Machine Learning for Classification and Control of Cardiac Arrhythmias

Sutha Subbian
Assistant Professor
Department of Instrumentation Engg., MITCampus, Anna University
Chennai-44, Tamilnadu,India
e-mail: sutha_mani@hotmail.com

Divya Govindaraju
Research Scholar
Department of Instrumentation Engg.
MIT Campus, Anna University
Chennai-44, Tamilnadu,India
e-mail: gdivyaice @gmail.com

Nambi Narayanan.S
Teaching Fellow
Department of Instrumentation Engg.
MIT Campus, Anna University
Chennai-44, Tamilnadu,India
e-mail: vinoth2852@gmail.com

***Abstract*—Cardiac arrest has become a primary cause of sudden death. Commonly it is caused by cardiac arrhythmias such as Tachycardia and Bradycardia. The aim of the paper is to propose machine learning algorithm for classifying the arrhythmias accurately using Electrocardiograph (ECG) signals and Clinical data. Further, a suitable model-based control scheme is developed for controlling Bradycardia. Firstly, the best features are extracted from the data set and are used for classifications of cardiac arrhythmias using Convolution Neural Network (CNN), Support Vector Machine (SVM) and CNN-SVM (SVM). The classification accuracy is compared for the proposed Machine Learning Algorithms with training dataset and test dataset. Secondly, after classifying cardiac arrhythmia, suitable model is identified by developing various nonlinear models namely Nonlinear Auto Regressive Exogenous (NARX), Nonlinear Hammerstein-Wiener (HW) and Recurrent Neural Network (RNN) for cardiac vascular system. The performances of the developed models are compared, and best model is intended to design controller. Finally, model-based control scheme is developed using the best model and closed loop studies are carried out. The simulation studies show the feasibility of the proposed control scheme.**

Keywords-cardiac arrhythmias; classification; cardiovascular system; CNN; SVM; bradycardia;, NARX; HW; RNN

I. INTRODUCTION

In developing countries, most of the deaths are due to Cardiac Vascular Diseases (CVD) such as coronary heart diseases, congential heart diseases,stroke, hypertensive heart disease inflammatory heart disease, rheumatic heart disease and cardiac arrythmias. In CVDs, Cardiac arrhythmias are prevalent among people across all ages. It may be very difficult to diagnose arrhythmias by doctors with high accuracy using clinical data and ElectroCardioGraphy (ECG) signals. Researchers in Standford university proved that 14 types of arrhythmias can be accurately detected from ECG signal by deep learning algorithms and outperforms them for most of the cases. Eduardo Jose da S. Luz et.al [1], states the recent methods in the classification of arrhythmia based on the heartbeat signal obtain from the ECG. Further, the author has suggested for future research in fully automatic arrhythmia classification. Oliver Faust et al [2], reviewed various deep learning methods applied for different types of physiological signals in healthcare domain. The author has discussed the importance of huge dataset for early detection of diseases using the above signals. Ali Isin et al [3], an automated ECG arrhythmia diagnosis for classification of ECG signal was carried out depending on the cardiac condition of the patient. The author has taken the clinical data of the cardiac patients with normal, paced or right bundle branch block condition and further used deep learning algorithms for diagnosis. G. Sannino et al [4], the author has proposed a novel deep learning algorithm for classification of arrhythmia. The irregularities in the arrhythmia detection are addressed by applying deep networks. The author has developed a Deep Neural Network (DNN) for classification using tensor flow and google library. The author has reported the accuracy, sensitivity and specificity have improved compared with the literature. Shraddha Singh [5], has classified arrhythmia ECG beat based on Recurrent Neural Networks (RNN). The time-series analysis was given as an input to LSTM network. Further, the comparison of different RNN models was carried out for quantitative data. Weiyi Yang [6], the author has applied feature extraction using PCA (Principal Component Analysis). Further, the author has used SVM for improving the speed of the classification. The author concludes the above algorithm are well suited for noisy data and skewed data applicability. Felipe Alonso-Atienza et al [7], the author has discussed the early detection of arrhythmia namely Ventricular Tachycardia (VT) and Ventricular Fibrillation (VF) using feature selection algorithm and classification using SVM (Support Vector Machine). The author concludes that by combining ECG parameters with SVM has significantly improved the efficiency for detection of arrhythmia. Farid Melgani et al

978-1-7281-4192-3/19/$31.00 ©2019 IEEE
DOI 10.1109/IRCE.2019.00021

[8], the author has done an experimental study on automatic classification using SVM and further used PSO (Particle Swarm Optimization) algorithm for optimizing the performance of the SVM classifier.

