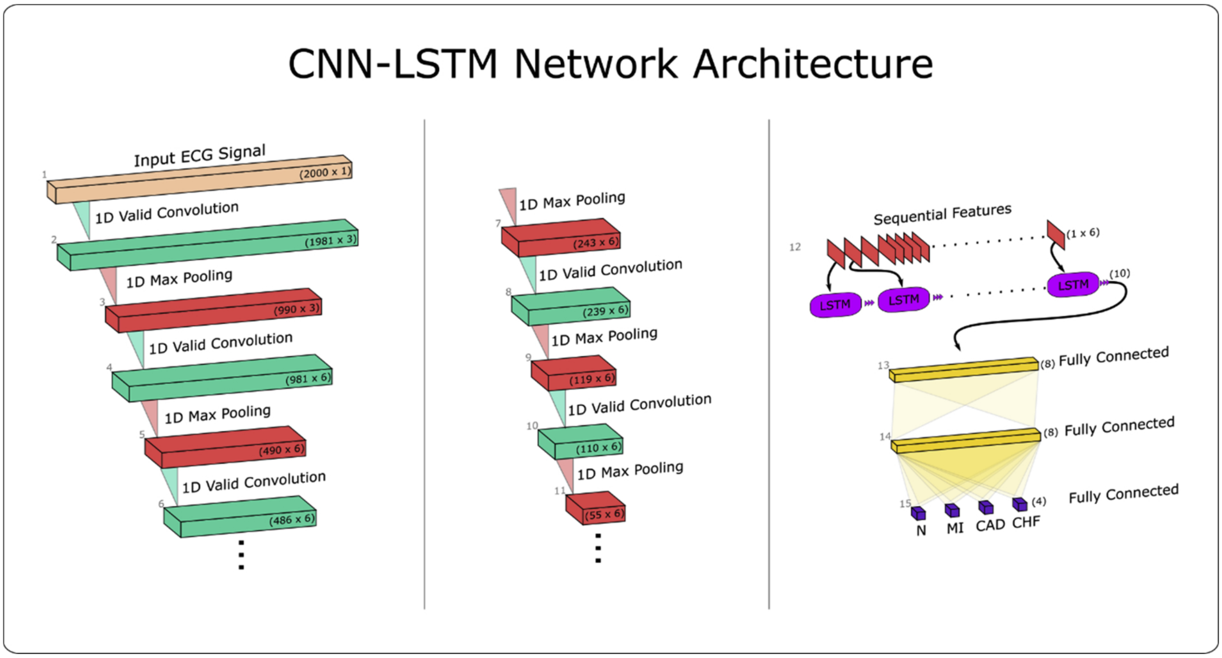

Fig. 5. Developed CNN-LSTM architecture.

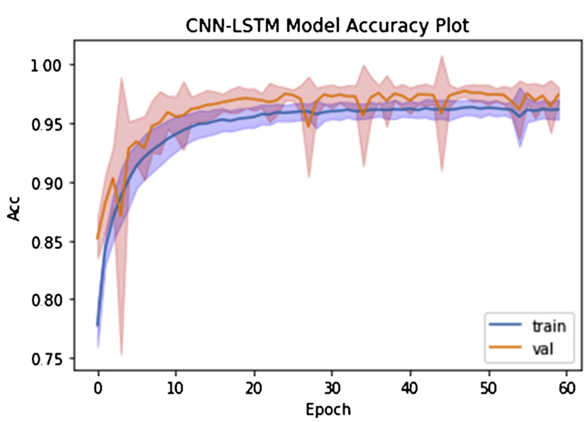

Fig. 6. Accuracy against varying epoch plot for CNN-LSTM model.

