ECG Heartbeat Classification: A Deep Transferable Representation

Mohammad Kachuee, Shayan Fazeli, Majid Sarrafzadeh
University of California, Los Angeles (UCLA)
Los Angeles, USA

Abstract—Electrocardiogram (ECG) can be reliably used as a measure to monitor the functionality of the cardiovascular system. Recently, there has been a great attention towards accurate categorization of heartbeats. While there are many commonalities between different ECG conditions, the focus of most studies has been classifying a set of conditions on a dataset annotated for that task rather than learning and employing a transferable knowledge between different tasks. In this paper, we propose a method based on deep convolutional neural networks for the classification of heartbeats which is able to accurately classify five different arrhythmias in accordance with the AAMI EC57 standard. Furthermore, we suggest a method for transferring the knowledge acquired on this task to the myocardial infarction (MI) classification task. We evaluated the proposed method on PhysionNet's MIT-BIH and PTB Diagnostics datasets. According to the results, the suggested method is able to make predictions with the average accuracies of 93.4% and 95.9% on arrhythmia classification and MI classification, respectively.

Index Terms—ECG, deep learning, transfer learning, heartbeat, myocardial infraction

I. INTRODUCTION

ECG is widely used by cardiologists and medical practitioners for monitoring the cardiac health. The main problem with manual analysis of ECG signals, similar to many other time-series data, lies in difficulty of detecting and categorizing different waveforms and morphologies in the signal. For a human, this task is both extensively time-consuming and prone to errors. Note that the proper diagnosis of cardiovascular diseases is of paramount importance since these are the cause of death for about one-third of all deaths around the globe [1]. For instance, millions of people experience irregular heartbeats which can be lethal in some cases. Therefore, accurate and low-cost diagnosis of arrhythmic heartbeats is highly desirable [2].

To address the problems raised with the manual analysis of ECG signals, many studies in the literature explored using machine learning techniques to accurately detect the anomalies in the signal [3], [4]. Most of these approaches involve a preprocessing phase for preparing the signal (e.g., passing it through band-pass filters, etc). Afterwards, the handcrafted features which are mostly statistical summarizations of signal windows are extracted from these signals and used in further analysis for the final classification task. As for the inference engine, conventional machine learning approaches for ECG analysis include Support Vector Machines, multi-layer perceptrons, decision trees, etc. [5], [6], [7]

These handcrafted features provide us with an acceptable representation of the signal, based on recent machine learning studies, automated feature extraction and representation methods are proven to be more scalable and are capable of making more accurate predictions. An end-to-end deep learning framework allows the machine to learn the features that are best suited to the specific task that it is dedicated to carry out [8], [9], [10]. This approach provides us with a more accurate representation of ECG signal using which the machine can compete with a human cardiologist in analyzing the signal [11]. Deep learning approaches, however, contain a tremendously large amount of variables which require massive amounts of data to be trained.

One way of dealing with the need to a massive amount of data is the concept of knowledge transfer between different tasks. In computer vision, as an example, ImageNet dataset along with the state of the art deep learning models have been used to transfer knowledge between different image understanding tasks [12]. As another example, it has been shown that different sentence categorization tasks can share a considerable amount of sentence understanding [13]. On the other hand, there has been limited uses of transfer learning in health informatics. For example, Alaa *et al.* [14] have used the parameters of a Gaussian expert process trained on patients with stable conditions for patients with deteriorating conditions.

In this paper, we propose a novel framework for ECG analysis that is able to represent the signal in a way that is transferable between different tasks. For this to happen, we describe a deep neural network architecture which offers a considerable capacity for learning such representations. This network has been trained on the task of arrhythmia detection for learning which it is plausible to assume that the model needs to learn most of the shape-related features of the ECG signal. Also, we have a large amount of labeled data for this task, which makes it easy to train a network with a large amount of parameters. Furthermore, we show that the signal representation learned from this task is successfully transferable to the task MI prediction using ECG signals. This method allows us to use these deep representations to share knowledge between ECG recognition tasks for which enough information may not be available for training a deep architecture.