Sudden cardiac death is caused by abnormal heart rhythm leading to cardiac arrhythmias. The arrhythmias can be classified into tachycardia and bradycardia. The control device used for Tachycardia (too fast heart rhythm) is defibrillator and for bradycardia (too small heart rhythm) is pacemaker. To develop pacemaker, modeling of CVD is essential. Various linear and non-linear models are developed for human heart by Fikret Yalcinkaya et al based on hydraulic, electrical and mechanical properties [9]. Martin J. conlon et al [10] developed Human Circulatory System (HCS) based on systematic HCS loop with physical parameters. A nonlinear feedback linearization technique is applied to force the output of the systems to generate artificial electrocardiogram signal using discrete data [11]. Different mathematical modeling approaches for human heart are reported in literatures [9-11]. Feedback linearization-based control scheme for human heart is proposed in [12]. Maximum likelihood based method for system identification of Hammerstein-Wiener (HW) model is illustrated in [13]. A new approach to estimate the response related to stimuli using Hammerstein-Wiener model is proposed in [14]. In [15], the importance of structure detection and a new method for structure detection of Nonlinear Auto Regressive Exogenous (NARX) model is addressed. A control-oriented modeling approach is proposed to depict nonlinear behaviors of cardiovascular response at both onset and offset of treadmill exercises by using support vector machine regression. Further, based on the established nonlinear time-variant model, a novel switching Model Predictive Control (MPC) algorithm has been proposed for the optimization of exercise efforts [16]. The main contribution of the work is to identify the suitable machine learning algorithm for accurate classification of cardiac arrhythmias and to develop suitable control scheme for controlling Bradycardia.

The rest of the paper is organized as follows: Section 2 presents description of dataset. Section 3 deals with classification of arrhythmias using Machine learning Algorithms. Non-linear modeling of the cardiovascular system is discussed in Section 4. Section 5 discusses the controller design of cardiovascular system using Hammerstein Wiener Model. Section 5 deals with the result and discussion. Section 6 concludes the paper.

II. Dataset Description

The section describes the data set used for classification and modeling of the cardiac arrythmias.

A. Cleveland Heart Disease Dataset

The Cleveland Dataset was obtained from UCI Machine Learning Repository. The dataset contains 14 attributes out of which 13 attributes were taken for input and the prediction parameter is taken as output. Therefore, the total of 258 values and their corresponding 13 input attributes were taken as training dataset. Further, the 45 values and their 13 attributes were taken for validation as shown in Table I.

TABLE I: Description of the Cleveland Dataset

Clinical Features	Description
Age	In Years
Chol (mg/dl)	Serum Cholesterol in mg/dl
Cp	Chest Pain type (typical- asymptomatic)
Exang	Exercise induced angina(1-yes;0-no)
Fbs	Fasting blood sugar(>120mg/dl)
Oldpeak	ST depression caused by exercise
Restecg	Resting ECG Results (Normal-Ventricular Abnormality)
Sex	Gender (1- male; 0 -female)
Slope	Upsloping - Down sloping
Thal	Normal – Reversible defect
Thalach	Maximum No.of Heartbeat
Trestbps (mmHg)	Resting BP(mm Hg)
Ca	Major Vessels (0-3)

B. MIT-BIH Arrhythmia Dataset

In this work, the MIT-Arrhythmia Database has been used. The dataset contains ECG recordings from 47 subjects from the Boston's Beth Israel Hospital. The ECG recording was a collection of both in-patient and outpatients. Further, the patients were 25 men aged between 32 to 89 years and 22 women aged between 23-89 years. The acquisition of ECG signal contains 445 two-channel Holter recorders and all the obtained signal were digitized. Each ECG signal from the Arrhythmia Database includes two important leads namely MLII and V1. The leads were used for detection of heartbeat using QRS complex and arrhythmia classification. Also, the normal sinus rhythm database was used for classification of normal and abnormal ECG signal. Both the physionet dataset contains different types of cardiac diseases namely bradycardia, normal sinus, atrial flutter and fibrillation, ventricular tachycardia and bigeminy.