study. From Table 1, it is observed that the CNN algorithm has been employed predominantly for the classification of normal versus MI, CAD of CHF ECG signals. Ten-fold validation is commonly used to assess the models. Apart from using CNN alone, Tan et al. [29], Feng et al. [21], and Liu et al. [25] also employed LSTM coupled with CNN models, portraying this combined approach to be the next best deep learning model. Accordingly, the CNN-LSTM model was exploited in this study and 10-fold validation was used to gauge the performance of our model. The classification accuracy obtained from our model is higher compared with the results of Feng et al. [21]; and the sensitivity and specificity values achieved from our model are also higher than Lui et al. [25], owing to the larger data size used in our study compared with these other studies. Although Tan et al. [29] achieved a higher classification accuracy of 99.9%, only 47 subjects were studied. In our study, we used a larger data size, and obtained a higher accuracy compared with most studies, for instance, Acharya et al. [28], Altan et al. [31], Feng et al. [21], Reasat et al. [20], Xu et al. [23], Diker et al. [22] and Kora et al. [26], which employed other deep learning

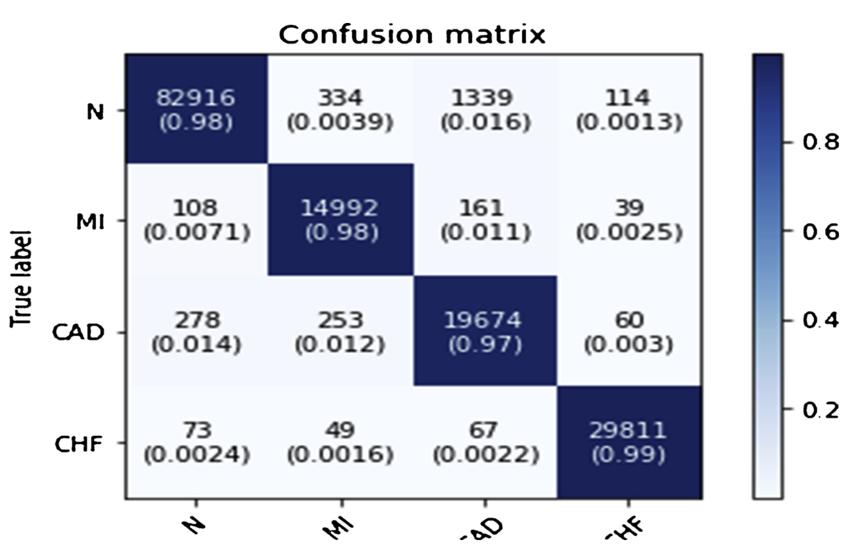

Fig. 7. Confusion matrix of the classified signals.

algorithms and small data sizes. Notably, while the studies in Table 1 all depict two-class results, ours is the first to discuss four-class results, which showed high classification accuracy. Our developed model attests to be robust in the classification of normal, CAD, MI, and CHF signals. Additionally, the developed system can be implemented in a wearable device (wireless patch) to enable monitoring of ECG signals for patient classification. The device would obtain the signals from a patient, which would be input to the developed deep learning model for classification. This model would be maintained on the cloud through the hospital server. Hence the diagnosis results would be available to the clinician through the hospital server within minutes. The developed system exhibits both advantages and disadvantages:

Advantages

I A high classification accuracy of 98.5% was achieved despite using signals with noise.
II The system established in this study is powerful as it has been validated using 10-fold.
III The entire data (large data) from PhysioNet was used.
IV With the developed model, three cardiac abnormalities can be detected.

Disadvantages

I Training of the model is time-consuming.
II Sizeable data is needed to train and test the model.
III Only a small data size of 7 was used to represent CAD patients.

5. Future work

In future work, we intend to develop a deep learning model that is better able to detect early stages of CAD, so that incident MI and CHF events can be averted. This would allow room for earlier diagnosis and timely treatment.

6. Conclusion

Cardiovascular diseases are the primary cause of death globally. When CAD is not identified during diagnostic testing, the disease can later manifest as MI and CHF. Cost-effective ECGs can be used to screen for CAD, so that treatment can be initiated to avert MI and CHF events. Application of deep learning algorithms to ECG interpretation can mitigate the pitfalls brought about by implementation of conventional machine learning algorithms. A 16-layer CNN-LSTM model was efficaciously used to classify CAD, MI, and CHF signals in our study, with a high precision rate of 98.5%. Ten-fold cross-validation provided confirmatory evidence as to the robustness of our proposed system. Hence it has the potential to be used as a diagnostic screening tool for CAD, which can lessen the workload of healthcare professionals. In the future, a deep learning model to detect early stages of CAD shall be developed.

