Remote Health Monitoring System

(Deep Neural Network Classifier)

A project submitted

in partial fulfillment of the requirements for the degree of

BEng (HONS) Electronic & Information Engineering

by

Liu Chenning (14111552D)

Supervised by

Dr. Lawrence Chi-Chung CHEUNG

Dr. Ka-Hong LOO (Assessor)

April 2018

The Hong Kong Polytechnic University
Department of Electronic and Information Engineering

EIE4430 / EIE4433 Honors Project

1. Student Name: __________Liu Chenning_____________ (Student No.:14111552D)

2. Programme Code: 42470

3. Project Title: Remote Health Monitoring System (Deep Neural Network Classifier)

4. Supervisor Name: Dr. Lawrence Chi-Chung Cheung

5. Project summary (State clearly the project objectives and results achieved by yourself.)

This project is to develop Deep Neural Network (DNN) classifiers to identify different types of heart arrhythmia through electrocardiogram (ECG) signals and enhance the outlook of the interface of the web server with the above classifiers. The DNN classifiers can identify heart arrhythmia with reasonable accuracy rates and the outlook of the interface of the web server has been enhanced.

DECLARATION OF ORIGINALITY

Except where reference is made in the text of this report, I declare that this report contains no material published elsewhere or extracted in whole or in part from any works or assignments presented by me or any other parties for another subject. In addition, it has not been submitted for the award of any other degree or diploma in any other tertiary institution.

No other person's work has been used without due acknowledgement in the main text of the Report.

I fully understand that any discrepancy from the above statements will constitute a case of plagiarism and be subject to the associated.

Liu Chenning

Signature

Abstract

This project aims to develop Deep Neural Network (DNN) classifiers to identify different types of heart arrhythmia from electrocardiogram (ECG) signals and enhance the outlook of the interface of the web server with the above classifiers. Following Association for the Advancement of Medical Instrumentation (AAMI) standard, a 5-class DNN classifier is built and its performance is optimized. Training and testing datasets were retrieved from MIT-BIH database. The feature extraction methods of the above datasets include wavelet transform detection and ECG morphology feature extraction. The DNN classifier achieves an overall of 92.30% with specificity of 99.995%. A remarkable improvement in the sensitivity of Supraventricular ectopic beat has reached 80% with specificity of 60%. The result shows that the accuracy of the classifier is high.

***Key words**: ECG, feature extraction, DNN implementation, DNN optimization*

Content

List of Tables

List of Figures

Acknowledgements

I would like to express my very gratitude to Dr. Lawrence Cheung, my FYP supervisor, for his patient guidance, enthusiastic encouragement and constructive suggestion in this project.

I would also like to offer my special thanks to Sean, for his advice and assistance in morphology data analysis. Advice given by Zlatan and Kiki has been a great support in this year.

My special thanks are extended to Fei, for her creative construction and devoted efforts in the hardware part.

Finally, I wish to thank my parents and friends, for their trust, encouragement and support throughout my study.

1. Introduction

1.1. Background

Arrhythmias are disturbances in normal cardiac electrical activity, and they usually occur in structurally normal hearts [1]. Among all kinds of cardiovascular disease, Cardiac Arrhythmia is the most common symptom that may lead to sudden death. The arrhythmia can be either mild or transient, and many cases of this disease show that patients usually share tiny but severe symptoms. Thus, it is very important for people to recognize various arrhythmias and get professional diagnosis as early as possible.

Currently, the detection of arrhythmia is mainly based on the electrocardiograph (ECG) classification. An ECG signal detector records the bioelectrical signals generated by the heart by placing electrode leads on the skin. In hospitals, the ECG morphology is recorded by a machine and doctors will read the ECG data manually to diagnose if the patient has Cardiac Arrhythmia. Usually, machines are expensive and the ECG signal requires beat-by-beat professional diagnosis. As a result, people with potential arrhythmia symptom usually are not willing to purchase one. A cheap, reliable and smart ECG signal classification system is important for them.

In this project, DNN classifiers have been developed to classify arrhythmia heartbeats. Wavelet transform is used to accurately locate the positions of a waveform within each heartbeat. This project, focusing on the development of DNN classifiers, cooperates with another final year project which is to build up a ECG signal detector with mobile app. Patients can wear a lightweight ECG detector and use the mobile app to check

ECG signals in real time. The web server can get ECG signals from the detector through the mobile app, classify the type of heart arrhythmia using DNN classifiers and return the result to the mobile app.

1.2. Objectives

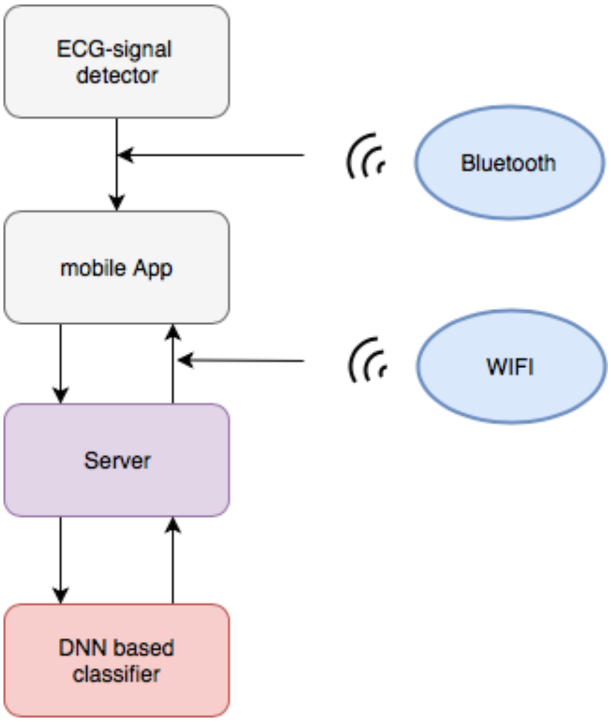


Figure 1 overview of this project

This project aims to develop DNN classifiers and enhance the web interface of the web server. It is divided into two parts:

- Build up five-class DNN classifiers (Class N, S, V, F and Q, especially in the sensitivity of Class S); and
- Improve the web interface of the web server.

2. Proposed Methodology

2.1. ECG data

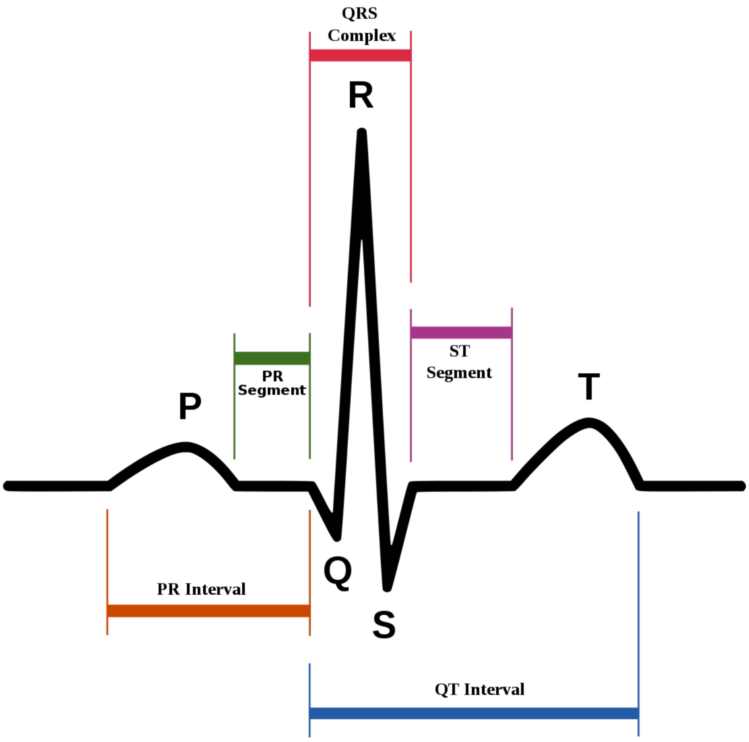


Figure 2 a whole complete morphology

The ECG signal for a normal heartbeat [3] is illustrated in the above figure. It includes:

- P wave: atria depolarization
- QRS complex: ventricular depolarization
- T wave: ventricular repolarization

Usually features extracted from this signal are:

- P peak: peak point of P wave
- PR interval: interval between P wave onset and R peak
- QRS duration: interval between Q wave onset and S wave offset
- Q peak: peak point of Q wave
- R peak: peak point of R wave
- S peak: peak point of S wave

- QT interval: interval between Q wave onset and T wave offset
- ST interval: interval between S wave onset and T wave offset
- RR interval: interval between successive heartbeat fiducial points

2.2. Data record selection

2.2.1. MIT-BIH database

Supported by Massachusetts Institute of Technology, MIT-BIH database [2] is a free and open-to-public database that provides 47 arrhythmia ECG recordings for analysis and testing in academia and industry. There are three standard ECG databases in the world, AHA [4], ST-T [5] and MIT-BIH. Among them, MIT-BIH database is the most widely used database in these years because it is free of charge.

The ECG recordings were detected from 47 subjects studied by BIH Arrhythmia Laboratory, including 25 men (32 to 89 years old) and 22 women (23 to 89 years old). 25 recordings (patient number 200 to 234) are selected to cover different kinds of rare but clinically important types of ECG signals. To reduce the size of the database file, MIT-BIH adopted a customized format for its 48 recordings. In each record, the ECG signal is digitized at 360 Hz sampling rate and is annotated by cardiologist for reference. Two-channel ambulatory signals were recorded, and in this project, modified limb lead II (MLII) is used for analysis.

2.2.2. Heartbeat type and labels

The MIT-BIH database uses 15 different types of heartbeat to label each ventricular beat (Table 1). According to the Association for the Advancement of Medical Instrument (AAMI) [6], the 15 types of heartbeats are mapped into five types of heartbeat.

AAMI heartbeat description	**N** Any heartbeat not in the S, V, F or Q	**S** Supraventricular ectopic beat	**V** Ventricular ectopic beat	**F** Fusion beat	**Q** Unknown beat
MIT-BIH heartbeat types	Normal beat (NOR)	Atrial premature beat (AP)	Premature ventricular contraction (PVC)	Fusion of ventricular and normal beat (fVN)	Paced beat (P)
	Left bundle branch block beat (LBBB)	Aberrated atrial premature beat (aAP)	Ventricular escape beat (VE)		Fusion of paced and normal beat (fPN)
	Right bundle branch block beat (RBBBB)	Nodal (junctional) premature beat (NP)			Unclassified beat (U)
	Atrial escape beats (AE)	Supraventricular premature beat (SP)			
	Nodal (junctional) escape beat (NE)				

Table 1 heartbeat types in AAMI standard

In this project, five categories are adopted in the classification process, which is the same as most of the published papers. This is important since it is easy to compare the optimized performance with other existing algorithms. In accordance with recommendations from AAMI, paced beats (P) are removed from analysis. Thus, the rest 14 types are used in this project, and they are mapped into 5 labels:

Heartbeat label	Description	Mapped label
N	Normal beat (NOR)	
L	Left bundle branch block beat (LBBB)	
R	Right bundle branch block beat (RBBB)	N
e	Atrial escape beat (AE)	
j	Nodal (junctional) escape beat (NE)	
A	Atrial premature beat (AP_	
a	Aberrated atrial premature beat (aAP)	S
J	Nodal (junctional) premature beat (NP)	

S	Supraventricular premature or ectopic beat (atrial or nodal) (SP)	
V	Premature ventricular contraction (PVC)	V
E	Ventricular escape beat (VE)	
F	Fusion of ventricular and normal beat (fVN)	F
f	Fusion of paced and normal beat (fPN)	Q
Q	Unclassifiable beat (U)	

Table 2 heartbeat types used in this project

2.2.3. Selection of training & testing recordings

To align with AAMI recommendation, four recordings (102, 104, 107 and 217) that contain paced beats are removed from analysis. Same as other research works, the rest 44 recordings are divided into 2 datasets (Table 3) [7]. Dataset 1 (DS1) is used for training and Dataset 2 (DS2) is used for testing. Arrhythmia recordings are evenly divided, and each dataset contains roughly 50,000 heartbeats. Table 4 shows the breakdown of each recording in five types. Table 5 shows the breakdown of each dataset in five types.

