e-ISSN: 2458-8156

Journal Homepage: http://dergipark.gov.tr/ejens

DergiPark AKADEMİK

EJENS

European Journal of Engineering and Natural Sciences

CODEN: EJENS

Diagnosis of Coronary Artery Disease Using Deep Belief Networks

Gokhan Altan[1,*], Novruz Allahverdi[2], Yakup Kutlu[3]

[1]*Mustafa Kemal University, Department of Informatics, 31000, Hatay, Turkey.*

[2]*KTO Karatay University, Department ofComputerEngineering, 42020, Karatay/Konya, Turkey.*

[3]*Iskenderun Technical University, Department of ComputerEngineering, 31200, Iskenderun/Hatay, Turkey.*

**CorrespondingAuthor email: gokhan_altan@hotmail.com*

Publication Info

Paper received:
29 May 2016

Revised received:
15 October 2016

Accepted:
01 March 2017

Abstract

In this study, a decision-support system is presented to aid cardiologists during the diagnosis and to create a base for a new diagnosis system which separates two classes (CAD and no-CAD patients) using an electrocardiogram (ECG).
24 hour filtered ECG signals from PhysioNet were used. 15 second short-term ECG segments were extracted from 24 hour ECG signals to increase the number of samples and to provide a convenient transformation in a short period of time. The Hilbert-Huang Transform, which is effective on non-linear and non-stationary signals, was used to extract the features from short-term ECG signals. Instinct Mode Function (IMF) was extracted by applying Empirical Mode Decomposition to short-term ECG signals. The Hilbert Transform (HT) was applied to each IMF to obtain instantaneous frequency characteristics of the signal. Dataset was created by extracting statistical features from HT applied to IMF. Deep Belief Networks (DBN) which have a common use in Deep Learning algorithms were used as the classifier. DBN classification accuracy in the diagnosis of the CAD is discussed. The extracted dataset was tested using the 10-fold cross validation method. The test characteristics (sensitivity, accuracy and specificity) that are the basic parameters of independent testing in the medical diagnostic systems were calculated using this validation method. Short-term ECG signals of CAD patients and no-CAD groups were classified by the DBN with the rates of 98.05%, 98.88% and 96.02%, for accuracy, specificity and sensitivity, respectively.
The DBN model achieved higher accuracy rates than the Neural Network

Key words

CoronaryArteryDisease, CAD, DeepBelief Networks, DBN, Deep Learning Algorithm,Hilbert-HuangTransform

1. INTRODUCTION

An electrocardiogram (ECG) is a uniform signal that records electrical changes at certain intervals of the heartbeat. It is also known as a map which shows the electrical activity of the heart. The ECG signals (ECGs), which are recorded with the aid of electrodes placed at different parts of the body, can have different electrical charges at any time. Electrical charges specify leads in the ECG. Bipolar and unipolar leads occur as a result of the position of the electrodes [1]. The ECG is a type of biomedical signal that is widely used for a variety of diagnosis and monitoring of cardiac diseases [2]. The electrical charges that are obtained directly as the result of discomfort of the heart and segments and intervals between waves on the ECG are of a paramount importance for the identification of the cardiac abnormalities in a subject [3].

Cardiac diseases have a high mortality rate worldwide. One of the most common cardiac diseases is Coronary Artery Disease (CAD). It is usually known as atherosclerosis in which there is a cardiac abnormality because of the narrowing of the arteries that feed the heart in time. In this cardiac disease, plaque caused by cholesterol collects in the coronary arteries and over time congestion results due to dilation of plaque. As a result of biological conditions, the arteries fail to feed the heart and it disrupts the rhythmic systole activity of the heart [4], [5]. CAD in adults results in heart attacks or a congestive heart. In the diagnosis of CAD, there are many clinical trials like physical examinations from which ST-deviation is measured, lab tests, Electrocardiogram (ECG), echocardiogram, stress tests and electron beam computed tomography, coronary angiography and cardiac catheterization [4]. In the literature, a diagnosis of the CAD has worked with various signal-processing methods using the clinical dataset (age, sex, history, physical exam, cardiac risk factors, exercise stress test data, heart rate, etc.) on the rule-based fuzzy classification [6], Other methods include using the clinical attributes on data mining techniques [7]–[14]; the chronic conditions, risk factors, and laboratory results of data mining techniques [15]; the clinical attributes on data discretization, data partitioning and reduced error pruning [16]; the history and physical examination data and the diastolic heart sounds [17]; both the heart rate variability (HRV) features and the non-linear features (Poincare Plot, entropy, etc.) [18]–[22]; the P wave features on the ECG [7]; the Wavelet Packet Transformation (10 levels) [23]; the cardiac analysis of ST measures [24], [25]; the Empirical Mode Decomposition (EMD) and Teager Energy operator [26]; the Discrete Wavelet Transformation (DWT) [27], [28] and using the Principal Component Analysis (PCA) [28], [29]. In this study, the Hilbert-Huang Transform (HHT) is applied to filtered short-term ECGs from the PhysioNet database to extract features.