References

[1] Mendis S. Global Status Report on Non-Communicable Diseases. Switzerland; 2014.
[2] Mozaffarian D, Benjamin EJ, Go AS, Arnett DK, Blaha MJ, Cushman M, et al. Heart disease and stroke statistics-2016 update a report from the american heart association. Circulation 2016;133(4):e38–48.
[3] Maximilian Buja L, Willerson JT. The role of coronary artery lesions in ischemic heart disease: insights from recent clinicopathologic, coronary arteriographic, and experimental studies. Hum. Pathol. 1987;18(5):451–61.
[4] Maximilian Buja L. Coronary artery disease: pathological anatomy and pathogenesis. Coronary Artery Disease. Springer; 2015. p. 1–20.
[5] Willerson JT. Ischemic Heart Disease: Clinical and Pathophysiological Aspects. Raven Press; 1982.
[6] Birnbaum Y, Wilson JM, Fiol M, De Luna AB, Eskola M, Nikus K. ECG diagnosis and classification of acute coronary syndromes. Ann. Noninvasive Electrocardiol. 2014;vol. 19(no. 1):4–14.
[7] Martis RJ, Acharya UR, Adeli H. Current methods in electrocardiogram characterization. Comput. Biol. Med. 2014;48:133–49.
[8] Liu B, Liu J, Wang G, Huang K, Li F, Zheng Y, et al. A novel electrocardiogram parameterization algorithm and its application in myocardial infarction detection. Comput. Biol. Med. 2015;61:178–84.
[9] Masetic Z, Subasi A. Congestive heart failure detection using random forest classifier. Comput. Methods Programs Biomed. 2016;130:54–64.
[10] Kotu V, Deshpande B. Deep learning. Data Sci. 2019;22(4):307–42.
[11] Lecun Y, Bengio Y, Hinton G. Deep learning. Nature 2015;vol. 521(no. 7553):436–44.
[12] Faust O, Hagiwara Y, Jen T, Shu O, Acharya UR. Computer Methods and Programs in Biomedicine Deep learning for healthcare applications based on physiological signals : a review. Comput. Methods Programs Biomed. 2018;161:1–13.
[13] Bouvrie J. Notes on convolutional neural networks. In Pract 2006:47–60.
[14] Hopfield JJ. Neural networks and physical systems with emergent collective computational abilities. Proc. Natl. Acad. Sci. U. S. A. 1982;vol. 79(no. 8):2554.
[15] Hochreiter S, Schmidhuber J. Long short-term memory. Neural Comput. 1997;9(November (no. 8)):1735–80.
[16] Naul B, Bloom J, Perez F, van Der Walt S. A recurrent neural network for classification of unevenly sampled variable stars. Nat. Astron. 2018;vol. 2(no. 2):151–5.
[17] Zia T, Zahid U. Long short-term memory recurrent neural network architectures for Urdu acoustic modeling. Int. J. Speech Technol. 2019;vol. 22(no. 1):21–30.
[18] Hinton GE, Salakhutdinov RR. Reducing the dimensionality of data with neural networks. Science 2006;vol. 313(no. 5786):504.
[19] Strodthoff N, Strodthoff C. Detecting and interpreting myocardial infarction using fully convolutional neural networks. Physiol. Meas 2019;40(1):1–11.
[20] Reasat T, Shahnaz C. Detection of inferior myocardial infarction using shallow convolutional neural networks. 5th IEEE Reg. 10 Humanit. Technol. Conf. 2017 2018. R10-HTC 2017, Vol. 2018-January, No. Imi, Pp. 718–721.
[21] Feng K, Pi X, Liu H, Sun K. Myocardial infarction classification based on convolutional neural network and recurrent neural network. Appl Sci (Basel) 2019;9(9):1–12.
[22] DİKER A, CÖMERT Z, AVCI E. A diagnostic model for identification of myocardial infarction from electrocardiography signals. Bitlis Eren Univ. J. Sci. Technol. 2017;7(2):132–9.