The rest of this paper is organized as follows. Section II explains the datasets used in this study. Section III presents the

TABLE I: Summary of mappings between beat annotations and AAMI EC57 [18] categories.

Category	Annotations
N	• Normal • Left/Right bundle branch block • Atrial escape • Nodal escape
S	• Atrial premature • Aberrant atrial premature • Nodal premature • Supra-ventricular premature
V	• Premature ventricular contraction • Ventricular escape
F	• Fusion of ventricular and normal
Q	• Paced • Fusion of paced and normal • Unclassifiable

Fig. 1: An example of a 10s ECG window and an extracted beat from it.

proposed method. Section IV presents results of the suggested method on different task and comparison of them with other works in the literature. Finally, Section V concludes the paper.

II. DATASETS

In this paper, we use PhysioNet MIT-BIH Arrhythmia and PTB Diagnostic ECG Databases as data source for labeled ECG records [15], [16], [17]. Furthermore, we demonstrate that the knowledge learned from the former database can be successfully transferred for training inference models for the latter. In all of our experiments, we have used ECG lead II re-sampled to the sampling frequency of $125Hz$ as the input.

The MIT-BIH dataset consists of ECG recordings from 47 different subjects recorded at the sampling rate of $360Hz$. Each beat is annotated by at least two cardiologists. We use annotations in this dataset to create five different beat categories in accordance with Association for the Advancement of Medical Instrumentation (AAMI) EC57 standard [18]. See Table I for a summary of mappings between beat annotations in each category.

The PTB Diagnostics dataset consists of ECG records from 290 subjects: 148 diagnosed as MI , 52 healthy control, and the rest are diagnosed with 7 different disease. Each record contains ECG signals from 12 leads sampled at the frequency of 1000Hz. In this study we have only used ECG lead II, and worked with MI and healthy control categories in our analyses.

III. METHODOLOGY

A. Preprocessing

As ECG beats are inputs of the proposed method we suggest a simple and yet effective method for preprocessing ECG
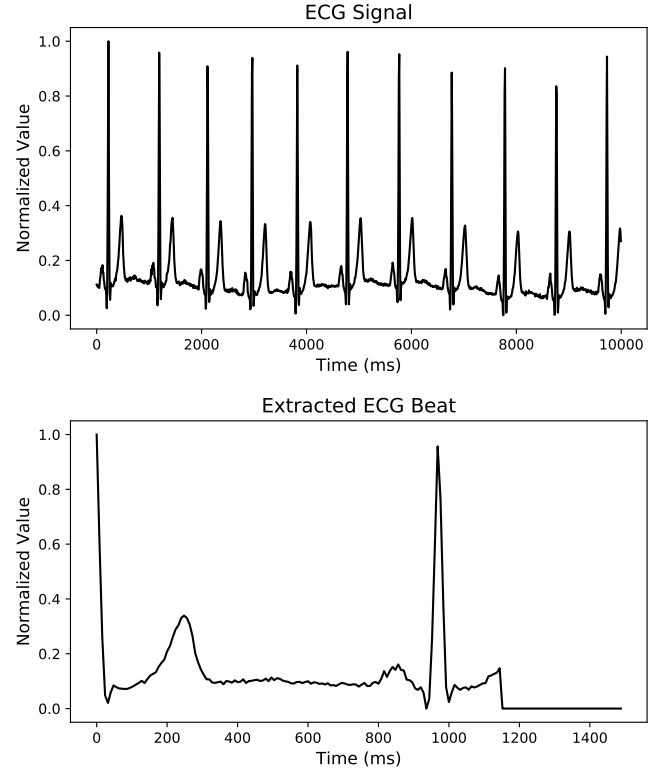
signals and extracting beats. The steps used for extracting beats from an ECG signal are as follows (see Fig. 1):

1) Splitting the continuous ECG signal to $10s$ windows and select a $10s$ window from an ECG signal.
2) Normalizing the amplitude values to the range of between zero and one.
3) Finding the set of all local maximums based on zero-crossings of the first derivative.
4) Finding the set of ECG R-peak candidates by applying a threshold of 0.9 on the normalized value of the local maximums.
5) Finding the median of R-R time intervals as the nominal heartbeat period of that window (T).
6) For each R-peak, selecting a signal part with the length equal to $1.2T$.
7) Padding each selected part with zeros to make its length equal to a predefined fixed length.