HHT is an adaptive method to analyze the nonlinear and non-stationary signals [30]. It has a common use in the biomedical signal analysis transformations due to these characteristics. The HHT is applied to the ECG for the diagnosis of Atrial Fibrillation [31] and Congestive Heart Failure [32]. In this study, the HHT would be applied to the filtered short-term ECGs for designing an effective statistical feature extraction model and classification of HHT-based statistical features using Deep Learning (DL) algorithms.

The DL is a comparatively new algorithm that is utilized to estimate the classification performances in many distributions such as speech recognition [33], [34], computer vision [35], [36], natural language processing [37]–[39], physiological data [40]–[42], biomedical datasets [43], [44] and non-linear signals [45], [46]. The DL aims to discover the multiple and deeper levels of distributions for a better classification performance. The basic concept of the DL is based on enhancing the classification performance using an artificial neural network model with multilayer hidden units. The most important difference of the DL is this: while both the multilayer neural network model and the DL has the same structure, it has at least two hidden layers and an unsupervised pre-training phase [47]. The depth of the DL is defined by the number of hidden layers on the model. The DL can also show a higher performance than neural networks with a small number of neurons. A fewer number of neurons on models can make it more convenient to calculate weights using supervised learning [48].

The aim of this study is to design a deep CAD diagnosis system that can be an alternative short-term ECG-based statistical feature based classification method to studies in the literature. 100 instances of 15 second ECG forms would be extracted from long-time ECGs in preprocessing. In this way, using short-term ECG forms would solve analysis problems and would enhance the number of samples up to 100x. The HHT would be applied to separated short-term ECGs and Instinct Mode Functions (IMFs) would be extracted. The system would extract statistical features from IMFs that are obtained applying the HHT. Each Instinct Mode Function (IMF) group obtained would be classified using the Deep Belief Networks (DBN) algorithm. Classification performances of the diagnosed subjects with or without the CAD would be examined.

2. MATERIALS AND METHODS

ECGs were used in the proposed diagnosis system. We preferred the moving window analysis technique and segmented long-time ECGs recorded into 15 second windows which were used in the preprocessing part. 15 second short-term ECGs were utilized in this study. IMFs were extracted applying HHT and statistical features were calculated for the obtained IMFs in feature extraction. Statistical features were classified using the DBN. A detailed description of the structure of the system is presented in the following sections.

2.1. Database

In the literature, different databases including different diagnosis systems were used for the diagnosis of the CAD. Clinical characteristics in particular [7]–[14] were used to separate subjects with or without CAD. Outside of the literature, the Long-Term ST Database [49] was used in the diagnosis of the CAD. The Long-Term ST Database contains 85 long-term ECGs from 80 subjects, chosen to exhibit a variety of events of ST segment changes. There are 25 subjects labeled as undiagnosed CAD patients and 60 subjects labeled as diagnosed CAD in this database. The individual recordings of the Long-Term ST Database are between 21 and 24 hours in duration.

2.2. Preprocessing

The information in a biomedical signal is unevenly distributed. We would like to call attention to the fact that all data records were from the PhysioNet databases as filtered long-term ECGs. Long-term ECGs may have too much noise while being recorded because of physical and recording conditions. The short-term ECGs usually have an ability to represent the Long-term ECG characteristics. The short-term ECG is a less affected form of represented Long-term ECG. That is why short-term ECGs were randomly segmented into 15 second short-term ECGs using the moving window analysis technique 100 times. In this way, the number of instances from each subject in the dataset could be increased by 100x.

2.3. Hilbert-Huang Transform

HHT is an effective analysis technique for non-linear signals. It has a common use in biomedical signals (ECG, EMG, EEG, etc.). It has a flexible mathematical formulation and is easily adaptable for various types of processes [32]. HHT is a two-stage transform. The first stage is EMD and the second stage is the Hilbert Spectrum Analysis (HSA). EMD extracts frequency-modulated signals that are named IMFs. After IMFs are extracted, HSA is applied to each IMF to calculate the instantaneous frequency and amplitude [50]. Considering all these characteristics, our work focuses on HHT analysis in short-term ECGs.