[23] Masetic Z, Subasi A. Congestive heart failure detection using random forest classifier. Comput. Methods Programs Biomed. 2016;vol. 130:54–64.
[24] Acharya UR, Fujita H, Oh SL, Hagiwara Y, Tan JH, Adam M. Application of deep convolutional neural network for automated detection of myocardial infarction using ECG signals. Inf. Sci. (Ny) 2017;vol. 415–416:190–8.
[25] Lui HW, Chow KL. Multiclass classification of myocardial infarction with convolutional and recurrent neural networks for portable ECG devices. Inf. Med. Unlocked 2018;13(June):26–33.
[26] Kora P, Kalva SR. Improved Bat algorithm for the detection of myocardial infarction. Springerplus 2015;vol. 4(no. 1):1–18.
[27] Safdarian N, Dabanloo NJ, Attarodi G. A new pattern recognition method for detection and localization of myocardial infarction using t-wave integral and total integral as extracted features from one cycle of ECG signal. J Biomed Sci Eng 2014;7(10):818–24.
[28] Acharya UR, Fujita H, Lih OS, Adam M, Tan JH, Chua CK. Automated detection of coronary artery disease using different durations of ECG segments with convolutional neural network. Knowl. Based Syst. 2017;132:62–71.
[29] Tan JH, et al. Application of stacked convolutional and long short-term memory network for accurate identification of CAD ECG signals. Comput. Biol. Med 2018;94(November 2017):19–26.
[30] Miao KH, Miao JH. Coronary heart disease diagnosis using deep neural networks. Int. J. Adv. Comput. Sci. Appl. 2018;vol. 9(no. 10):1–8.
[31] Altan G, Allahverdi N, Kutlu Y. Diagnosis of coronary artery disease using deep belief networks. Eur. J. Eng. Nat. Sci. 2017;2(1):29–36.
[32] Caliskan A, Yuksel ME. Classification of coronary artery disease data sets by using a deep neural network. EuroBiotech J. 2017;vol. 1(no. 4):271–7.
[33] Acharya UR, et al. Deep convolutional neural network for the automated diagnosis of congestive heart failure using ECG signals. Appl. Intell 2019;vol. 49(no. 1):16–27.
[34] Kwon J, et al. Development and validation of deep-learning algorithm for electrocardiography-based heart failure identification. Korean Circ. J. 2019;49(7):629.
[35] Khade S, Subhedar A, Choudhary K, Deshpande T, Kulkarni U. A system to detect heart failure using deep learning techniques. Int. Res. J. Eng. Sci. Technol. Innov. 2019;6(June):384–7.
[36] Baloglu UB, Talo M, Yildirim O, Tan RS, Acharya UR. Classification of myocardial infarction with multi-lead ECG signals and deep CNN. Pattern Recognit. Lett. 2019;vol. 122:23–30.
[37] Acharya UR, Fujita H, Oh SL, Raghavendra U, Tan JH, Adam M, et al. Automated identification of shockable and non-shockable life-threatening ventricular arrhythmias using convolutional neural network. Future Gener. Comput. Syst. 2018;79:952–9.
[38] Oh SL, Hagiwara Y, Raghavendra U, Yuvaraj R, Arunkumar N, Murugappan M, et al. A deep learning approach for Parkinson's disease diagnosis from EEG signals. Neural Comput. Appl. 2018:1–7.
[39] Acharya UR, Oh SL, Hagiwara Y, Tan JH, Adeli H, Subha DP. Automated EEG-based screening of depression using deep convolutional neural network. Comput. Methods Programs Biomed. 2018;161:103–13.
[40] Oh SL, Vicnesh J, Ciaccio EJ, Yuvaraj R, Acharya UR. Deep convolutional neural network model for automated diagnosis of schizophrenia using EEG signals. Appl. Sci. 2019;9(14):2870.
[41] Duda RO, Hart PE, Stork DG. Pattern Classification. second edition New York: John Wiley and Sons; 2001.