	Recordings	**Total**
Full database	100, 101,102, 103, 104, 105, 106, 107, 108, 109, 111, 112, 113, 114, 115, 116, 117, 118, 119, 121, 122, 123, 124, 200, 201, 202, 203, 205, 207, 208, 209, 210, 212, 213, 214, 215, 217, 219, 220, 221, 222, 223, 228, 230, 231, 232, 233, 234	48
DS1	101, 106, 108, 109, 112, 114, 115, 116, 118, 119, 122, 124, 201, 203, 205, 207, 208, 209, 215, 220, 223, 230	22
DS2	100, 103, 105, 111, 113, 117, 121, 123, 200, 202, 210, 212, 213, 214, 219, 221, 222, 228, 231, 232, 233, 234	22

Table 3 Dataset 1 & Dataset 2

Record No.	N	S	V	F	Q	Total
100	2239	33	1	0	0	2273

101	1860	3	0	0	2	1863
103	2082	2	0	0	0	2084
105	2526	0	41	0	5	2567
106	1507	0	520	0	0	2027
108	1740	4	17	2	0	1763
109	2492	0	38	2	0	2532
111	2123	0	1	0	0	2324
112	2537	2	0	0	0	2539
113	1789	6	0	0	0	1795
114	1820	12	43	4	0	1899
115	1953	0	0	0	0	1953
116	2302	1	109	0	0	2412
117	1534	1	0	0	0	1535
118	2166	96	16	0	0	2278
119	1543	0	444	0	0	1987
121	1861	1	1	0	0	1862
122	2476	0	0	0	0	2476
123	1515	0	3	0	0	1518
124	1536	31	47	5	0	1620
200	1743	30	826	2	0	2601
201	1635	128	198	2	0	1963
202	2061	55	19	1	0	2136
203	2529	2	444	1	4	2980
205	2571	3	71	11	0	2656
207	1543	107	210	0	0	1860
208	1586	2	992	373	2	2955
209	2621	383	1	0	0	3005
210	2423	22	195	10	0	2850
212	2748	0	0	0	0	2748
213	2641	28	220	362	0	3251
214	2003	0	256	1	2	2262
215	3195	3	164	1	0	3363
219	2082	7	64	1	0	2154
220	1954	94	0	0	0	2048

221	2031	0	396	0	0	2427
222	2274	209	0	0	0	2483
223	2045	73	473	14	0	2605
228	1688	3	362	0	0	2053
230	2255	0	1	0	0	2256
231	1568	1	2	0	0	1571
232	398	1382	0	0	0	1783
233	2239	7	831	11	0	3087
234	2700	50	3	0	0	2753
Total	**90125**	**2781**	**7009**	**803**	**15**	**100733**

Table 4 heartbeat type distribution in each recording

Recordings	N	S	V	F	Q	Total
DS1	45866	944	3788	415	8	51021
DS2	44259	1837	3221	388	7	49705
44 recordings	90125	2781	7009	803	15	100733

Table 5 heartbeat type distribution in 2 datasets

2.3. Toolbox

2.3.1. LIBSVM

LIBSVM [8], developed by Lin Chi-Jen from Taiwan University, is a fast, efficient and easy-to-use Support Vector Machine (SVM) software toolbox. It is widely used to solve classification problems, including regression and one-class-SVM program. Four modes for training supported by this toolbox are: linear, polynomial, radial basis function-RBF and sigmoid. These four modes are used to effectively solve multi-class problems and probability estimation in multi-class with unbalanced data. Manual adjustments of parameters in training a SVM model are greatly reduced. The toolbox provides default parameters and make it convenient to use.

When deploying a SVM to perform classification and regression, there is no uniform format for selection of methods, parameters and models. Thus, current optimized SVM algorithms are still trained based on past experience, experimental comparisons or other resources.

2.3.2. ecgpuwave

The *ecgpuwave* [9] toolbox is a lightweight Matlab toolbox supported by PhysioNet [10], a research resource for complex physiologic signals. This toolbox is used to detect QRS complex locations and the peak locations. It has been successfully tested on several ECG databases: MIT-BIH, CSE [11] and ST-T [5].

The QRS detector is based on the Pan and Tompkins algorithm (PT algorithm) [12], and it uses slope information [13] to improve the performance. This toolbox is often used to find location information of QRS complex.

2.3.3. ECG-kit

ECG-kit [14], developed by Mariano Llamedo Soria, is a lightweight development open-source toolbox with a common API that can be implemented in Matlab. This toolbox provides several popular algorithms in ECG signal analysis: QRS detection, wavelet-based ECG delineator [15] and so on. It also allows inputs from several ECG databases, including MIT-BIH.

The wavelet-based ECG delineator algorithm is used to detect the location of P wave, QRS complex and T wave. Over 20 features can be derived from the combination of location information.

2.3.4. Wavelet transform toolbox

Wavelet Transform [16] is a useful time-frequency domain analysis method after Fourier Transform. Before the appearance of Wavelet Transform, Fourier Transform is the most widely used and most effective analysis in the field of signal processing. However, compared with Wavelet Transform, Fourier Transform still has limitations in time-frequency localizing and unstable signal analysis.

Wavelet Transform is a local transition between time domain and frequency domain. Thus, it can effectively extract information from signal using multiscale analysis. Moreover, Wavelet Transform can effectively filter out local noise, and thus perform well in feature extraction [17]. Thus, the wavelet transform toolbox [18] is good for analysis on ECG data, an unstable signal that requires localized analysis.

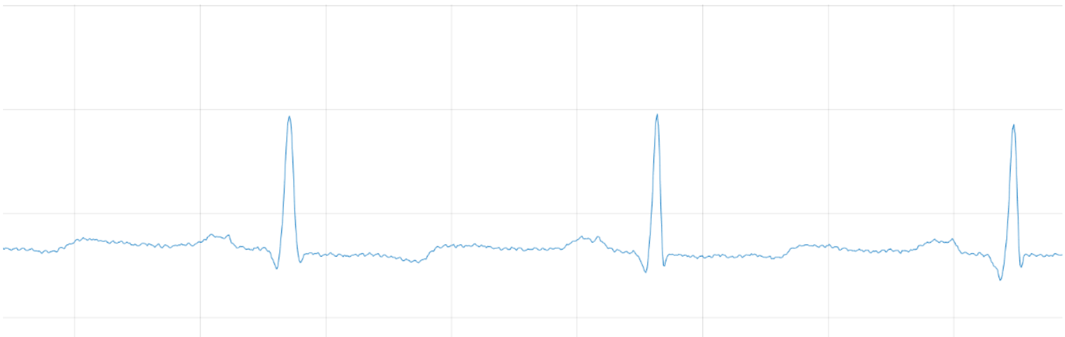


Figure 3 original ECG signal

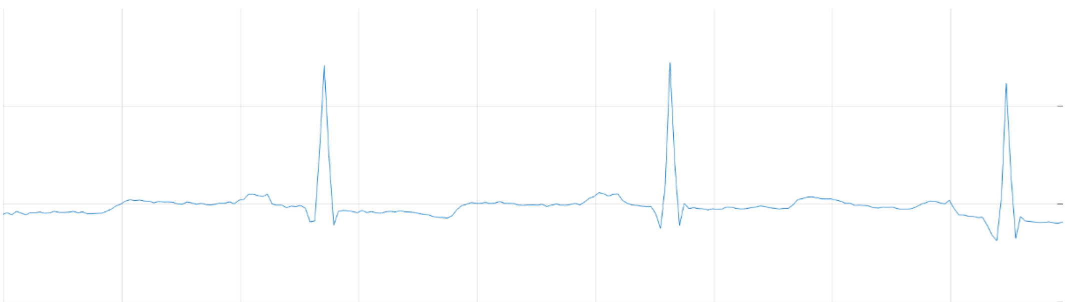


Figure 4 signal with noise filtered

This toolbox is used to locate the begin, peak and end location of P wave, QRS complex and T wave. To prove that this toolbox can correctly locate the position of each wave, the accuracy of detected beats is calculated. The criteria for the comparison is 5 sample points (0.014s). If the distance between annotated R peak and detected R peak is less than 5 sample points, this heartbeat is marked to be correctly detected. Since the beginning and ending of a record may contain incomplete heartbeats, they are not considered as detected beats.

2.3.5. Feature extraction

This project contains three parts in automated heartbeat classification: data preprocessing, feature extraction and DNN classification. In the first stage, high-frequency noise is removed by Wavelet Transform. The ECG detector then detects if there are P wave, QRS complex and T wave in each heartbeat, and gives their corresponding positions. Note that raw signals are not directly fed into DNN classifier. In practice, processing the raw data in a classifier is technically possible and can achieve a satisfied result. However, researchers found that discriminative features could often help to achieve a higher level of performance [19]. Results in published papers have shown a very promising application of using features as the input to construct a two-class classifier. Therefore, feature extraction is performed before classification.

The annotation file in MIT-BIH only contains R peak locations. Thus, it is impossible to directly compare the precision of detected P and T wave. Although the QRS complex can now be accurately detected (usually with 99+% accuracy using PT algorithm), currently there is no superior algorithms to detect the exact position of P and T waves.

Clinically, cardiologists manually find the exact positions and ratios of each part, and then give corresponding diagnosis. Thus, with help from cardiologists, the detected location was evaluated and proved to be accurate.

2.3.6. Synchronization of QRS location

In the first phase of feature extraction, QRS complex is found. The peak positions of Q, R and S peak are detected.

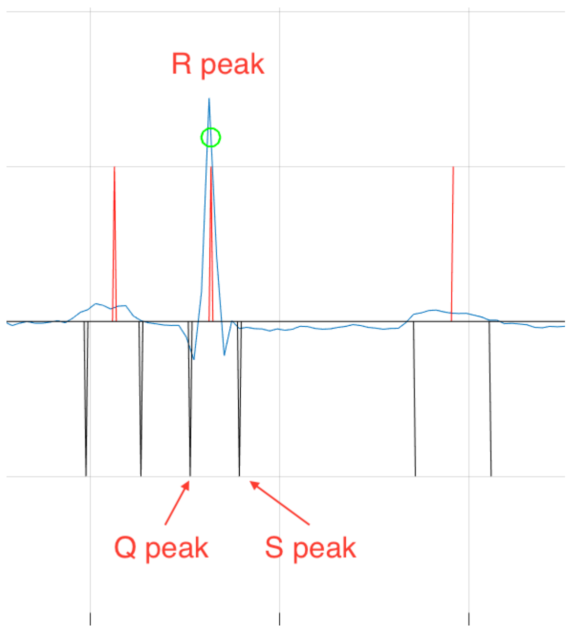


Figure 5 detection of peak, wave onset and offset

Although positions of three peaks can be found properly, synchronization is still necessary before features are extracted. Below is the logic of synchronization of peak locations in each heartbeat:

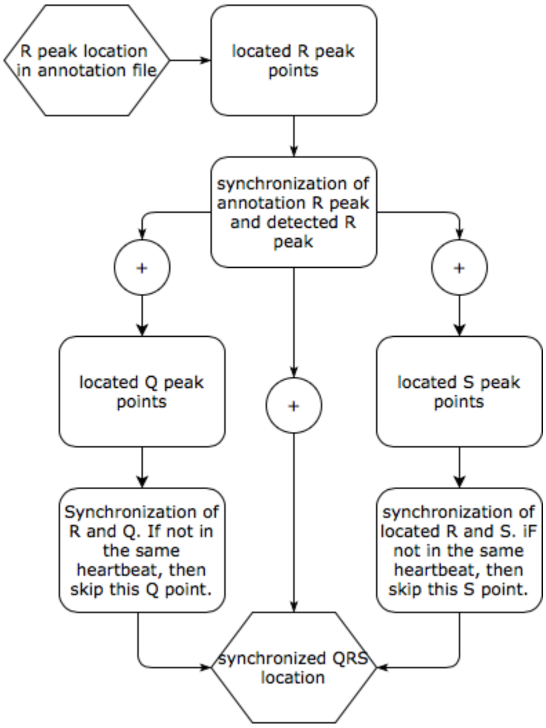


Figure 6 flow chart of feature extraction

Five cases may exist for synchronization of R peak location in annotation file and R peak points located by the toolbox. The green color represents R peak location in annotation file and the orange color represents R peak points located by the toolbox.