C. Action potential Dataset

Action potential used modeling cardiac vascular system is generated by Hodgkin and Huxley model. The detail description of the model is given in [16]

III. Classification of Arrhythmias Using Machine Learning Algorithms

This section is comprised of the deep learning algorithms used for classification of Cardiac arrhythmias.

D. CNN (Convolutional Neural Network)

CNN is one of the forms of deep learning algorithm which contains many hidden layers and parameters. The CNN does not need any supervision since it can learn by itself and organize on its own. CNN can be applied to various applications namely automatic detection in medical filed, image classification and pattern recognition. CNN mainly performs the following operations namely convolution, detect non-linearity in the data, pooling and classification of the data. The main features of CNN were 1. It contains 3D namely width, height and depth of the neurons inside the layers. 2. It has local connectivity between the neurons and the neurons of the adjacent layers. 3. It also has shared weights to form the feature map.

E. SVM (Support Vector Machine)

SVM is the most popular classifier used in many classification methods. It is also a linear classifier which uses support vectors to find the best possible distance to the closest data set during training within the data points. Mostly kernel function like linear, radial basis function and quadratic are used. Further, the combination of CNN-SVM is attempted for improving the accuracy.

The nonlinear modeling of cardiac vascular system is described in the subsequent section.

IV. Nonlinear Modeling of Cardiac Vascular System

A. NARX Model

The NARX Model consists of regression block and estimate block. The current and past input values and past output data are to the regression block. Regression block output is applied to estimate block. Non-Linear function is estimated using the estimator with different activation functions like wavelet network, tree partition, sigmoid network, neural network and custom network. The NARX model equation is given in (1)

$$y(t) = L^T(u-r) + d + g(Q(u-r)) \quad (1)$$

where $y(t)$ is output, $r(t)$ is regressor $u(t)$ is input and $L^T(u-r)$ is output of Linear function block, L is a autoregressive with exogenous (ARX) linear function, d is a scalar offset, $g(Q(u-r))$ represent output of non-linear function and Q is projection matrix.

B. HW Model

The block diagram of HW is shown in Fig. 1

The Hammerstein model equation is given in (2)

$$u_1(t) = f(u(t)) \quad (2)$$

For Second block $y_1(t) = q(u_1(t))$ (3)

where $q = \frac{B(q)}{F(q)}$

Therefore, B and F are Similar to polynomials in the linear output Error model

The Weiner Model equation is given in (4)

$$y(t) = h(y_1(t)) \quad (4)$$

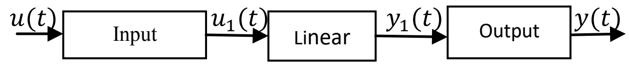

Figure 1. Block diagram of HW model

C. RNN Model

Recurrent Neural Network is used to represent time series data, in this network present value depends on its past values. One of the most important recurrent neural networks in deep learning algorithm is Long-Short Term Memory (LSTM) Network. The structure of LSTM is given in Fig.2.

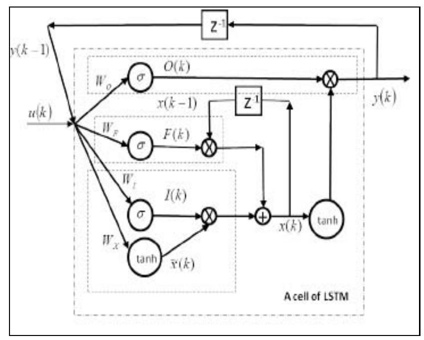

Figure 2. Structure of LSTM

The development of control scheme based on the HW model is discussed in the subsequent section.

V. Control of Cardiacvascular System

Once Sinoatrial (SA) node failure is diagnosed, pacemaker can be used to control cardiac vascular system. The design of pacemaker using HW model is discussed in this section. The Fig 3 shows the block diagram of cardiac vascular control system using Hammerstein Weiner model.