It is worth mentioning that the suggested beat extraction method is simple and effective in extracting R-R intervals from signals with different morphologies. For example, we have not used any form of filtering or any processing that makes any assumption about the signal morphology or spectrum. Additionally, all the extracted beats have identical lengths which is essential for being used as inputs to the subsequent processing parts.
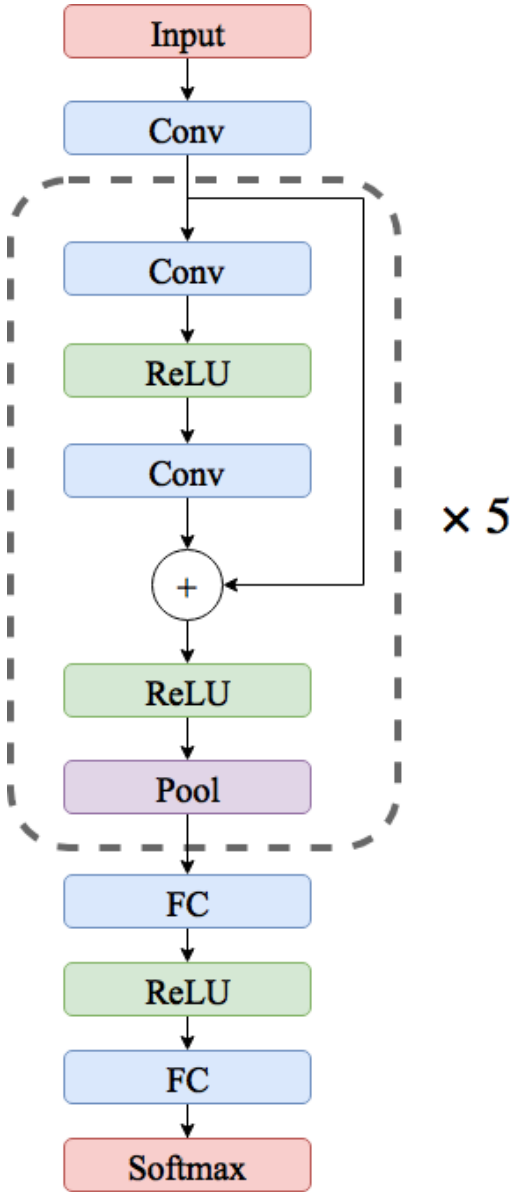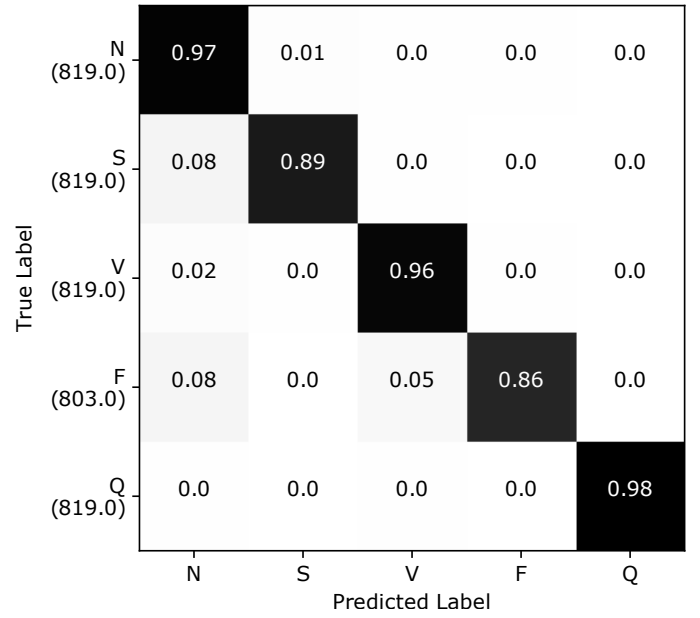
Fig. 2: Architecture of the proposed network.