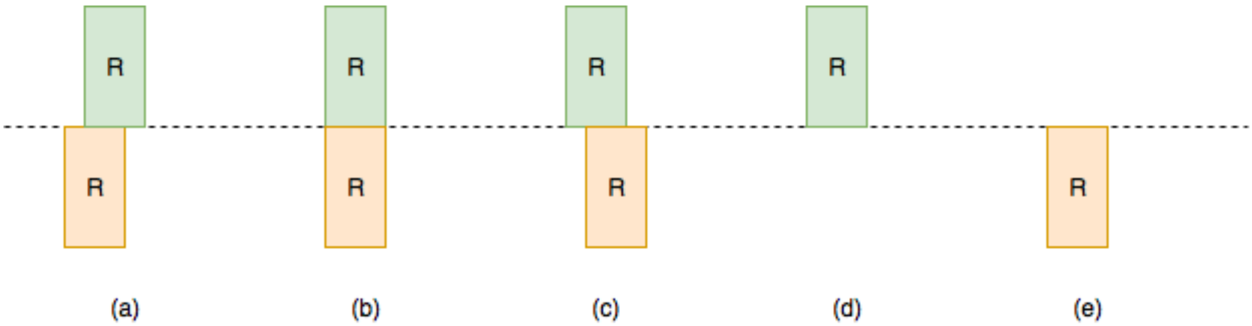


Figure 7 5 situations of synchronization

In the first three cases, if the distance of the green R and the orange R is less than five sample points, this located R peak will be marked as correctly detected and combined with annotated R peak for next stage. If annotated R exists but located R does not (case d), the located R will be marked with zero value in that heartbeat. The process is same for case e. The synchronization of Q peak and S peak is the same, the green block then changes to located R position, and the orange block changes to located Q and S position respectively.

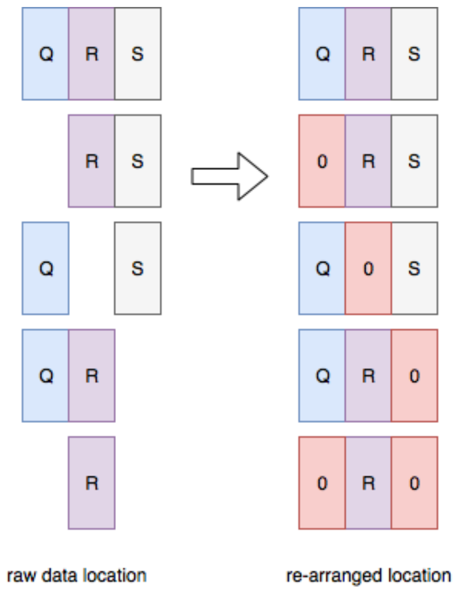


Figure 8 located positions & mapped QRS location

The detected QRS location is stored in the format of the right side in Figure 8. There could be five cases in mapping locations: all QRS exist, Q missing, R missing, S missing and QS missing. For convenience, any missing part of Q, R or S positions will be given a default value 0 (red block in re-arrange format).

2.3.7. Feature extraction I

The QRS complex usually includes important information. Features extracted from the QRS complex can be used to isolate arrhythmia beats. From observations, abnormal heartbeats usually have missing Q and S peaks. The use of QRS complex features and the DNN classifier were adopted to perform normal and PVC (premature ventricular contraction) classification [20]. However, it has not been used in classifying five types of heartbeats before. Thus, in the first phase of feature extraction, five features are collected: Q peak, R peak, S peak, QRS duration and RR interval.

Having QRS location detected, it is easy to extract these five features. As mentioned in the previous part, some of the heartbeats may have missing Q, R, or S peak locations, and a default value of 0 is assigned to those points. Table 6 shows the default value assignment in the feature set when some points are missing:

Description	Default value
Q location missing	Q peak = 0, QRS duration = 0
R location missing	R peak = 0, 2 consecutive RR intervals = 0
S location missing	S peak = 0, QRS duration = 0

Table 6 default value setting in feature set

The feature set is arranged in the sequence: Q peak, R peak, S peak, QRS duration and RR interval. The feature set is then combined with QRS location to label with the heartbeat type. It will be normalized using two methods in the next stage.

2.3.8. Synchronization of P wave and T wave

More features are extracted in the second phase of synchronization. P wave onset, P wave offset, P peak and T peak are further combined with original location information.

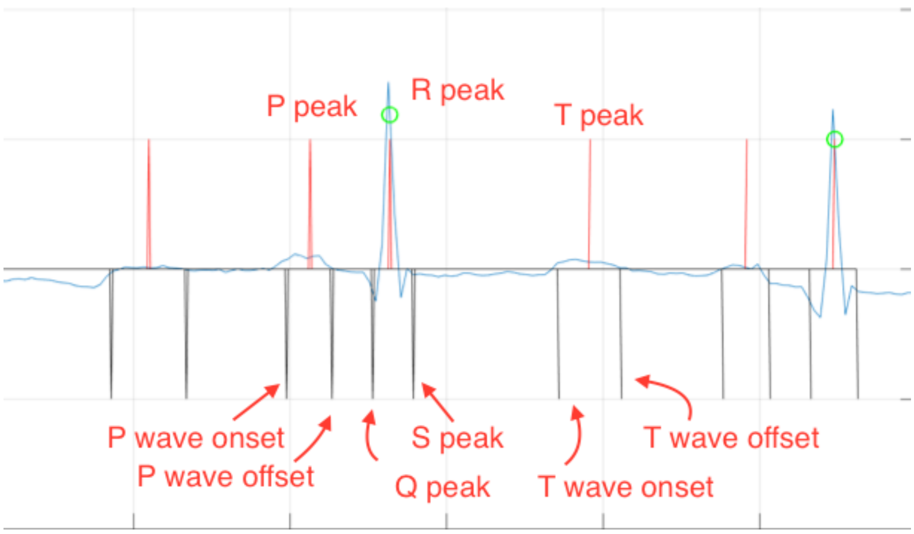


Figure 9 extracted wave locations in a complete morphology

The location of detected R peak position is served as a reference point within each heartbeat. Positions of P and T waves are compared with detected R peak. If the distance between current P wave and R peak is less than 170 sample points (derived from observations), current P wave will be marked in same heartbeat. Threshold value of 225 is used for T wave. Locations are saved in sequence: P wave onset, P peak, P wave offset, Q peak, R peak, S peak and T peak. If any location is missing, default value 0 will be assigned to it.

2.3.9. Feature extraction II

In the second phase of the feature extraction, seven more features are found: P peak, T peak, PR interval, P wave duration, ST interval and PP interval. In this stage, information outside the QRS complex are included. In total, twelve features served as the input of the DNN classifier. Table 7 shows the default value assignment in the feature set when points are missing:

Description	Pre-assigned value
P peak location missing	P peak = 0, PR interval = 0
T location missing	T peak = 0, ST segment = 0
P wave onset location missing	P wave duration = 0
P wave offset location missing	P wave duration = 0

Table 7 default value setting

From observations, in most of recordings, the initial five heartbeats cannot locate exact positions of P and T wave. Those beats are abandoned to avoid using ambiguous data in the DNN classifier. Also, if the input record is too long, the precision of detected P and T wave drops significantly. Therefore, for a better performance, each recording is chopped down into 4 parts for detection and combined together for normalization in the next stage. Usually in a record, roughly 15 heartbeats are abandoned, taking up 0.6% of total beats. The elimination of those beats will not greatly influence the overall result in classification.

2.3.10. Normalization

Before the feature sets are fed into the classifier, z-score normalization (standardization) in the feature set is performed. Using normalization, a more stable convergence of weight and biases can be achieved. Relationships that are more robust can be reached. The z-score normalization standardizes features column by column, to make features center at zero with a standard deviation of one. For each feature j in dataset 1, the i-th value is normalized according to the following formula:

$$z_{ij} = \frac{x_{ij} - \mu_{xj}}{\sigma_{xj}}$$

where x is the value in dataset 1, μ_{xj} is the mean value of that feature j in dataset 1, and σ_{xj} is the standard deviation in dataset 1. Values of feature j in dataset 2 is normalized in two ways:

$$z_{ij} = \frac{y_{ij} - \mu_{xj}}{\sigma_{xj}} \qquad \text{(a)}$$

$$z_{ij} = \frac{y_{ij} - \mu_{yj}}{\sigma_{yj}} \qquad (b)$$

The first method is to use the mean and standard deviation of dataset 1 to normalize dataset 2. This assumes that the distribution of dataset 2 is the same as dataset 1. The second method is to normalize dataset 2 using the same way as dataset 1. In this situation, dataset 2 may have a different distribution. For the 5-feature and 12-feature datasets, two normalization methods are compared in DNN classification.

2.4. Classification

Support Vector Machine (SVM) and Deep Neural Network (DNN) are deployed as the classifier in this project.

2.4.1. DNN classifier

The Deep Neural Network (DNN) is designed on Artificial Neural Network (ANN), usually with multiple hidden layers. In this project, a supervised model with a backpropagation algorithm is used. Currently, DNN classifiers have been applied to many applications, such as image and speech recognitions [18].