EMD is an algorithm that breaks down natural form non-linear signals without leaving the time domain. EMD assumes a random signal which consists of its own self-oscillation at different frequencies. Oscillations are symmetrical to the mean of the local minimum and the local maximum at a t time. EMD extracts Instinct Mode functions which are a complete and nearly orthogonal basis at different frequencies [31]. IMFs are all in the time-domain and of the same length as the original signal. EMD has a detailed formula in the literature [30]. In the formula, X represents the original signal, c represents extracted IMF and r_n represents the residual signal [30], [31].

$$X(t) = \sum_{j=1}^{n} c_{j+r_n} \tag{1}$$

After sifting through IMFs by obtaining a monotonic residual signal, the Hilbert Transform (HT) can be applied to each IMF to compute the instantaneous frequencies spectral analysis. The instantaneous frequencies give most important information about the signal characteristics. After performing the HT to each IMF component, the amplitude and frequency of each component as functions of time is [30]:

$$x(t) = \mathbb{R}\left\{\sum_{i=1}^{n} a_i\,(t) e^{j\int \omega(t)dt}\right\} \tag{2}$$

The frequency-time distribution of the amplitude is designated as the HSA, $H(\omega, t)$ and the marginal spectrum is $h(\omega)$as follows [30] :

$$h(\omega) = \int_0^t H(\omega, t)dt \tag{3}$$

2.4. Deep Belief Networks

The DBN is a machine learning algorithm become more popular because of its semi-supervised learning methods. The DBN has a two stage learning process: unsupervised learning followed by supervised learning. In the first stage, it evaluates weights and biases between visible and hidden layers using an unsupervised pre-training of stacked Restricted Boltzmann Machines (RBM). RBMs are stacked between two adjacent layers which are visible-hidden layers or hidden-hidden layers. RBMs are energy-based functions and have only connections between adjacent nodes. Weights and biases between hidden and visible layers are evaluated with the aid of the probability of greedy layer-wise method. At the second stage, pre-training is followed by supervised fine-tuning with weighted neurons and biases to improve parameters.

The DBN can be defined as a specialized model with many hidden layers of DL. The upper layers of the DBN may hold more detailed and descriptive features to identify the solution of diagnosing systems, whereas lower layers may not. The DBN has more important advantages than the classical neural networks such as achieving high performance with a small number of training sets, and having the ability of utilizing the connections between the features in deeper processes. In the supervised learning phase, weights and biases are updated using fine-tuning in which the gradient descent or ascent algorithms are used for improving the accuracies and sensitivities of models [51], [52]. The DBN is a probabilistic joint distribution of input vector x and the ℓ hidden layers as follows:

$$P(x, h^1, \ldots, h^\ell) = \left(\prod_{k=0}^{\ell-2} P(h^k \,|h^{k+1})\right) P(h^{\ell-1}, h^\ell) \tag{4}$$

$P(h^{\ell-1}, h^\ell)$is the probability of conditional distribution between the adjacentlayers and h^0 is the input vector.

The energy function of the state (h^{k-1}, h^k) is defined as:

$$E(h^{k-1}, h^k; \theta) = -\sum_{s=1}^{D_{k-1}} \sum_{t=1}^{D_k} w_{st}^k h_s^{k-1} h_t^k - \sum_{s=1}^{D_{k-1}} b_s h_s^{k-1} - \sum_{t-1}^{D_k} c_t\, h_t \tag{5}$$

where $\theta = (w_{st}, b, c)$ which are the parameters of the DBN;$w_{st}^k$is the weight between s^{th} neuron in the layerh^{k-1} and t^{th} neuron in the layer h^k; $b_s$is the s^{th} bias of layer h^{k-1} and $c_t$is the t^{th} bias of layer h^k. $D_k$is the number of neurons in the k^{th} layer. The probabilistic distribution of the energy function is:

$$P(h^{k-1}; \theta) = \frac{\sum_{h^k} \exp(-E(h^{k-1}, h^k; \theta))}{\sum_{h^{k-1}} \sum_{h^k} \exp(-E(h^{k-1}, h^k; \theta))} \tag{6}$$

After layer-wise unsupervised learning, calculated weights are refined using supervised learning based on gradient descent. This fine-tuning process updates w parameters for a better discriminative ability and for obtaining higher classification performances [53].

Accuracy (ACC), Specificity (SPE) and Sensitivity (SEN) are used to evaluate the performance of medical diagnosis systems. Calculation of these performance measurements are described in [22], [54].