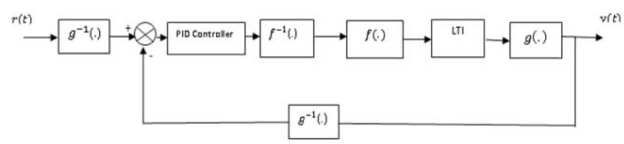

Figure 3. Control scheme for Hammerstein Wiener Model

VI. Results and Discussion

This section contains results of classifications of arrhythmias, modeling of cardiac vascular system, controller design and closed loop performance of control loop with cardiac vascular system.

A. Classification of Heart Disease using Clinical Data

CNN has three important parts

1. Data Pre-processing of the input ECG signal – To understand the different types of diseases

2. Convolution and Pooling: To extract maximum features for prediction of the disease

3. Fully connected Layer: Activating the SoftMax function for prediction of the disease.

1) Performance metrics

The dataset is evaluated using four different metrics namely accuracy, precision, recall and F1-measures. TP is denoted as True Positive, TN is denoted as True Negative, FP is denoted as False Positive, FN is denoted as False Negative.

Accuracy = TP+TN/TP=TN=FP=FN

Precision = TP/TP+FP

Recall = TP/TP+FN

F1-Measure = 2 x Precision x Recall/Precision +Recall

The performance metrics are performed to determine the success rate of prediction of the cardiovascular disease with respect to the clinical data as input. In order to improve the accuracy during prediction, four performance metrics were calculated.

2) Experimental results

Initially, CNN network is trained with the help of 2 convolutional layers and later the layers were increased up to 5 for the same number of feature map. There were 93 feature map and four fully connected layers and 1031 hidden layers. The fourth conventional layer has the best accuracy around 97% compared with all the other layers. Table 2 shows the comparison of different layers of various layers of CNN.

TABLE II: COMPARISON OF PERFORMANCE METRICS OF HEART DISEASE USING DIFFERENT LAYERS OF CNN

Number of CNN layers	Accuracy	Recall	Precision	F1-Measure
2	81	88.32	86.34	89.27
3	92	97.56	96.54	97.12
4	97	96.15	97.60	97.01
5	95	93.11	96.33	95.43

B. *Classifications of Arrhythmias using ECG signal*

The whole dataset taken for classification contains about 94,615 subject data. Some dataset which were not having enough information was removed. Finally, the total dataset taken for computation is 4500 subjects. The normal subjects were 2150 and the abnormal were around 2350. Further, the dataset has been divided for 60% for training the data and 40% for testing data.

1) Performance evaluation metrics:

The classification performance is evaluated using four different metrics namely accuracy, sensitivity, specificity and positive predictive value. The specificity refers to some part of the negative results from testing that are identified correctly. Sensitivity refers to the probability of correct classification of the arrhythmia as normal or abnormal (tachycardia or bradycardia). Positive predictive value is the result of the positive result from the test data of the dataset. Table 3 shows the performance metrics for classification of arrythmias.

TABLE III. COMPARISON OF PERFORMANCE METRICS OF THE PROPOSED MACHINE LEARNING ALGORITHM FOR ARRHYTHMIA CLASSIFICATION

	Accuracy (%)	Sensitivity (%)	Specificity (%)	PPV (%)
SVM	91.02	84.25	93.46	90.41
CNN	93.23	89.79	94.51	92.71
CNN-SVM	94.50	90.01	95.34	95.88

C. *Modeling of Cardiac Vascular System*

Action potential generated by the SA node and ECG wave are given as input to the machine learning algorithms and NARX, HW and RNN nonlinear models are developed. The developed models are compared qualitatively as well as quantitatively. The qualitative comparisons are shown in Fig.4.

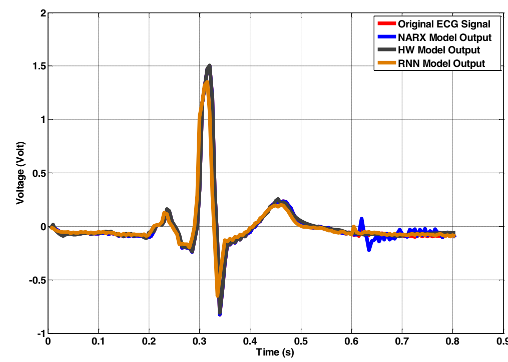

Figure 4. Qualitative comparisons of model responses

From the Fig. 4, it is found that NARX model output does not superimpose with original model after 0.6second. Similarly, QRS wave form of RNN model is not reached the peak value of original signal. Meanwhile HW outperforms other models. The quantitative comparisons are performed by computing Mean Square Error (MSE), Normalized Mean Square Error (NRSE), Root Mean Square Error (RMSE), Normalized Root Mean Square Error (NRMSE) and accuracy. The values are tabulated in Table 4.