For each neuron, the structure is shown as follows:

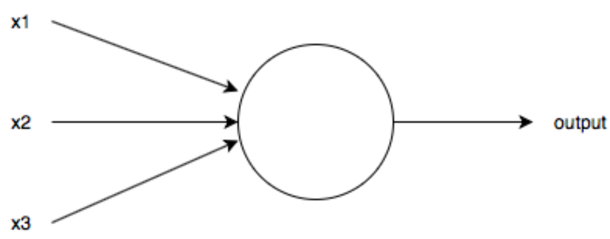


Figure 10 structure of a single neuron

$$z = \sum_{i=1}^{m} w_i x_i + b$$

where z is the output value, w_i is the weight of input x_i and b is the bias value. Initial weights are generated randomly for training. Rectified Linear Unit (ReLU) serves as the activation function. Maximum-likelihood is used for estimating the class of output. The Softmax function is used here:

$$\mathrm{p_j} = \frac{\exp\,(x_j)}{\sum_k \exp\,(x_k)}$$

where p_j is the probability of an object that belongs to class j, x_j and x_k are the inputs of neuron i and k. Stochastic gradient descent is then used to reduce error (backpropagation):

$$\Delta w_{ij}(t+1) = \Delta w_{ij}(t) + \eta \frac{\partial C}{\partial w_{ij}}$$

where w_{ij} is the weight between neuron i and j, η is the learning rate and C is the cross entropy. The calculation of C is shownbelow:

$$C = -\sum_j d_j \log p_j$$

A sample structure is shown as follows:

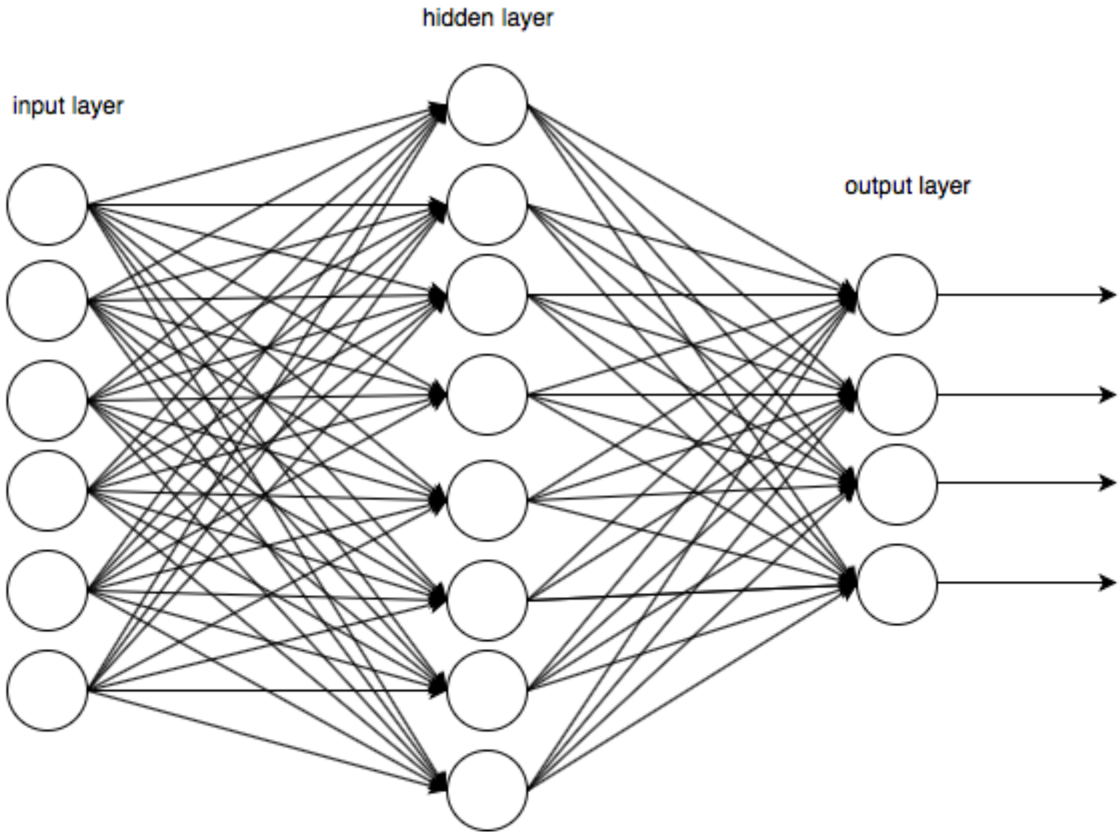


Figure 11 sample structure of a DNN classifier

2.4.2. Designed structure

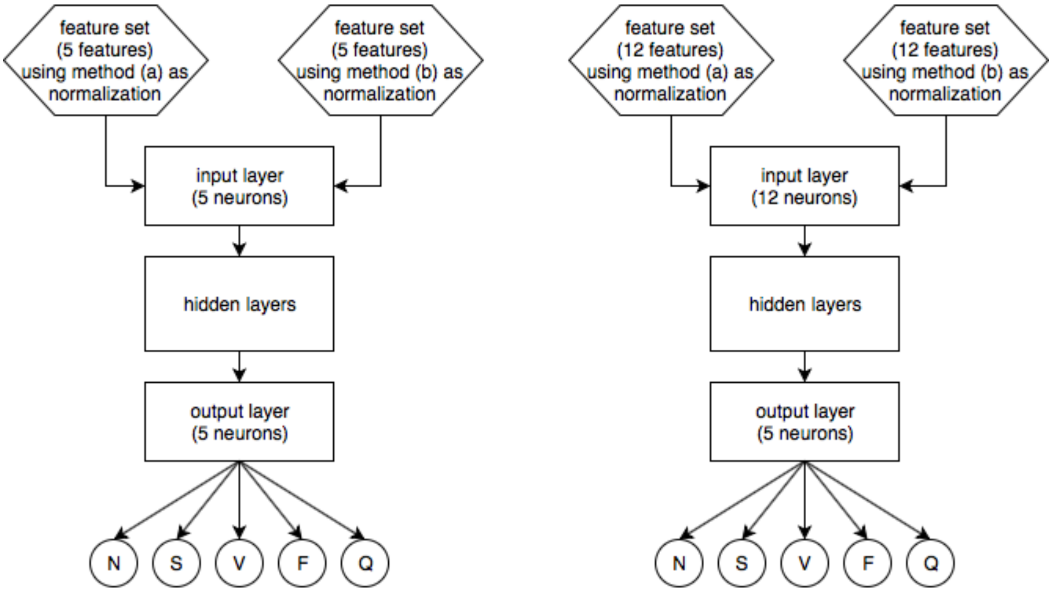


Figure 12 designed structures of DNN classifiers

Two DNN structures are designed for 5-feature dataset and 12-feature dataset respectively. The main difference is the number of neurons in the input layer. These two internal structures, including number of hidden layers and hidden nodes within each layer, are different and DNN optimization will be performed. In total, four input feature sets are used for building four DNN classifiers. In the next stage, four DNN classifiers are optimized for a final output.

2.5. DNN optimization

DNN optimization mainly focuses on the internal structure of the DNN classifier. The number of hidden layers and hidden nodes within each layer are investigated, and their performance is compared.

From the whole dataset, the quantity of N type heartbeats (90125) is far larger than S (2781) and V (7009) types. Moreover, F (803) and Q (15) type only take up less than 1% of the data. The unbalanced data usually results in a low sensitivity in a class with very few training and testing samples. The number of normal heartbeats is far larger than arrhythmia heartbeats. Therefore, it is easy to achieve a very high overall accuracy, because the classifier can classify all normal heartbeats properly. However, at the same time, the sensitivity of arrhythmia beats could be very low. However, using overall accuracy cannot give reliable comparisons between classifiers with different structures.

The previous project used PT algorithm to identify each heartbeat, and raw data were directly fed into a DNN model for classification. It has achieved a relatively high accuracy in normal type (N, sensitivity = 98.70%). However, the sensitivity and accuracy in S type (sensitivity = 5.66%) and F type (sensitivity = 5.67%) are not good.

As illustrated by AAMI, most of patients with dangerous arrhythmia usually have S or V type heartbeats. Thus, the sensitivity and specificity of class S and V are more important than class F and Q in the unbalanced data. To fine-tune the model, the sensitivity of class S and V is the primary concern.

It is generally accepted that when the number of hidden layers is large, the testing error is usually small and the accuracy is high. However, the complexity of the neural network is high and the training time is long. Additionally, the number of hidden nodes can also greatly affect the performance of the neural network. Too few hidden nodes can directly result in a poor performance. Meanwhile, although too many nodes may reduce system error, the training time may be long. It is also possible that in the training process, testing error gets trapped into a local minimum instead of the global minimum.

Some formulas are provided for guiding the design of a neural network:

$$n_h = \sqrt{n_i + n_o} + \alpha$$

$$n_h = \log_2 n_i + \alpha$$

$$n_h = \sqrt{n_i \times n_o} + \alpha$$

where n_i, n_h, n_o are the number of neurons in the input layer, the hidden layer and the output layer, and α is usually between 1 and 10. The above formulas indicate that the optimal points could be between 5 and 20.

The Receiver Operating Characteristic (ROC) curve is usually used to visualize the performance of classifier because a suitable decision threshold can then be selected based on it [21]. However, comparing two classifiers from a 2 dimensions would be difficult while a single number is a more favorable choice [22]. Therefore, for imbalanced datasets, a remedy could be Area Under the ROC Curve (AUC) [23]. Thus, in this project, ROC curve and AUC are combined to find the optimal classifier.

3. Evaluation

3.1. Evaluation of wavelet transform toolbox

The feasibility of this toolbox is tested. The result is compared through their annotation files. The MIT-BIH database provides reference points of R peaks. Beats are marked as correctly detected if the detected location is within five samples (0.0014 seconds) of the annotated location. Thus, the accuracy is calculated as follow:

$$\text{Accuracy} = \frac{correct\ detected\ beats + 2}{total\ beats} \times 100\%$$

3.2. Evaluation of DNN classifier

The performance of DNN classifiers is evaluated by accuracy, sensitivity, specificity and AUC. ROC curves of class S and V are used to evaluate each DNN structure. The structure with the highest AUC value of class S and V is selected as the optimized structure. Each of the four datasets has an optimized DNN structure, which is evaluated using AUC as primary selection criteria. For the optimized structure, confusion matrix is formed as below: correctly classified normal and arrhythmia heartbeats (true positive, TP), wrongly classified normal heartbeats (false positive, FP) and wrongly classified arrhythmia heartbeats (false negative, FN).

	N	S	V	F	Q
N	TP	FP	FP	FP	FP
S	FN	TP	FN	FN	FN
V	FN	FN	TP	FN	FN
F	FN	FN	FN	TP	FN
Q	FN	FN	FN	FN	TP

Overall accuracy is calculated as follow:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \times 100\%$$

Specificity

$$\text{Specificity}_{\text{normal}} = \frac{TP_{normal}}{total\ number\ of\ detected\ normal\ beats} \times 100\%$$

$$\text{Specificity}_{\text{abnormal}} = \frac{TP_{abnormal}}{total\ number\ of\ detected\ abnormal\ beats} \times 100\%$$

Sensitivity

$$\text{Sensitivity}_{\text{normal}} = \frac{TP_{normal}}{total\ number\ of\ normal\ beats} \times 100\%$$

$$\text{Sensitivity}_{\text{abnormal}} = \frac{TP_{normal}}{total\ number\ of\ abnormal\ beats} \times 100\%$$

The optimization result is evaluated using the ROC curve and the graph is swapped in horizontal direction for a better visual comparison. It is in a 2-D form using $(1 - \text{specificity})$ as x axis and sensitivity as y axis. Area under the ROC curve (AUC) is calculated:

$$\text{AUC} = \text{area}(\text{ROC curve}), \qquad \text{x} \in [0,1], \text{y} \in [0,1]$$

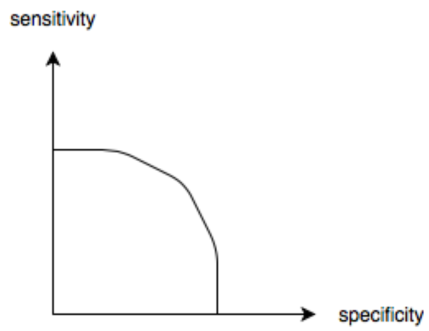
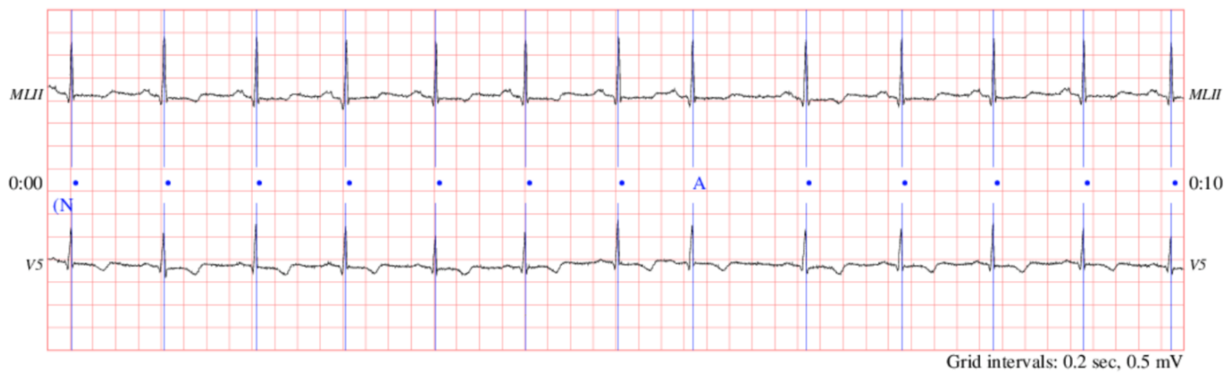


Figure 13 ROC curve

The final optimal ROC curve should be protruding to the right upper side as much as possible. The threshold (operating point) is chosen from the curve, a point where the area under that point is largest.