Fig. 3: Confusion matrix for heartbeat classification on the test set. Total number of samples in each class is indicated inside parenthesis. Numbers inside blocks are number of samples classified in each category normalized by the total number of samples and rounded to two digits.

B. Training the Arrhythmia Classifier

In this paper we suggest training a convolutional neural network for classification of ECG beat types on the MIT-BIH dataset. The trained network not only can be used for the purpose of beat classification, but also in the next section we show that it can be used as an informative representation of heartbeats.

Fig. 2 illustrates the network architecture proposed for the beat classification task. Extracted beats, as explained in Section III-A, are used as inputs. Here, all convolution layers are applying 1-D convolution through time and each have 32 kernels of size 5. We also use max pooling of size 5 and stride 2 in all pooling layers. The predictor network consists of five residual blocks followed by two fully-connected layers with 32 neurons each and a softmax layer to predict output class

probabilities. Each residual block contains two convolutional layers, two ReLU nonlinearities [19], a residual skip connection [20], and a pooling layer. In total, the resulting network is a deep network consisting of 13 weight layers.

C. Training the MI Predictor

After training the network suggested in Section III-B, we use the output activations of the very last convolution layer as a representation of input beats. Here, we use this representation as input to a two layer fully-connected network with 32 neurons at each layer to predict MI. It is noteworthy to mention that during the training for the MI prediction task, we freeze the weights for all other layers aside from the last two. In other words, we only train the last two network layers and use the learned representation of Section III-B.

D. Implementation Details

In all experiments, TensorFlow computational library [21] is used for model training and evaluation. Cross entropy loss on the softmax outputs is used as the loss function. For training the networks, we used Adam optimization method [22] with the learning rate, beta-1, and beta-2 of 0.001, 0.9, and 0.999, respectively. Learning rate is decayed exponentially with the decay factor of 0.75 every 10000 iterations. Training all the networks took less than two hours on a GeForce GTX 1080Ti processor.

TABLE II: Comparison of heartbeat classification results.

Work	Approach	Average Accuracy (%)
This Paper	**Deep residual CNN**	**93.4**
Acharya *et al.* [23]	Augmentation + CNN	93.5
Martis *et al.* [24]	DWT + SVM	93.8
Li *et al.* [25]	DWT + random forest	94.6

TABLE III: Comparison of MI classification results.

Work	Accuracy (%)	Precision (%)	Recall (%)
This Paper[1]	**95.9**	**95.2**	**95.1**
Acharya *et al.* [27][1]	93.5	92.8	93.7
Safdarian *et al.* [28][1]	94.7	—	—
Kojuri *et al.* [29][2]	95.6	97.9	93.3
Sun *et al.* [30][3]	—	82.4	92.6
Liu *et al.* [31][3]	94.4	—	—
Sharma *et al.* [26][3]	96	99	93

[1]: PTB dataset, ECG lead II
[2]: dataset collected by authors, 12-lead ECG
[3]: PTB dataset, 12-lead ECG

IV. Results

A. Arrhythmia Classification and learning the representation

We evaluated the arrhythmia classifier of Section III-B on 4079 heartbeats (about 819 from each class) that are not used in the network training phase. Note that the dataset is being augmented to reach a balance in the number of beats in each category. Fig. 3 presents the confusion matrix of applying the classifier on the test set. As it can be seen from this figure, the model is able to make accurate predictions and distinguish different classes. all 5 classes well done, except S and F (less than 90%)

Table II presents the average accuracy of the proposed method and compares it with other relevant methods in the literature. While suggesting a predictor for MIT-BIH is not the sole purpose of this study, according to the results, the accuracies achieved in this paper are competitive to the state of the art methods. The main reason behind this might be the fact that we have used residual connections in our network architecture which allows us to train deeper networks compared to using traditional convolutional architectures.