3. RESULTS

Various biomedical signals and clinical characteristics are analyzed with lots of digital signal processing methods and data mining algorithms in medical diagnosis systems. Computer-based diagnosis systems may have support decision systems for improving clinicians' performance and classification performance. Computer-based diagnosis systems can also be enhanced using biomedical signal processing methods and can be utilized as an alternative or additional method to the clinical characteristics of the subjects. In this study, the ECG is used to diagnose the subjects with or without CAD. The DBN structure of the proposed diagnosis system is seen in Fig.1.

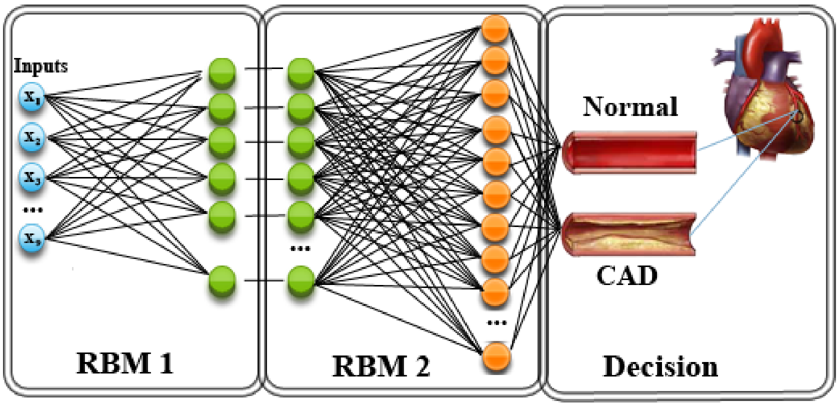

Figure 1. Structure of the DBN Classifier to diagnose CAD

Long-term ECGs may have much more noise than the short-term ECGs. Physical conditions such as coughing, changing the standing possession, instantaneous movements and recording conditions such as dislocation of the probe, etc. can handle noise in the long-term ECGs. The short-term ECG is less affected by these kinds of conditions. The short-term ECGs usually have an ability to represent the long-term ECG characteristics in most cardiac diseases. 15 second short-term ECGs were segmented from 85 long-term ECGs using the moving window analysis technique 100 times. The number of instances was increased to 8500 ECGs.

EMD was applied to each short-term ECG and the IMFs ranging in number from 10 to 14 were extracted for 8500 ECGs. Obtained IMFs can be seen in Fig.2. HT was applied to each IMF and the instantaneous frequencies spectral features were computed. HT allows deriving the analytic representation of a signal and includes phase features.

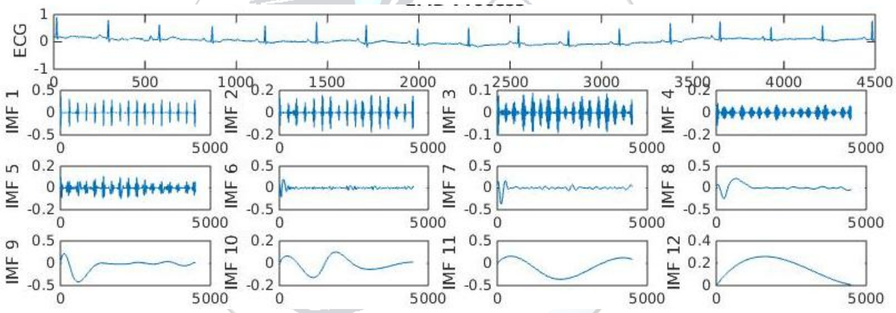

Figure 2. A randomly selected Short-term ECG and Extracted IMFs by applying EMD

HHT is used for feature extraction using the Hilbert Spectral Analysis in most of the studies. In this study, the IMFs that were extracted after the HHT process were used as the base of the features, but not directly used as the features. Statistical features (minimum (Min), maximum (Max), skewness (Skw), median, mean, Standard Deviation (SD), correlation (Corr), mode and energy) were calculated from each IMF for creating the diagnosis dataset. Each short-term ECG had a number of 9 statistical features multiplied by the number of the extracted IMF. The MATLAB statistical toolbox is used for calculation of statistical features. As it is seen in Table I, the highest five responsible features in the diagnosis of the CAD are Max and Min of the 3rd IMF, Max and Corr of the 4th IMF and Min of the 5th IMF. The lowest responsible features are mode values of IMFs.