4. Project implementation

4.1. Evaluation of wavelet transform toolbox



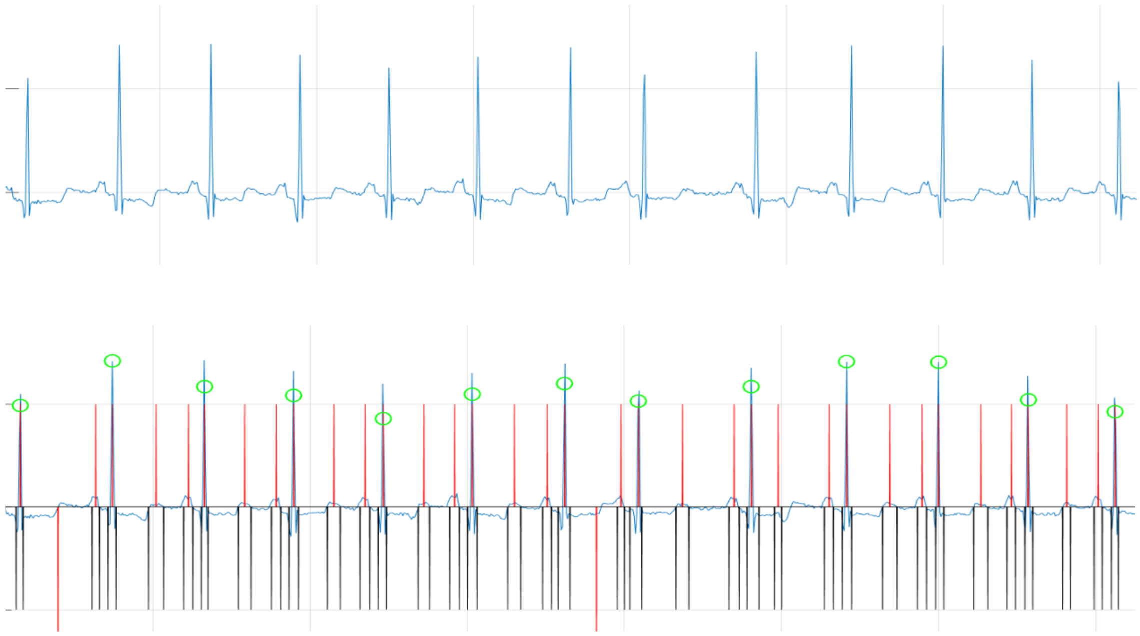


Figure 14 comparison between MIT-BIH, ECG signal with noise filtered and extracted location

The accuracy of the wavelet transform toolbox is evaluated. Above are three comparison graphs: data from MIT-BIH records, data with noise filtered, and detected positions of each part.

Record No.	Total beats	Correct detected beats	Accuracy
100	2273	2270	99.96%
101	1865	1861	99.89%
103	2084	2080	99.90%
105	2572	2538	98.76%
106	2027	1978	97.68%
108	1763	1722	97.79%
109	2532	2527	99.88%
111	2124	2117	99.76%
112	2539	2530	99.72%
113	1795	1790	99.83%
114	1879	1472	78.45%
115	1953	1945	99.69%
116	2412	2381	98.80%
117	1535	1528	99.67%

118	2278	2264	99.47%
119	1987	1927	97.08%
121	1863	1852	99.52%
122	2476	2469	99.80%
123	1518	1510	99.60%
124	1619	1601	99.01%
200	2601	2212	85.12%
201	1963	1910	97.40%
202	2163	2122	98.20%
203	3000	2128	71.00%
205	2656	2627	98.98%
207	1860	1674	90.11%
208	2955	2866	97.06%
209	3005	2960	98.57%
210	2650	2578	97.36%
212	2748	2733	99.53%
213	3251	3206	98.61%
214	2262	2235	98.89%
215	3363	3321	98.81%
219	2154	2129	98.93%
220	2048	2025	98.97%
221	2427	2395	98.76%
222	2483	2353	94.84%
223	2605	2556	98.20%
228	2053	1769	86.26%
230	2256	2067	91.71%
231	1571	1562	99.55%
232	1780	1759	98.93%
233	3079	3006	97.69%
234	2753	2740	99.60%
Total	**100733**	**97295**	**96.67%**

Table 8 tested accuracy of wavelet transform toolbox

Record 114 and 203 are two special records with large amount of ambiguous R peaks. From the graphs, those beats are usually V type heartbeats

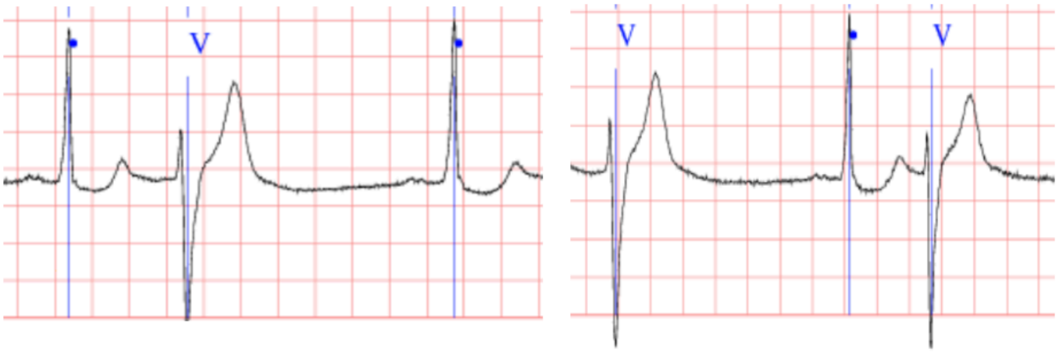


Figure 15 fragments of Record 114

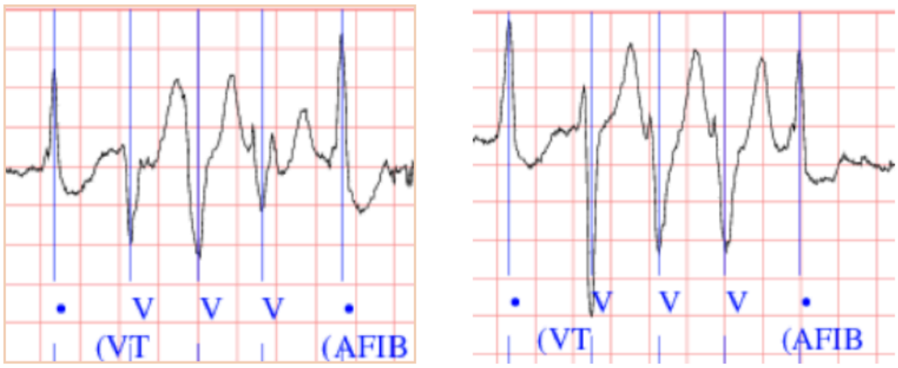


Figure 16 fragments of Record 207

The overall accuracy of wavelet transform toolbox is 96.67%, and this indicates that the detection of heartbeat is accurate. Thus, this toolbox is used for feature extraction in the next stage.

4.2. Evaluation on DNN classifier optimization

The evaluation of optimal DNN classifiers is divided into four parts: 5-feature dataset with two normalization methods, 12-feature dataset with two normalization methods. In each part, DNN internal structure is carefully designed and tested. For each internal structure, the epoch with the best performance (least testing error) is recorded. Then, the trained DNN model with the optimal epoch is used to find the AUC of class S and V.

For each part, the internal structure with the best AUC is selected as the optimal structure for the corresponding dataset.

4.2.1. 5-feature dataset with normalization method (a)

Internal structure	Epoch	AUC
[10 10 10]	5	S=0.2686, V=0.4399
[20 20 20]	5	S=0.2537, V=0.4652
[30 30 30]	5	S=0.2780, V=0.4661
[40 40 40]	5	S=0.2865, V=0.5020
[50 50 50]	5	S=0.2852, V=0.4699
[30 30 30 30]	8	S=0.3840, V=0.5369
[35 35 35 35]	8	S=0.3700, V=0.5472
[40 40 40 40]	**7**	**S=0.5728, V=0.6596**
[45 45 45 45]	7	S=0.2951, V=0.4888
[50 50 50 50]	8	S=0.3600, V=0.5463
[40 40 40 40 40]	17	S=0.3882, V=0.5685

Table 9 optimization result using 5-feature dataset with normalization method (a)

(Some internal structure testing results are not shown here due to low sensitivity)

In pre-optimization, it can be found that [30 30 30 30] has the highest AUC value for both S and V. Thus, in the next stage, the optimization mainly focuses on internal structures that are similar to the structure of [30 30 30 30]. However, after testing with other number of hidden nodes (e.g. 29 and 31), it can be concluded that, internal structures with similar number of hidden nodes usually perform similarly. The same optimization process is also applied to the following three parts.

The ROC curve of the best internal structure with the optimal epoch is as below:

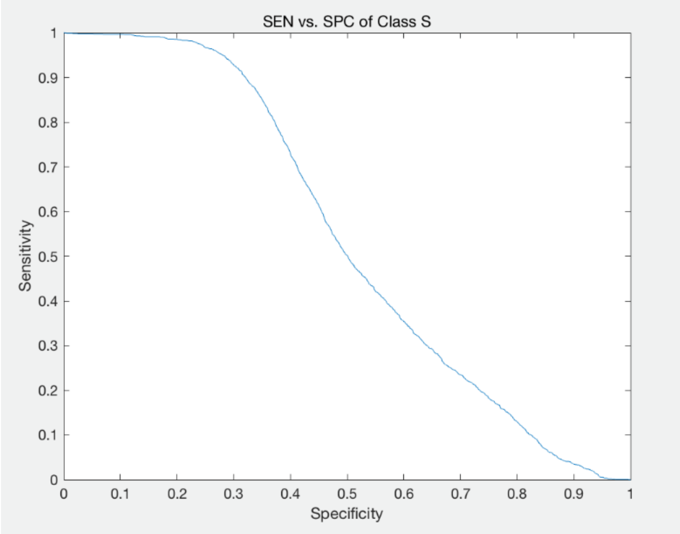


Figure 17 ROC curve of Class S

$$\mathrm{AUC_S} = 0.5728$$

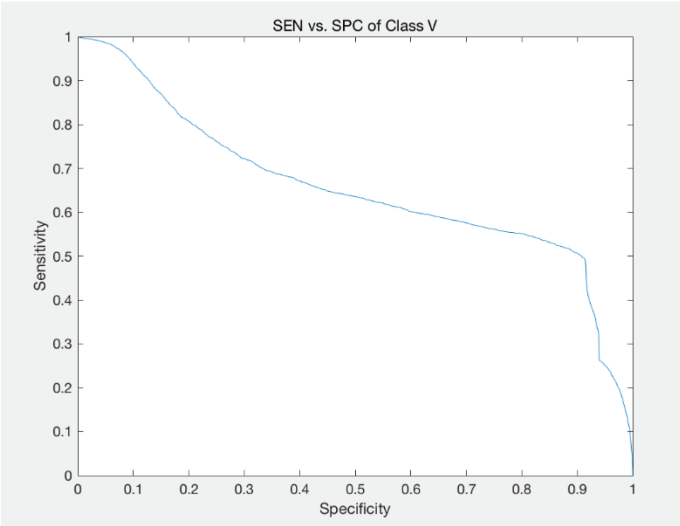


Figure 18 ROC curve of Class V

$$\mathrm{AUC_V} = 0.6596$$

4.2.2. 5-feature dataset with normalization method (b)