B. MI Classification using the learned representation

We have trained our MI predictor using the learned representations, and took 80% of the PTB dataset as our training set. We have used the remaining 20% to test our model. Table III presents a comparison between the average accuracy, precision, and recall of the proposed method for MI classification and other work in the literature. The performance of the proposed method is better than all other works except the method suggested by Sharma *et al.* [26] that reports higher accuracy and precision values. However, it noteworthy to mention that Sharma *et al.* use 12-lead ECG as opposed to us using only the lead II.

C. Visualization of the learned representation

In order to visualize the learned representation, we have used t-SNE visualization method [32] to map high-dimensional vector created by the last convolutional layer to the 2D space. In a nutshell, t-SNE creates a mapping such that the joint probability of data-points appearing close to each other in the high-dimensional space is similar to the same probability distribution in the low-dimensional mapped points.

Fig. 4a illustrates the visualization of the learned representation on the MIT-BIH dataset samples. As it can be seen from this figure, data-points from different classes are easily separable using the learned representation. Fig. 4b shows the visualization of the MI classification task on the PTB samples using the representation trained on MIT-BIH. It can be inferred from this figure that the transferred representation for the beat classification task is able to provide a reasonable separation for the MI classification task. It should be noted that here we only use class labels to colorize the plots and other than this we do not use sample labels in the visualizations.

V. Conclusion

In this study we have presented a method for ECG heartbeat classification based on a transferable representation. Specifically, we have trained a deep convolutional neural network with residual connections for the arrhythmia classification task and shown that the representation learned for this task can be used as a base to train accurate classifiers for the classification of MI. According to the results, the suggested method is able to make predictions on both tasks with accuracies comparable to the state of the art methods in the literature. Furthermore, we visualized the learned representation using t-SNE method and illustrated the effectiveness of the proposed approach.

References

[1] 2018. [Online]. Available: http://www.who.int/mediacentre/factsheets/fs317/en/

[2] H. Society, "Heart diseases and disorders," 2018. [Online]. Available: https://www.hrsonline.org/Patient-Resources/Heart-Diseases-Disorders

[3] A. Esmaili, M. Kachuee, and M. Shabany, "Nonlinear cuffless blood pressure estimation of healthy subjects using pulse transit time and arrival time," *IEEE Transactions on Instrumentation and Measurement*, vol. 66, no. 12, pp. 3299–3308, 2017.

[4] A. E. Dastjerdi, M. Kachuee, and M. Shabany, "Non-invasive blood pressure estimation using phonocardiogram," in *Circuits and Systems (ISCAS), 2017 IEEE International Symposium on*. IEEE, 2017, pp. 1–4.

[5] O. T. Inan, L. Giovangrandi, and G. T. Kovacs, "Robust neural-network-based classification of premature ventricular contractions using wavelet transform and timing interval features," *IEEE Transactions on Biomedical Engineering*, vol. 53, no. 12, pp. 2507–2515, 2006.

[6] O. Sayadi, M. B. Shamsollahi, and G. D. Clifford, "Robust detection of premature ventricular contractions using a wave-based bayesian framework," *IEEE Transactions on Biomedical Engineering*, vol. 57, no. 2, pp. 353–362, 2010.

[7] M. Kachuee, M. M. Kiani, H. Mohammadzade, and M. Shabany, "Cuffless blood pressure estimation algorithms for continuous health-care monitoring," *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 4, pp. 859–869, 2017.