Table I. Performance of each statistical feature for reduced the short-term ECG in diagnosis of the CAD

	Max	**Min**	**Corr**	**Energy**	**Mean**	**SD**	**Skw**	**Median**	**Mode**	**All**
SEN	85.05	80.98	78.72	60.67	52.70	57.10	43.50	39.77	21.05	96.02
SPE	92.56	81.40	62.52	78.52	73.20	30.24	22.52	5.36	28.00	98.88
ACC	87.26	81.11	73.95	65.92	58.73	49.20	37.33	29.65	23.09	98.05

The system was tested using a 10-fold cross validation. The dataset was randomly divided into 10 folds with the same number of subjects with and without the CAD. 9-folds of dataset were used for the training of the DBN classifier and one fold was used for testing the DBN. We subdivided the 9-folds of the dataset into 100 batches to speed-up and update weights step by step on the learning phase. The DBN that was utilized in the proposed diagnosis model has one input layer, 2 hidden layers and one output layer (Fig.1). The input layer has 9 input units for statistical features.

Greedy layer-wise pre-training is used in this model at the unsupervised learning stage of the DBN with 5 epochs. The DBN has 2 hidden layers with 100 hidden units for each. All feature sets were normalized to 0-1. The output layer has two outputs (subjects with and without CAD). To unfold the DBN to a neural network for the supervised learning stage of DBN, model parameters were selected by iterations. The learning rate is 2 and the activation function of the hidden layers on the supervised learning phase is the hyperbolic tangent function to avoid bias in the gradients and to have a stronger gradient. And a sigmoid output function was utilized. After training the DBN, test results were performed and the reliability and statistical performances of the model were calculated.

Table II. Comparison of the models on diagnosis of the CAD

Authors	ACC	SPE	SEN	Method	Classification	Data
Lee at. al[21]	85.00-90.00	-	-	HRV measurements	SVM	HRV
Dua et. al. [20]	89.50	-	-	HRV measurements	ANN	HRV
Kim et. al. [22]	75.00	-	-	HRV measurements	MDA	HRV
Giri et. al. [28]	96.80	93.70	100.00	DWT, LDA, PCA	GMM	HRV
Patidar et. al. [29]	99.70	99.80	99.60	Co-entropy, PCA	SVM	HRV
Arafat et. al. [24]	86.00	-	-	ST measures	Fuzzy Clustering	ECG
Alizadensani et. al. [25]	94.08	-	-	Q waveand ST measures	SVM	ECG
Thisstudy	98.05	98.88	96.02	HHT, Statistical Features	DBN	ECG

*MDA: Multiple Discriminant Analysis, SVM: Support Vector Machines, ANN: Artificial Neural Network, LDA: Linear Discriminant Analysis, GMM: Gaussian Mixture Model, HRV: Heart Rate Variability

As it is seen in Table II, when we compared the classification performances in literature, high accuracy rates were achieved using various classification methods on both ECG and HRV data. The highest classification performance on ECG data is achieved using the DBN on ECG based features. The achieved performance measurements that are achieved using the proposed method have a remarkable point in both HRV and ECG studies. Subjects with and without CAD were separated with a classification accuracy rate of 98.05%, a specificity of 98.88% and a sensitivity of 96.02% using statistical features from IMFs.

4. CONCLUSIONS

Linear and non-linear HRV features, various analysis techniques on ECG, heart sounds and clinical characteristics were used to evaluate diagnosis of the CAD. Different classification methods were utilized on these features to achieve high classification performances. In this study, the HHT that has a widespread utilization on non-linear signals was used to extract features from filtered ECGs. The proposed system diagnoses the subjects with or without the CAD. In this integrated system, DBN was used to classify the statistical features of IMFs. Accuracy, specificity and sensitivity achievements were used to evaluate system performance.

HRV are the Poincare plots, cross Corr, SD, arithmetic mean, Skw, kurtosis, and approximate entropy measurements between R waves which are extracted from 5 min short-term ECGs. Using HRV does not take into consideration the durations and intervals of the other waves except R waves on the ECGs. In this study, the ECG was used considering that all waves (whole ECG) may carry significant characteristics in diagnosis of the CAD. As seen in Table 1, an accuracy rate of 98.05%, a specificity rate of 98.88% and a sensitivity rate of 96.02% were achieved in the diagnosis of the CAD. It is difficult to compare the classification accuracies with the literature, because of different databases. Achieved performances show that the proposed method has an ability to separate the subjects with and without the CAD. Thus, the highest accuracy is achieved in the studies using ECG signals and most of the studies using HRV. The biggest advantage of the proposed method compared to HRV is using 15 second short-term ECG segments.