Internal structure	Epoch	AUC
[30 30 30]	5	S=0.3035, V=0.4692
[40 40 40]	5	S=0.3102, V=0.5214
[50 50 50]	5	S=0.3074, V=0.4824
[20 20 20 20]	10	S=0.4646, V=0.6305
[25 25 25 25]	9	S=0.5016, V=0.6214

[30 30 30 30]	**9**	**S=0.5463, V=0.6559**
[40 40 40 40]	7	S=0.4280, V=0.5974
[30 30 30 30 30]	16	S=0.3670, V=0.5350

Table 10 optimization result using 5-feature dataset with normalization method (b)

(Some internal structure testing results are not shown here due to low sensitivity)

The ROC curve of the best internal structure with the optimal epoch is as below:

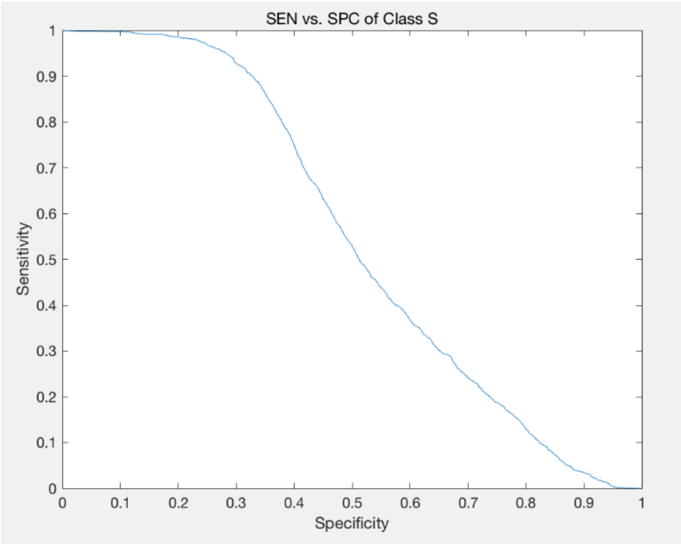


Figure 19 ROC curve of Class S

$$\mathrm{AUC_S} = 0.5463$$

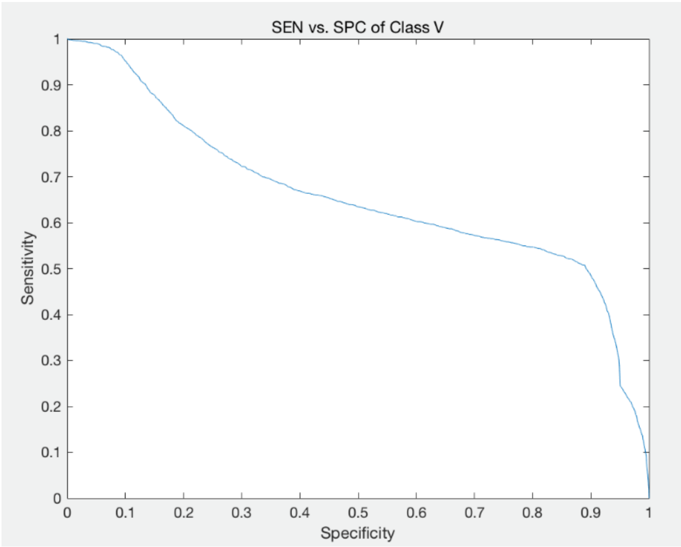


Figure 20 ROC curve of Class V

$$\mathrm{AUC_V} = 0.6559$$

4.2.3. 12-feature dataset with normalization method (a)

Internal structure	Epoch	AUC
[10 10 10]	9	S=0.6707, V=0.9133
[20 20 20]	6	S=0.6012, V=0.8939
[30 30 30]	6	S=0.5803, V=0.9000
[10 10 10 10]	5	S=0.6609, V=0.8713
[20 20 20 20]	6	S=0.5939, V=0.8974
[9 10 10]	**8**	**S=0.6788, V=0.9135**
[10 9 10]	5	S=0.5623, V=0.8758
[10 10 9]	5	S=0.5912, V=0.8633

Table 11 optimization result using 12-feature dataset with normalization method (a)

(Some internal structure testing results are not shown here due to low sensitivity)

The ROC curve of the best internal structure with the optimal epoch is as below:

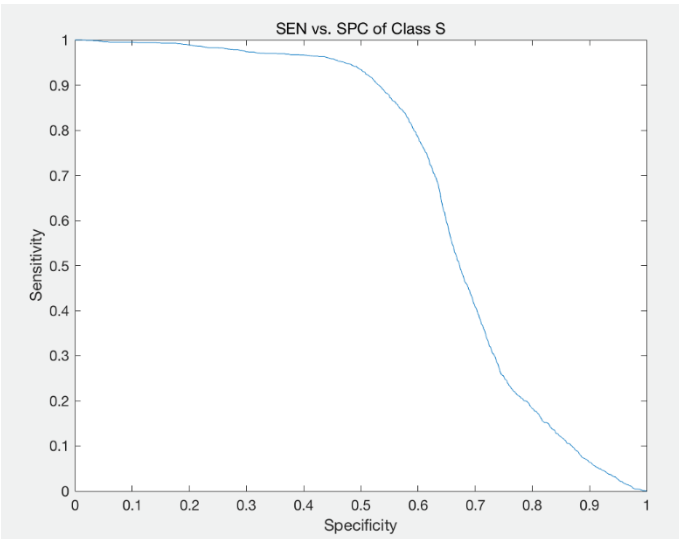


Figure 21 ROC curve of Class S

$$\mathrm{AUC_S} = 0.6788$$

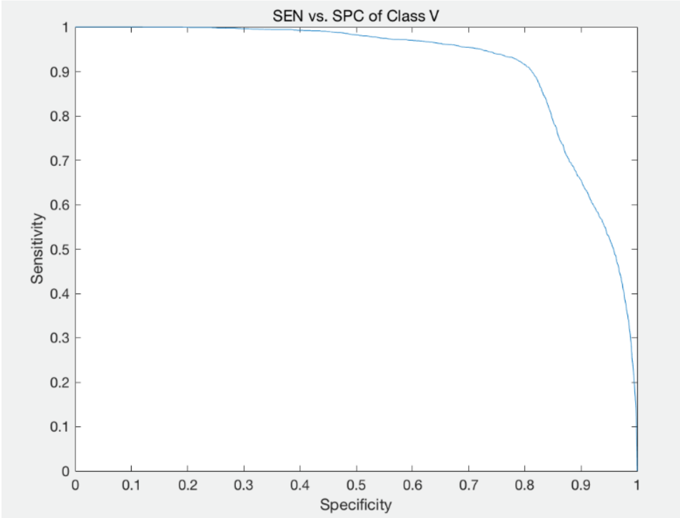


Figure 22 ROC curve of Class V

$$\mathrm{AUC_V} = 0.9135$$

4.2.4. 12-feature dataset with normalization method (b)

Internal structure	Epoch	AUC
[5 5 5]	5	S=0.3764, V=0.7995
[10 10 10]	**6**	**S=0.6294, V=0.9028**
[20 20 20]	5	S=0.5589, V=0.8602
[30 30 30]	5	S=0.5192, V=0.8340
[10 10 10 10]	5	S=0.6599, V=0.8778
[20 20 20 20]	5	S=0.5493, V=0.8665

Table 12 optimization result using 12-feature dataset with normalization method (b)

(Some internal structure testing results are not shown here due to low sensitivity)

The ROC curve of the best internal structure with the optimal epoch is as below:

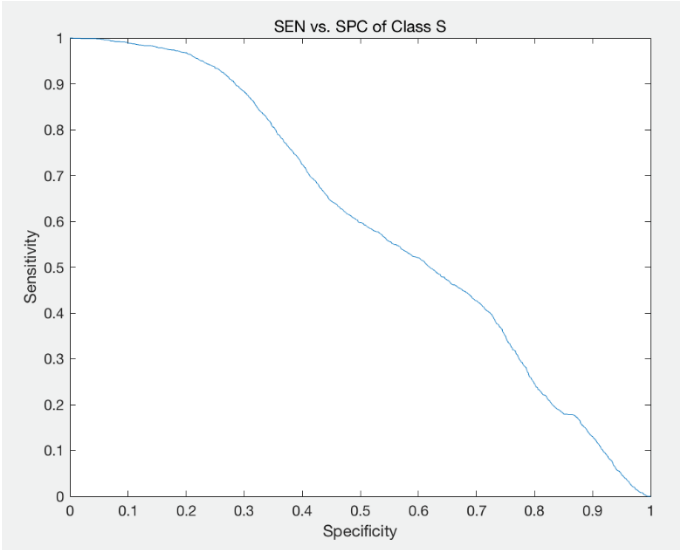


Figure 23 ROC curve of Class S

$$\mathrm{AUC_S} = 0.6294$$

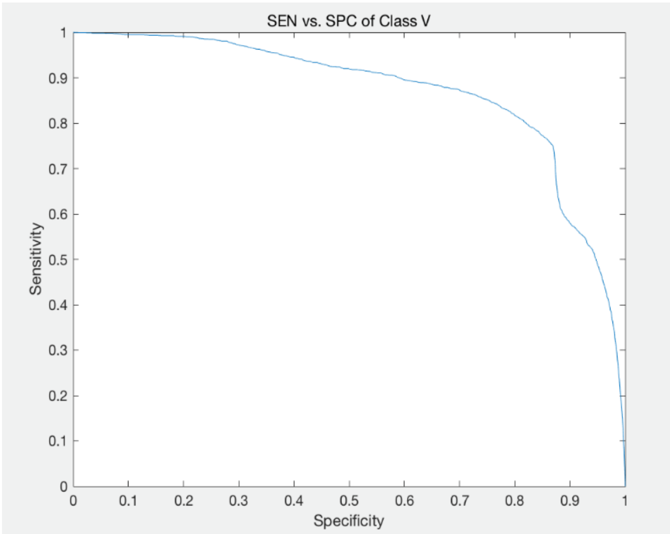


Figure 24 ROC curve of Class V

$$\mathrm{AUC_V} = 0.9028$$

4.2.5. Analysis of result

Below is the comparison between 4 datasets:

Datasets	Optimal structure	Best echo	Best AUC
5-feature dataset (a)	[40 40 40 40]	7	S=0.5728, V=0.6596
5-feature dataset (b)	[30 30 30 30]	9	S=0.5463, V=0.6559
12-feature dataset (a)	**[9 10 10]**	**8**	**S=0.6788, V=0.9135**
12-feature dataset (b)	[10 10 10]	6	S=0.6294, V=0.9028

Table 13 optimized result comparison

From the above results and graphs, there are five interesting observations in the optimization process:

a) The DNN with the 5-feature dataset performs worse than that with the 12-feature dataset.

   It is easy to find that the last two DNNs perform much better than the first two, especially for class V. The ROC curve in the first two datasets is still concave, or nearly straight (Figure 17-20). However, for the last two DNNs, the ROC curve is protruding to right upper corner, the expected shape of a good classifier (Figure 21-24).

b) The optimal internal structure of a DNN with the 12-feature dataset is usually less complicated than that with the 5-feature dataset.

   From Table 13, the optimal structures of a DNN with the 5-feature dataset are usually four hidden layers with 30 to 40 hidden nodes in each layer. On the contrary, the structures of DNNs with the 12-feature dataset are usually three hidden layers with approximately 10 hidden nodes in each layer, which are simpler than that with the 5-feature dataset.

c) Similar internal structures of a DNN usually give similar performance.