[42] Kingma DP, Ba JL. Adam: A Method for Stochastic Optimization. 2015. p. 1–15.
[43] Acharya UR, et al. Automated characterization and classification of coronary artery disease and myocardial infarction by decomposition of ECG signals: a comparative study. Inf. Sci. (Ny). 2017;377:17–29.
[44] Acharya UR, et al. Automated characterization of coronary artery disease, myocardial infarction, and congestive heart failure using contourlet and shearlet transforms of electrocardiogram signal. Knowl. Based Syst. 2017;132:156–66.
[45] Oh SL, Ng EYK, Tan RS, Acharya UR. Automated diagnosis of arrhythmia using combination of CNN and LSTM techniques with variable length heart beats. Comput. Biol. Med. 2018;vol. 102(no. April):278–87.
[46] Gao J, Zhang H, Lu P, Wang Z. An effective LSTM recurrent network to detect arrhythmia on imbalanced ECG dataset. Res. Article 2019;2019.
[47] Yildirim O, Baloglu UB, Tan R-S, Ciaccio EJ, Acharya UR. A new approach for arrhythmia classification using deep coded features and LSTM networks. Comput. Methods Programs Biomed. 2019;vol. 176:121–33.
[48] Pławiak P, Acharya UR. Novel deep genetic ensemble of classifiers for arrhythmia detection using ECG signals. Neural Comput. Appl. 2019(no. March).
[49] Han C, Shi L. ML – ResNet : a novel network to detect and locate myocardial infarction using 12 leads ECG. Comput. Methods Programs Biomed. 2019;185.
[50] Liu W, Wang F, Huang Q, Chang S, Wang H, He J. MFB-CBRNN: a hybrid network for MI detection using 12-lead ECGs. IEEE J. Biomed. Heal. Inf. 2019:1.
[51] Acharya UR, et al. Automated detection and localization of myocardial infarction using electrocardiogram: a comparative study of different leads. Knowl. Based Syst. 2016;99(2016):146–56.
[52] Kumar M, Pachori RB, Acharya UR. Automated diagnosis of myocardial infarction ECG signals using sample entropy in flexible analytic wavelet transform framework. Entropy 2017;vol. 19(no. 9).
[53] Jayachandran ES, Joseph P, Acharya R. Analysis of myocardial infarction using discrete wavelet transform. J. Med. Syst 2010;34(6):985–92.
[54] Sharma M, Tan RS, Acharya UR. A novel automated diagnostic system for classification of myocardial infarction ECG signals using an optimal biorthogonal filter bank. Comput. Biol. Med 2018;vol. 102(no. July):341–56.
[55] Acharya UR, et al. Application of higher-order spectra for the characterization of Coronary artery disease using electrocardiogram signals. Biomed. Signal Process. Control 2017;vol. 31(no. January):31–43.
[56] Kumar M, Pachori RB, Acharya UR. Characterization of coronary artery disease using flexible analytic wavelet transform applied on ECG signals. Biomed. Signal Process. Control 2017;vol. 31:301–8.
[57] Acharya UR, et al. Entropies for automated detection of coronary artery disease using ECG signals: a review. Biocybern. Biomed. Eng. 2018;38(2):373–84.
[58] Sharma M, Rajendra Acharya U. A new method to identify coronary artery disease with ECG signals and time-Frequency concentrated antisymmetric biorthogonal wavelet filter bank. Pattern Recognit. Lett. 2019;vol. 125:235–40.
[59] Bhurane AA, Sharma M, San-Tan R, Acharya UR. An efficient detection of congestive heart failure using frequency localized filter banks for the diagnosis with ECG signals. Cogn. Syst. Res. 2019;vol. 55(no. January):82–94.
[60] Acharya UR, et al. Automated identification of shockable and non-shockable life-threatening ventricular arrhythmias using convolutional neural network. Futur. Gener. Comput. Syst. 2018;vol. 79:952–9.
[61] Sharma M, Tan RS, Acharya UR. Detection of shockable ventricular arrhythmia using optimal orthogonal wavelet filters. Neural Comput. Appl. 2019.
[62] Sharma M, Singh S, Kumar A, San Tan R, Acharya UR. Automated detection of shockable and non-shockable arrhythmia using novel wavelet-based ECG features. Comput. Biol. Med. 2019;vol. 115:103446.
[63] SM K, Rajendra AU, Oliver F, Lim M. Analysis of cardiac signals using spatial filling index and time-frequency domain. Biomed. Eng. Online 2004;vol. 3(no. 1):30.