REFERENCES

[1]. M. Gabriel Khan, *Rapid ECG Interpretation(Contemporary Cardiology)*, 3rd edition. Humana Press, 2007.

[2]. U. F. Chan, W. W. Chan, S. H. Pun, M. I. Vai, and P. U. Mak, “Flexible implementation of front-end bioelectric signal amplifier using fpaa for telemedicine system,” in *Annual International Conference of the IEEE Engineering in Medicine and Biology - Proceedings*, 2007, pp. 3721–3724.

[3]. Y. Özbay and G. Tezel, “A new method for classification of ECG arrhythmias using neural network with adaptive

activation function," *Digit. Signal Process.*, vol. 20, no. 4, pp. 1040–1049, 2010.

[4]. "CAD Information." [Online]. Available: http://www.nhlbi.nih.gov/health/health-. [Accessed: 01-Jan-2016].

[5]. G. K. Hansson, "Inflammation, atherosclerosis, and coronary artery disease," *N. Engl. J. Med.*, vol. 352, no. 16, pp. 1685–1695, 2005.

[6]. K. Polat and S. Güneş, "A hybrid approach to medical decision support systems: Combining feature selection, fuzzy weighted pre-processing and AIRS," *Comput. Methods Programs Biomed.*, vol. 88, no. 2, pp. 164–174, 2007.

[7]. R. Yilmaz and R. Demirbag, "P-wave dispersion in patients with stable coronary artery disease and its relationship with severity of the disease," *J. Electrocardiol.*, vol. 38, no. 3, pp. 279–284, 2005.

[8]. İ. Babaoglu, O. Findik, and E. Ülker, "A comparison of feature selection models utilizing binary particle swarm optimization and genetic algorithm in determining coronary artery disease using support vector machine," *Expert Syst. Appl.*, vol. 37, no. 4, pp. 3177–3183, 2010.

[9]. Y. N. Devi and S.Anto, "An Evolutionary-Fuzzy Expert System for the Diagnosis of Coronary Artery Disease," *, An Evol. Fuzzy Expert Syst. Diagnosis Coron. Artery Dis.*, vol. 3, no. 4, pp. 1478–1484, 2014.

[10]. M. G. Tsipouras, T. P. Exarchos, D. I. Fotiadis, A. P. Kotsia, K. V. Vakalis, K. K. Naka, and L. K. Michalis, "Automated Diagnosis of Coronary Artery Disease Based on Data Mining and Fuzzy Modeling," *IEEE Trans. Inf. Technol. Biomed.*, vol. 12, no. 4, pp. 447–458, 2008.

[11] . Noor Akhmad Setiawan, P. A. Venkatachalam, and Ahmad Fadzil M.Hani, "Diagnosis of coronary artery disease using artificial intelligence based decision support system," *Proc. Int. Conf. Man-Machine Syst.*, 2009.

[12]. N. Ghadiri Hedeshi and M. Saniee Abadeh, "Coronary artery disease detection using a fuzzy-boosting PSO approach," *Comput. Intell. Neurosci.*, vol. 2014, 2014.

[13]. M. C. Colak, C. Colak, H. Kocatürk, S. Sağiroğlu, and I. Barutçu, "Predicting coronary artery disease using different artificial neural network models.," *Anadolu Kardiyol. Derg.*, vol. 8, no. 4, pp. 249–254, 2008.

[14] . A. Rajkumar and M. G. S. Reena, "Diagnosis Of Heart Disease Using Datamining Algorithm," *Glob. J. Comput. Sci. Technol.*, vol. 10, no. 10, pp. 38–43, 2010.

[15] . N. Lavesson, A. Halling, and M. Freitag, "Classifying the Severity of an Acute Coronary Syndrome by Mining Patient Data," *26th Annu. Work. Swedish Artif. Intell. Soc.*, vol. 35, pp. 55–63, 2009.

[16] . M. Shouman, T. Turner, and R. Stocker, "Using data mining techniques in heart disease diagnosis and treatment," *Japan-Egypt Conf. Electron. Commun. Comput.*, pp. 173–177, 2012.

[17] . M. Akay, W. Welkowitz, J. L. Semmlow, and J. Kostis, "Application of the ARMA method to acoustic detection of coronary artery disease," *Med. Biol. Eng. Comput.*, vol. 29, no. 4, pp. 365–372, 1991.