TABLE IV. QUANTITATIVE COMPARISON OF MODEL PERFORMANCES

Goodness of Fitting	HW Model	NARX Model	RNN Model
MSE	2.6242e-04	4.8983e-04	0.0153
NMSE	0.0036	0.0067	0.2095
RMSE	0.0162	0.0221	0.1235
NRMSE	0.0600	0.0820	0.4578
PRD (accuracy)	94%	91.8%	54.22%

From the quantitative comparison, it is observed that HW model provides best performance than other models. Hence, HW model is selected for controller design.

D. Design of HW Controller for Cardiac Vascular System

The HW controller is designed using inverse HW model. The input nonlinearity and output nonlinearity of inverse model are shown in Fig. 5 and Fig.6.

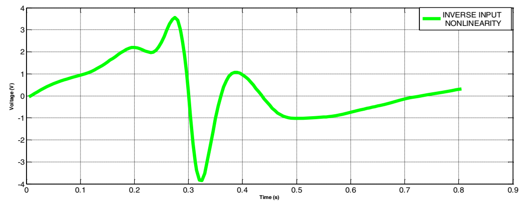

Figure 5. Input nonlinearity of inverse model

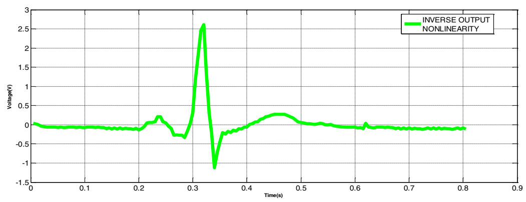

Figure 6. Output nonlinearity of inverse model

The transfer function of linear block is given in Equation (5)

$$\mathbf{a} = \frac{-0.9759\, z^{-1} + z^{-2}}{1 - 2.773 z^{-1} + 2.589 z^{-2} - 0.813 z^{-3}} \tag{5}$$

The PID controller parameters obtained using linear block is tabulated in Table V.

TABLE V. CONTROLLER PARAMETERS OF HW

Time(S)	k_P	k_I	k_D
1	0.000366	0.000366	0.0000914

The designed HW controller is used for closed loop analysis of cardiac vascular system in the next section.

E. Closed Loop Response of Cardiac Vascular System

The closed loop response and associated manipulated variable of control system designed using the HW model is shown in Fig.7 and Fig.8

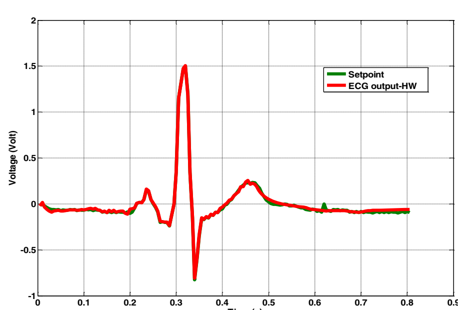

Figure 7. Closed ECG signal of HW model based cardiac vascular system

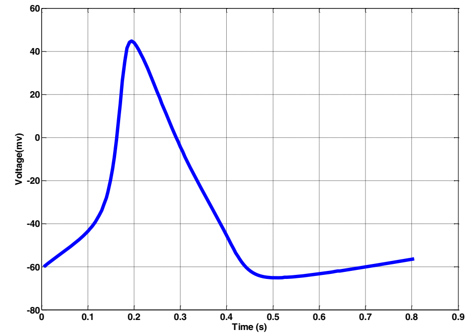

Figure 8. Manipulated variable associated with the control system

VII. CONCLUSION

This paper is aimed to classify the arrhythmias from the clinical data and ECG signals using CNN,SVM and CNN-SVM. From the performance comparisons, it is found that CNN-SVM outperforms than CNN and SVM. Then nonlinear models such as NARX,HW and RNN are developed for cardiacvascular system to design controller for Bradycardia and their performances are compared qualitatively as well as quantitatively. From the performance comparison it is found that HW model provides best performances than other models. Finally, the best fit HW model is used for pacemaker/controller design for bradycardia patients. The closed loop simulation result shows HW model based controller provides good performance.