[8] U. R. Acharya, S. L. Oh, Y. Hagiwara, J. H. Tan, M. Adam, A. Gertych, and R. S. Tan, "A deep convolutional neural network model to classify heartbeats," *Computers in Biology and Medicine*, vol. 89, pp. 389–396, 2017.
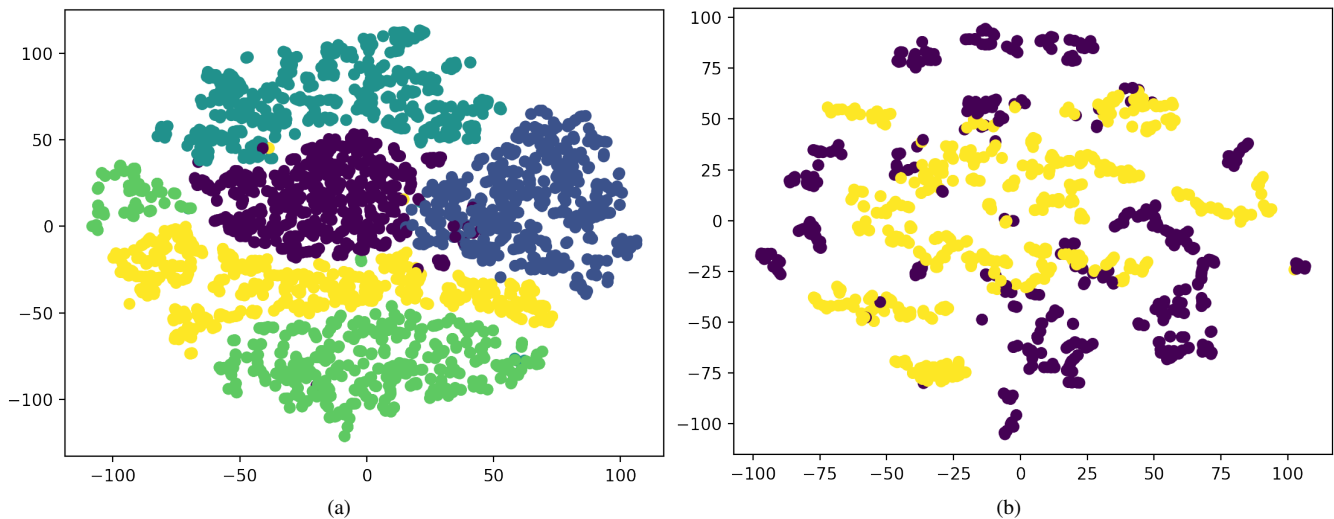
Fig. 4: t-SNE visualization of the learned representation: (a) samples from MIT-BIH for ECG beat classification (b) samples from PTB dataset for MI classification. Labels for each task are indicated with colors (best viewed in color).

[9] S. Kiranyaz, T. Ince, and M. Gabbouj, "Real-Time Patient-Specific ECG Classification by 1-D Convolutional Neural Networks," *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 3, pp. 664–675, 2016.

[10] L. Jin and J. Dong, "Classification of normal and abnormal ECG records using lead convolutional neural network and rule inference," *Science China Information Sciences*, vol. 60, no. 7, 2017.

[11] P. Rajpurkar, A. Y. Hannun, M. Haghpanahi, C. Bourn, and A. Y. Ng, "Cardiologist-level arrhythmia detection with convolutional neural networks," *arXiv preprint arXiv:1707.01836*, 2017.

[12] M. Oquab, L. Bottou, I. Laptev, and J. Sivic, "Learning and transferring mid-level image representations using convolutional neural networks," *2014 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1717–1724, 2014.

[13] A. Conneau, D. Kiela, H. Schwenk, and y. Loïc Barrault and Antoine Bordes, booktitle=EMNLP, "Supervised learning of universal sentence representations from natural language inference data."

[14] A. M. Alaa, J. Yoon, S. Hu, and M. van der Schaar, "Personalized risk scoring for critical care prognosis using mixtures of gaussian processes," *IEEE Transactions on Biomedical Engineering*, vol. 65, no. 1, pp. 207–218, 2018.