[18]. U. R. Acharya, O. Faust, V. Sree, G. Swapna, R. J. Martis, N. A. Kadri, and J. S. Suri, "Linear and nonlinear analysis of normal and CAD-affected heart rate signals," *Comput. Methods Programs Biomed.*, vol. 113, no. 1, pp. 55–68, 2014.

[19]. H. G. Lee, K. Y. Noh, and K. H. Ryu, "Mining biosignal data: Coronary artery disease diagnosis using linear and nonlinear features of HRV," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2007, vol. 4819 LNAI, pp. 218–228.

[20]. S. Dua, X. Du, S. Vinitha Sree, and V. I. Thajudin Ahamed, "Novel classification of coronary artery disease using heart rate variability analysis," *J. Mech. Med. Biol.*, vol. 12, no. 4, p. 1240017, 2012.

[21]. H. G. L. H. G. Lee, K. Y. N. K. Y. Noh, and K. H. R. K. H. Ryu, "A Data Mining Approach for Coronary Heart Disease Prediction using HRV Features and Carotid Arterial Wall Thickness," *2008 Int. Conf. Biomed. Eng. Informatics*, vol. 1, 2008.

[22]. H. C. W. Kim, S. Jin, Y. Park, "A study on development of multi-parametric measure of heart rate variability diagnosing cardiovascular disease," *IFMBE Proc. 14*, pp. 3480–3483, 2007.

[23]. M. Karimi, R. Amirfattahi, S. Sadri, and S. A. Marvasti, "Noninvasive detection and classification of coronary artery occlusions using wavelet analysis of heart sounds with neural networks," in *Medical Applications of Signal Processing, 2005. The 3rd IEE International Seminar on (Ref. No. 2005-1119)*, 2005, pp. 117–120.

[24]. S. Arafat, M. Dohrmann, and M. Skubic, "Classification of Coronary Artery Disease Stress ECGs using Uncertainty Modeling," in *2005 ICSC Congress on Computational Intelligence Methods and Applications*, 2005, pp. 1–4.

[25]. R. Alizadehsani, J. Habibi, M. J. Hosseini, H. Mashayekhi, R. Boghrati, A. Ghandeharioun, B. Bahadorian, and Z. A. Sani, "A data mining approach for diagnosis of coronary artery disease," *Comput. Methods Programs Biomed.*, vol. 111, no. 1, pp. 52–61, 2013.

[26]. Z. Zhao and C. Ma, "An Intelligent System for Noninvasive Diagnosis of Coronary Artery Disease with EMD-TEO and BP Neural Network," in *2008 International Workshop on Education Technology and Training & 2008 International Workshop on Geoscience and Remote Sensing*, 2008, pp. 631–635.

[27]. U. R. Acharya, S. V. Sree, M. Muthu Rama Krishnan, N. Krishnananda, S. Ranjan, P. Umesh, and J. S. Suri, “Automated classification of patients with coronary artery disease using grayscale features from left ventricle echocardiographic images,” *Comput. Methods Programs Biomed.*, vol. 112, no. 3, pp. 624–632, 2013.

[28]. D. Giri, U. Rajendra Acharya, R. J. Martis, S. Vinitha Sree, T. C. Lim, T. Ahamed, and J. S. Suri, “Automated diagnosis of Coronary Artery Disease affected patients using LDA, PCA, ICA and Discrete Wavelet Transform,” *Knowledge-Based Syst.*, vol. 37, pp. 274–282, 2013.

[29]. S. Patidar, R. B. Pachori, and U. Rajendra Acharya, “Automated diagnosis of coronary artery disease using tunable-Q wavelet transform applied on heart rate signals,” *Knowledge-Based Syst.*, vol. 82, pp. 1–10, 2015.

[30]. N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung, and H. H. Liu, “The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis,” *Proc. R. Soc. A Math. Phys. Eng. Sci.*, vol. 454, no. 1971, pp. 903–995, 1998.

[31]. U. Maji, M. Mitra, and S. Pal, “Automatic Detection of Atrial Fibrillation Using Empirical Mode Decomposition and Statistical Approach,” *Procedia Technol.*, vol. 10, pp. 45–52, 2013.

[32]. G. Altan, A. Yayik, Y. Kutlu, S. Yildirim, and E. Yildirim, “Analyse of Congestive Heart Failure Using Hilbert-Huang Transform,” *Dokuz Eylul Univ. Eng. Sci.*, vol. 16, pp. 94–103, 2014.