REFERENCES

[1] Luz, Eduardo José da S., William Robson Schwartz, Guillermo Cámara-Chávez, and David Menotti. "ECG-based heartbeat classification for arrhythmia detection: A survey." *Computer methods and programs in biomedicine* 127 (2016): 144-164.

[2] Faust, Oliver, Yuki Hagiwara, Tan Jen Hong, Oh Shu Lih, and U. Rajendra Acharya. "Deep learning for healthcwere applications based on physiological signals: A review." *Computer methods and programs in biomedicine* 161 (2018): 1-13.

[3] Isin, Ali, and SelenOzdalili. "Cardiac arrhythmia detection using deep learning." *Procedia computer science* 120 (2017): 268-275.

[4] Sannino, Giovanna, and Giuseppe De Pietro. "A deep learning approach for ECG-based heartbeat classification for arrhythmia detection." *Future Generation Computer Systems*86 (2018): 446-455.

[5] Singh, Shraddha, Saroj Kumar Pandey, UrjaPawar, and Rekh Ram Janghel. "Classification of ECG arrhythmia using recurrent neural networks." *Procedia computer science* 132 (2018): 1290-1297.

[6] Yang, Weiyi, Yujuan Si, Di Wang, and Buhao Guo. "Automatic recognition of arrhythmia based on principal component analysis network and linear support vector machine." *Computers in biology and medicine* 101 (2018): 22-32.

[7] Acharya, U. Rajendra, Hamido Fujita, Shu Lih Oh, U. Raghavendra, Jen Hong Tan, Muhammad Adam, Arkadiusz Gertych, and Yuki Hagiwara. "Automated identification of shockable and non-shockable life-threatening ventricular arrhythmias using convolutional neural network." *Future Generation Computer Systems* 79 (2018): 952-959.

[8] Melgani, Farid, and YakoubBazi. "Classification of electrocardiogram signals with support vector machines and particle swarm optimization." *IEEE transactions on information technology in biomedicine* 12, no. 5 (2008): 667-677.

[9] Fikret Yalcinkaya, Ertem Kizilkaplan, and Ali Erbas, “Mathematical Modelling of Human Heart as a Hydroelectromechanical System” , *Proceedings of the International conference on Electrical and Electronics Engg*,Brusa, Turkey, (2013) :362-366

[10] Martin J. conlon, Donald L.Russell and Tofy Mussivand , “Development of a Mathematical Model of the Human Circulatory System”, *Annals of Biomedical Engineering,* 34(2006): 1400–1413.

[11] Witt Thanom and Robert,N. K. LOH “Nonlinear Control of Heartbeat Models” *Journal of Systemics cybernetics and Informatics*,9(2011):1400–1413.

[12] Adrian Wills, Thomas B. Schon, Lennart Ljung and Brett Ninness , “Identification of Hammerstein model” Automatica, 49(2013):70-81.

[13] Reza AbbasiAsl, Rahman Khorsandi, Bijan Vosooghi-Vahdat, “Hammerstein-Wiener Model: A New Approach to the Estimation of Formal Neural Information”, *Journal of Basics and Clinical Nero Science*, 3(2012):45-51.

[14] E. Radmaneshfar, M. Karrari , “A New Method for Structure Detection of Nonlinear ARX model: ANOVA_BSD” ,*Proceedings of the World Congress on Engineering*, , London, U.K , 2017:407-411.

[15] Teddy M.cheng Andrey, V Savkin senior, Branko G celler, Steven W.su, and Lu Wung, “Nonlinear Modelling and Control of Human Heart Rate Response During Exercise with various work load Intensities”, IEEE Transactions on Biomedical Engineering, 55(2008), 2499-2508.

[16] Jose A. Dfaz, M.Oscar Tequita and Fernando N aranio, “Neuronal Synchronization of Electrical Activity using the Hodgkin-Huxley Model and RCLSJ Circuit”, Journl of Ingenieria Y Ciencia,12(2016):93-106.