[15] A. L. Goldberger, L. A. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. Stanley, "Physiobank, physiotoolkit, and physionet," *Circulation*, vol. 101, no. 23, pp. e215–e220, 2000.

[16] G. B. Moody and R. G. Mark, "The impact of the mit-bih arrhythmia database," *IEEE Engineering in Medicine and Biology Magazine*, vol. 20, no. 3, pp. 45–50, 2001.

[17] R. Bousseljot, D. Kreiseler, and A. Schnabel, "Nutzung der ekg-signaldatenbank cardiodat der ptb über das internet," *Biomedizinische Technik/Biomedical Engineering*, vol. 40, no. s1, pp. 317–318, 1995.

[18] A. for the Advancement of Medical Instrumentation *et al.*, "Testing and reporting performance results of cardiac rhythm and st segment measurement algorithms," *ANSI/AAMI EC38*, vol. 1998, 1998.

[19] V. Nair and G. E. Hinton, "Rectified linear units improve restricted boltzmann machines," in *Proceedings of the 27th international conference on machine learning (ICML-10)*, 2010, pp. 807–814.

[20] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.

[21] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin *et al.*, "Tensorflow: Large-scale machine learning on heterogeneous distributed systems," *arXiv preprint arXiv:1603.04467*, 2016.

[22] D. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.

[23] U. R. Acharya, S. L. Oh, Y. Hagiwara, J. H. Tan, M. Adam, A. Gertych, and R. San Tan, "A deep convolutional neural network model to classify heartbeats," *Computers in biology and medicine*, vol. 89, pp. 389–396, 2017.

[24] R. J. Martis, U. R. Acharya, C. M. Lim, K. Mandana, A. K. Ray, and C. Chakraborty, "Application of higher order cumulant features for cardiac health diagnosis using ecg signals," *International journal of neural systems*, vol. 23, no. 04, p. 1350014, 2013.

[25] T. Li and M. Zhou, "Ecg classification using wavelet packet entropy and random forests," *Entropy*, vol. 18, no. 8, p. 285, 2016.

[26] L. Sharma, R. Tripathy, and S. Dandapat, "Multiscale energy and eigenspace approach to detection and localization of myocardial infarction," *IEEE Transactions on Biomedical Engineering*, vol. 62, no. 7, pp. 1827–1837, 2015.

[27] U. R. Acharya, H. Fujita, S. L. Oh, Y. Hagiwara, J. H. Tan, and M. Adam, "Application of deep convolutional neural network for automated detection of myocardial infarction using ecg signals," *Information Sciences*, vol. 415, pp. 190–198, 2017.

[28] N. Safdarian, N. J. Dabanloo, and G. Attarodi, "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," *Journal of Biomedical Science and Engineering*, vol. 7, no. 10, p. 818, 2014.

[29] J. Kojuri, R. Boostani, P. Dehghani, F. Nowroozipour, and N. Saki, "Prediction of acute myocardial infarction with artificial neural networks in patients with nondiagnostic electrocardiogram," *Journal of Cardiovascular Disease Research Vol*, vol. 6, no. 2, p. 51, 2015.

[30] L. Sun, Y. Lu, K. Yang, and S. Li, "Ecg analysis using multiple instance learning for myocardial infarction detection," *IEEE transactions on biomedical engineering*, vol. 59, no. 12, pp. 3348–3356, 2012.

[31] B. Liu, J. Liu, G. Wang, K. Huang, F. Li, Y. Zheng, Y. Luo, and F. Zhou, "A novel electrocardiogram parameterization algorithm and its application in myocardial infarction detection," *Computers in biology and medicine*, vol. 61, pp. 178–184, 2015.

[32] L. v. d. Maaten and G. Hinton, "Visualizing data using t-sne," *Journal of Machine Learning Research*, vol. 9, no. Nov, pp. 2579–2605, 2008.