[33]. G. Hinton, L. Deng, D. Yu, G. Dahl, A. Mohamed, N. Jaitly, V. Vanhoucke, P. Nguyen, T. Sainath, and B. Kingsbury, “Deep Neural Networks for Acoustic Modeling in Speech Recognition,” *IEEE Signal Process. Mag.*, vol. 29, no. 6, pp. 82–97, 2012.

[34]. S. M. Siniscalchi, D. Yu, L. Deng, and C. H. Lee, “Exploiting deep neural networks for detection-based speech recognition,” *Neurocomputing*, vol. 106, pp. 148–157, 2013.

[35]. C. Farabet, C. Couprie, L. Najman, and Y. Lecun, “Learning hierarchical features for scene labeling,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 8, pp. 1915–1929, 2013.

[36]. P. Sermanet, K. Kavukcuoglu, S. Chintala, and Y. Lecun, “Pedestrian detection with unsupervised multi-stage feature learning,” in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2013, pp. 3626–3633.

[37]. Q. Le and T. Mikolov, “Distributed Representations of Sentences and Documents,” *Int. Conf. Mach. Learn. - ICML 2014*, vol. 32, pp. 1188–1196, 2014.

[38]. I. Sutskever, O. Vinyals, and Q. Le, “Sequence to sequence learning with neural networks,” *Adv. Neural Inf. …*, p. 9, 2014.

[39]. H. Lee, R. Grosse, R. Ranganath, and A. Y. Ng, “Convolutional deep belief networks for scalable unsupervised learning of hierarchical representations,” *Proc. 26th Annu. Int. Conf. Mach. Learn. ICML 09*, vol. 2008, pp. 1–8, 2009.

[40]. P. W. Mirowski, Y. LeCun, D. Madhavan, and R. Kuzniecky, “Comparing SVM and convolutional networks for epileptic seizure prediction from intracranial EEG,” in *Proceedings of the 2008 IEEE Workshop on Machine Learning for Signal Processing, MLSP 2008*, 2008, pp. 244–249.

[41]. P. W. Mirowski, D. Madhavan, and Y. Lecun, “Time-delay neural networks and independent component analysis for eeg-based prediction of epileptic seizures propagation,” in *Advancement of Artificial Intelligence Conference*, 2007, pp. 1892–1893.

[42]. D. Wang and Y. Shang, “Modeling Physiological Data with Deep Belief Networks.,” *Int. J. Inf. Educ. Technol.*, vol. 3, no. 5, pp. 505–511, 2013.

[43]. P. Drotár, J. Gazda, and Z. Smékal, “An experimental comparison of feature selection methods on two-class biomedical datasets,” *Comput. Biol. Med.*, vol. 66, pp. 1–10, 2015.

[44]. P. Tamilselvan and P. Wang, “Failure diagnosis using deep belief learning based health state classification,” *Reliab. Eng. Syst. Saf.*, vol. 115, pp. 124–135, 2013.

[45]. E. de la Rosa and W. Yu, “Randomized algorithms for nonlinear system identification with deep learning modification,” *Inf. Sci. (Ny).*, p. -, 2015.

[46]. M. Längkvist, L. Karlsson, and A. Loutfi, “A review of unsupervised feature learning and deep learning for time-series modeling,” *Pattern Recognit. Lett.*, vol. 42, no. 1, pp. 11–24, 2014.

[47]. Y. Bengio and O. Delalleau, “Justifying and generalizing contrastive divergence,” *Neural Comput.*, vol. 21, no. 6, pp. 1601–1621, 2009.

[48]. G. Hinton, G. Hinton, T. Sejnowski, and T. Sejnowski, *Learning and relearning in Boltzmann machines*, vol. 1. 1986.

[49]. A. L. Goldberger, L. A. N. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. . Stanley, “PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals,” *Circulation*, vol. 101, pp. 215–220, 2000.

[50]. N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung, and H. H. Liu, “The

empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 454, no. 1971. pp. 903–995, 1998.

[51]. G. E. Hinton, S. Osindero, and Y.-W. Teh, "A fast learning algorithm for deep belief nets.," *Neural Comput.*, vol. 18, no. 7, pp. 1527–54, 2006.

[52]. Y. Bengio, P. Lamblin, D. Popovici, and H. Larochelle, "Greedy Layer-Wise Training of Deep Networks," *Adv. Neural Inf. Process. Syst.*, vol. 19, no. 1, p. 153, 2007.

[53]. Y. Liu, S. Zhou, and Q. Chen, "Discriminative deep belief networks for visual data classification," *Pattern Recognit.*, vol. 44, no. 10–11, pp. 2287–2296, 2011.

[54]. R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*. 2000.