   From Table 11, the performance of internal structure [10 10 10], [9 10 10] and [10 10 9] are similar. However, the performance of internal structure [10 10 10], [20 20 20] and [30 30 30] varies.

d) The optimal structure of a DNN has same input nodes in each hidden layer.

   From Table 13, all of the four tested optimal structures have the similar shape: the number of hidden nodes are same within the same structure.

e) Setting default value to missing locations performs better.

   At the beginning, a default value 0 is assigned to missing locations and its corresponding features before normalization. The performance of a DNN with and without default setting is compared. The performance investigation shows that the performance of a DNN with default setting is usually better.

They can be explained as below:

a) For the 5-feature dataset, the ratio of hidden nodes to features are $30:5 = 6:1$ & $40:5 = 8:1$. This indicates that the internal structure is very complicated for the 5-feature dataset, and many noise nodes are included inside the trained model. Those noise nodes are irrelevant and may affect the nearby nodes. They contribute very little to the classification result and thus should be removed from the structure. The whole structure is very complicated because the input features are not enough to classify the heartbeat types. Thus, the performance is not good and the model is hard to train.

   For the 12-feature dataset, the ratio becomes $10:12 = 5:6$, which is approximately 1 and much smaller than the previous dataset. Meanwhile, the internal structure is less complicated with small number of hidden layers. This indicates that most of the 12 input features are useful for classification, and the number of noise nodes decreases. The ROC curve of class S and V also shows the improvement in performance.

   Thus, it can be concluded that, when training the DNN classifier, the DNN structure can help us to design the number of input features. When the ratio of hidden nodes to input features is too large, the classification performance could be unsatisfactory. More features should be added to obtain a better result.

b) As mentioned in (a), too many hidden nodes may increase the number of noise nodes in the structure. Thus, small change in hidden nodes can have limited influence on the performance.

c) From most of previous researches, the expected DNN structure should be a raised shape. The middle-hidden layers always contain more neurons than layers near input and output.

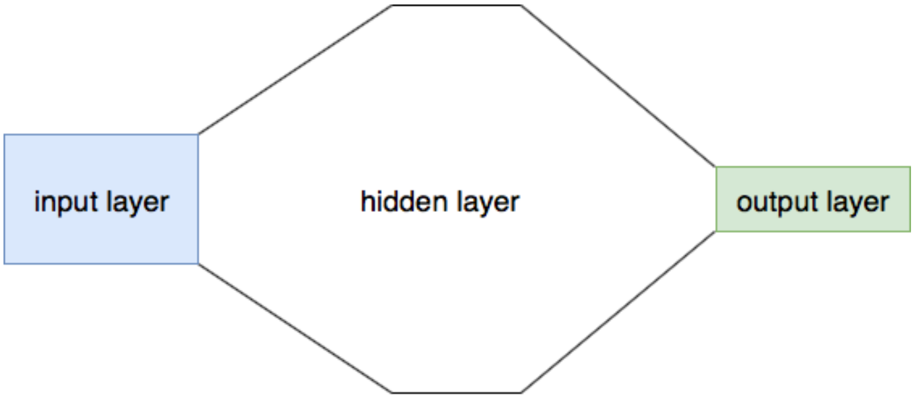


Figure 25 structure of common DNN classifiers

However, the optimal structure in this project is not a raised shape but a rectangle shape. The shape is shown as follow:

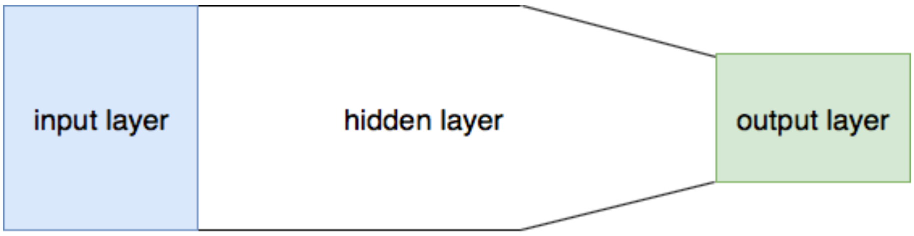


Figure 26 structure obtained in this project

This can be explained that in this project, weights of neurons in layers near input and out could be ineffective. This means some of the weights could be very close to zero, and thus may have very small influence on the whole structure. The weight of the first and last hidden layers is shown as follow (12-feature dataset with normalization method (a), internal structure [9 10 10], epoch = 8):

Weights of first hidden layer	Weights of last hidden layer
0.0424	0.1506
-1.9712	-0.0049
-0.5150	-0.1735
1.3031	-0.0022
0.3098	0.1010
-0.1306	-0.3777
3.2793	0.0245
0.0737	-0.4834
-0.2391	-0.3947
-1.6939	-0.2508
2.2771	-0.1289

Table 14 weight of first & last hidden layers optimal structure in Table 11

Weights with absolute value less than 0.1 are marked in yellow color. Those weights are less important than other neurons. Thus, the shape of the trained DNN model is roughly the same as most of the DNN structures in other papers.

4.3. Evaluation on DNN classifiers

4.3.1. 5-class DNN classifier

From the previous testing, the best DNN classifier in this project is using the 12-feature dataset and the normalization method (a), with the internal structure [9 10 10] and the optimal number of epochs 8. The AUC values for class S and class V are 0.6788 and 0.9135 respectively.

Confusion matrix is built for the selected optimal DNN structure:

	N	S	V	F	Q	Total
N	43345	1784	1511	278	5	46923
S	0	5	0	0	0	5
V	10	18	1117	99	2	1246
F	0	0	0	0	0	0
Q	0	0	0	0	0	0
Total	43355	1807	2628	377	7	48174

$$\text{Accuracy} = 92.30\%$$

$$Spec\text{ificity} = 99.995\%$$

For the best operating points of Class S, the ROC curve (Figure 21) shows that class S can achieve sensitivity above 80% with specificity over 60%, which is better than the previous sensitivity 5.66%.

For the best operating points of Class V, the ROC curve (Figure 22) shows that class V can achieve sensitivity above 90% with specificity over 80%. This is roughly the same performance (91.01%) as the previous project. Although the overall accuracy (92.30%) is less than the previous project (94.03%), the sensitivity of class S has improved from 5.66% to 80% with a specificity no less than 60%. This result is acceptable and there are two reasons for this:

a) The usage of Wavelet Transform toolbox

   The accuracy of using this toolbox to detect heartbeats is 96.67%, a bit lower when comparing with previous PT algorithm (98.46%). Thus, this could result in around 750 heartbeats mistakenly detected before feature extraction. Those beats are eliminated before they are finally built up the feature set. Thus, the testing sample is slightly smaller than the original whole dataset 2. However, PT algorithm can only detect QRS complex but not P and T waves. To extract detail position of P and T wave, other supplementary algorithms need to be tested. Although the use of Wavelet Transform is not seen in any other researches before, tested result shows that Wavelet Transform toolbox is an optimal choice for this project.

b) The reduction of input features.

417 data points were used as the input layer in previous project, which is roughly 35 times of our input features. The 417 data points are raw data, and serve as 417 features for the whole DNN classifier. As a result, the training time of the previous model is very long and its complexity is much higher. Moreover, from previous literature review, most of the algorithms can achieve 80+% accuracy in classification using over 22 features. Correspondently, using roughly half of the number of features, this project has achieved a very high performance in Class S and Class V.

c) Improvement of sensitivity and specificity in Supraventricular ectopic beat (Class S)

From literature review, due to the unbalanced data, Class S is more difficult to train than Class V. There are only 944 Class S in training data and 1807 in testing data. The sensitivity has greatly improved from 5.66% to 80% with a specificity over 60%, which is a big leap in the enhancement of classification.

4.3.2. 2-class DNN classifier

A two-class classifier is built and tested for classifying normal and abnormal heartbeats. Below is the confusion matrix:

	Normal	Abnormal	Total
Normal	42752	3838	46590
Abnormal	603	981	1584
Total	43355	4819	48174

$$\mathrm{Accuracy} = 90.78\%$$

$$\mathrm{Specificity}_{\mathrm{normal}} = 91.76\%$$

$$\mathrm{Specificity}_{\mathrm{abnormal}} = 61.93\%$$

$$\mathrm{Sensetivity}_{\mathrm{normal}} = 98.61\%$$

$$\mathrm{Sensitivity}_{\mathrm{abnormal}} = 20.36\%$$

The result shows that feature extraction may not be a desirable way to construct a 2-class classifier. The number of input features is far less than the previous project, thus, the performance could be no better than the previous one.

4.4. Improvement of web interface

The web interface is improved for commercial use.

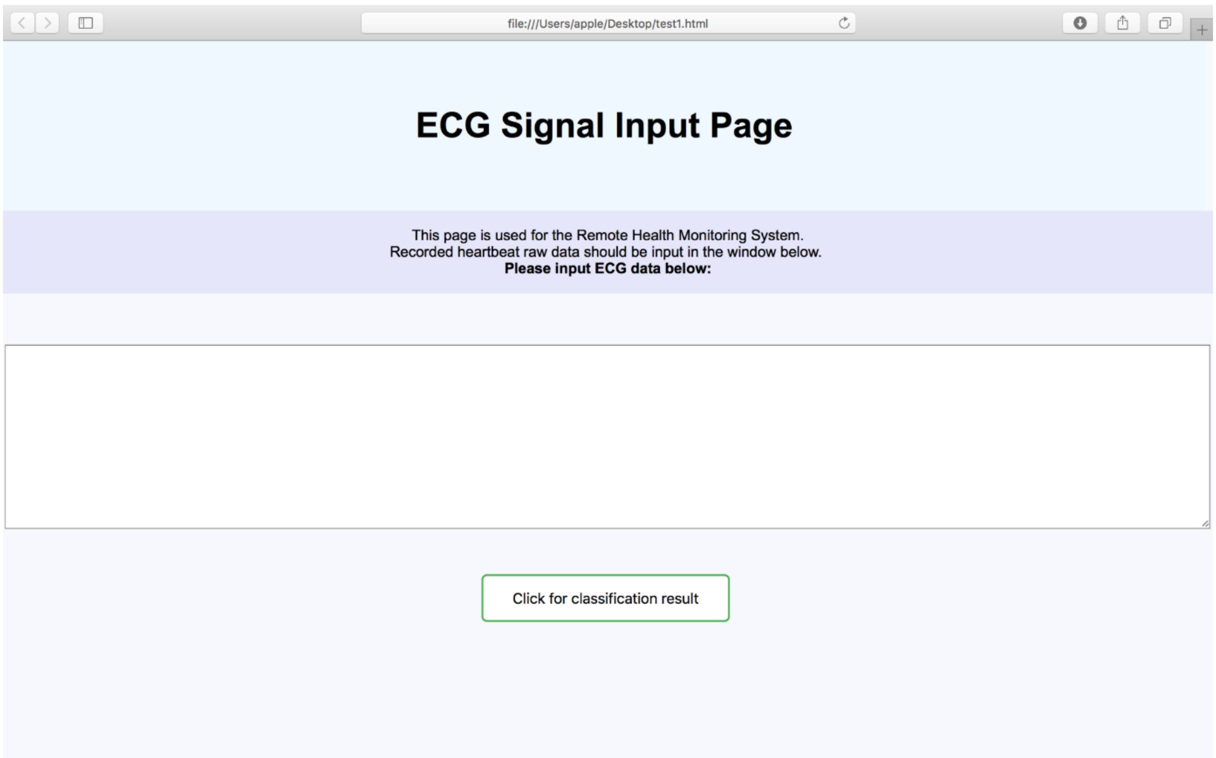


Figure 27 open page of web interface

When a normal or abnormal beat is detected:

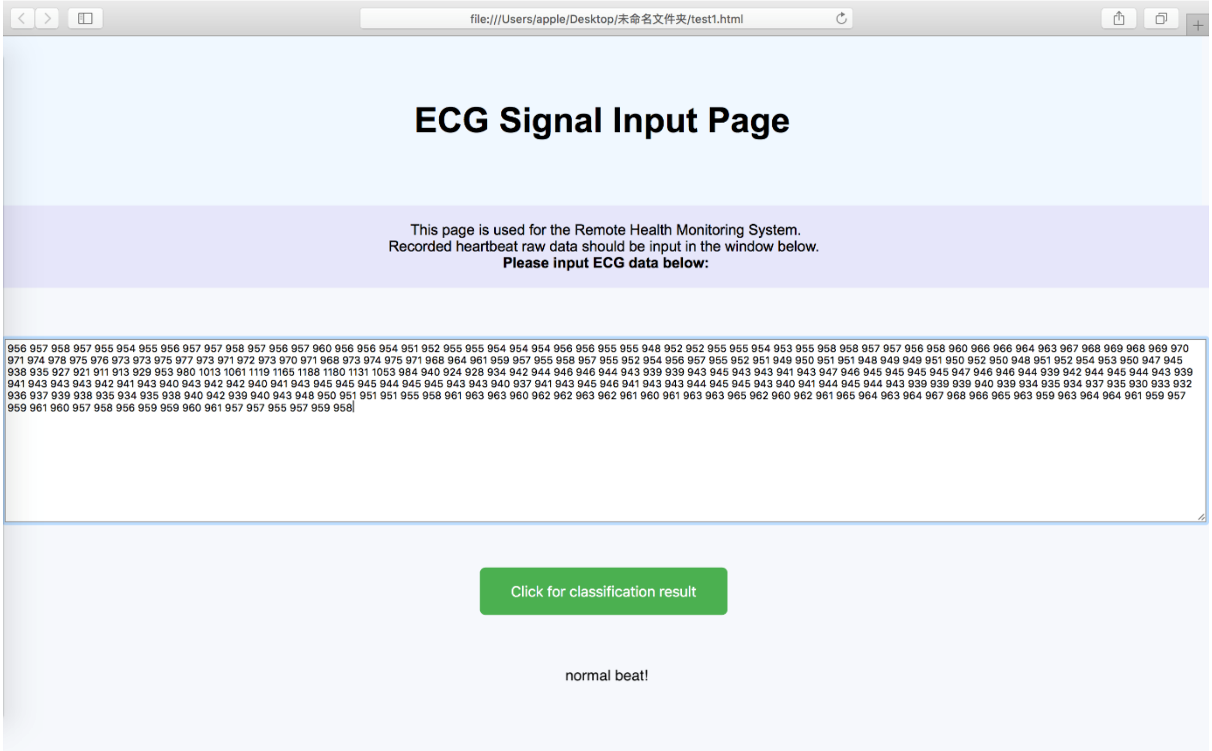


Figure 28 web page when normal beat is detected

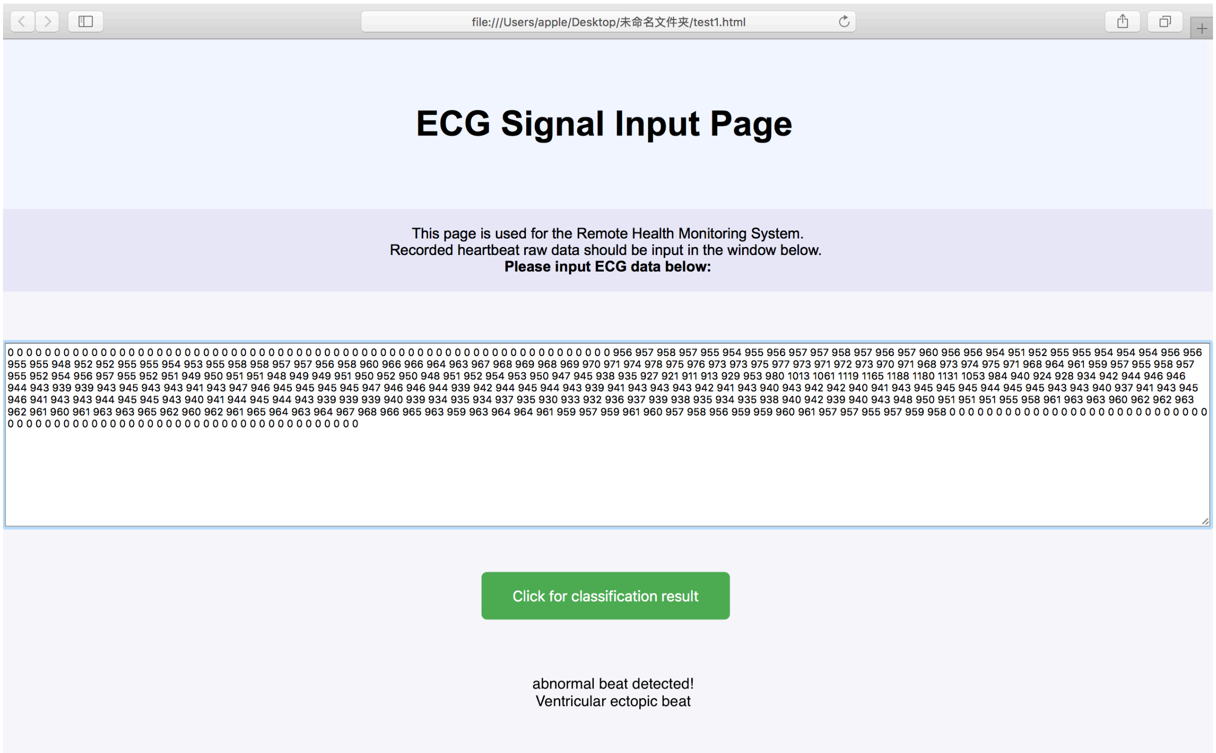


Figure 29 web page when abnormal beat is detected

4.5. Discussion: difficulties encountered and solutions

During the whole project, many difficulties were encountered. Among all the difficulties, there are three key turning points:

The first difficulty is the selection of ECG wave detection. From literature review, ecgpuwave and ECG-kit was first selected as target toolbox. However, in the testing period, both two toolboxes are difficult to use, and some of those functions are disabled due to unknown reasons. Initially, adjustments on toolbox algorithms are expected. However, with more than 30 functions in each toolbox, it is difficult to modify the program in a short time. Thus, other possible algorithms are sought, and finally selected the wavelet transform toolbox.

The second difficulty is the classifier selection. As planned, SVM should be used to improve the performance of the previous project. Thus, in the first stage, SVM was used for training and testing for 10 recordings (roughly 8000 beats for training and testing respectively), and the accuracy turned from 60% to 70%. However, when using the whole dataset, convergence cannot be achieved within the maximum number of iterations allowed. Although SVM can converge and give a relatively acceptable result when using 8000 data points, the whole dataset is still too complicated to find the decision boundary. Thus, SVM is abandoned and DNN classifier is chosen instead.

The third difficulty is the examination of optimization. Due to the imbalanced data, it is difficult to find the well-trained DNN model. Using maximum-likelihood method is difficult because the trained model always tends to classify the arrhythmia beats into normal ones. To find a better criterion, the ROC curve and the PR curve were then selected and compared. It turns out that ROC curve and AUC are the better ways to show the overall performance.

5. Conclusion

This project aims to develop DNN classifiers for ECG signals. All ECG data are obtained from the MIT-BIH database and are processed using the following procedure:

- Wave detection using the Wavelet Transform toolbox
- Heartbeat synchronization and default value setting
- Normalization using two methods

The well-trained DNN classifier is optimized and the details of the optimal structure are:

- Internal structure: 3 hidden layers, 9 nodes in first layer, 10 nodes in second layer and 10 nodes in third layer
- Overall accuracy of heartbeat classification: 92.30%
- Improved sensitivity and specificity of Supraventricular ectopic beat: 80% specificity with no less than 60% sensitivity

The overall performance of the DNN classifier can surely meet the objectives of the project. Considerable improvement has been made in the sensitivity of Supraventricular entopic beat (Class S). For future improvements, optimization in combining the two classifiers' results and integration of web server in a cloud system could be possible directions for a more mature, user-friendly and reliable product.

6. Reference

[1] J. White, “Arrhythmia”, *InnovAiT,* Vol.9(10), pp.717-625, 2016

[2] G.B. Moody & R.G. Mark, “The impact of the MIT-BIH Arrhythmia Database,” *IEEE Eng in Med and Biol,* vol. 20, no. 3, pp. 45-50, May – June 2001. https://222.physionet.org/physiobank/database/mitdb/

[3] ECG morphology, online: https://en.wikipedia.org/wiki/Electrocardiography#/media/File:SinusRhythmLabels.svg

[4] AHA database, online: https://www.aha.org/data-insights/aha-data-products, 2003

[5] European ST-T Database, online: http://www.physionet.org/physiobank/database/edb, 2003

[6] “Testing and Reporting Performance of Cardiac Rhythm and ST Segment Measurement Algorithm”, *Association for the Advancement of Medical Instrumentation*, 1998

[7] Y. Can, V.B. Vijaya Kumar & M.T. Coimbra, “Heartbeat Classification Using Morphological and Dynamic Features of ECG Signals”, *IEEE Trans. Biomedical Eng.*, vol. 59, no. 10, pp. 2930 – 2941, Oct. 2012

[8] LIBSVM 2000-2017 c.c. Chang & C.J. Lin, online: https://www.csie.ntu.edu.tw/%7Ecjlin/libsvm/

[9] L. Pablo, J. Raimon & B. Eudald, online: https://physionet.org/physiotools/ecgpuwave/src/matlab/

[10] PhysioNet, *National Institute of General Medical Science (NIGMS) and National Institute of Biomedical Imaging and Bioengineering (NIBIB)*, online: https://www.physionet.org

[11] CSE Database, France, online: http://www.iath.virginia.edu/~spw4s/CSE/CSE_database.frame.html, 2003

[12] J. Pan and W.J. Tompkins, "A Real-Time QRS Detection Algorithm", *IEEE Transactions on Biomedical Engineering*, 32(3):230-236, 1985

[13] P. Laguna, "Application to Long-term Records", *New Electrocardiographic Signal Processing Techniques*, 1990

[14] A.J. Demski & S.M. Llamedo, "ecg-kit a Matlab Toolbox for Cardiovascular Signal Processing," *J. Open Research Software,* vol. 4, no. 1, p. e8, 2016 http://doi.org/10.5334/jors.86

[15] J.P. Martinez, R. Almeida & A.P. Rocha, "A wavelet-based ECG delineator: evaluation on standard databases," *IEEE trans. Biomedical Eng.*, vol. 51, no. 4, pp. 570 – 581, Apr. 2004

[16] F. Jingwei & T. Fengchun, "a background elimination method based on wavelet transform in wound infection detection by electronic nose", *Sensors and Actuators B: Chemical,* vol. 157, no. 2, pp 395 – 400, Oct. 2011

[17] C. Distante, M. Leo, P. Siciliano, K.C. Persaud, "On the study of feature extraction methods for an electronic nose", *Sensors and Actuators B: Chemical,* vol. 87, pp 274 – 288, 2002

[18] W. yuansan, online source, available at: *http://www.lxway.com/492491962.htm*

[19] J. Tae Joon, P. Hyun Ji & M. Nguyen Hoang, “Premature ventricular contraction beat detection with deep neural networks”, *IEEE International Conference on Machine Learning and Applications,* 2016

[20] L. Kunyang, P. Weifeng & L. Yifan, “A method to detect sleep apnea based on deep neural network and hidden Markov model using single-lead ECG signal”, *Neurocomputing*, Vol.294, pp.94-101, 2018

[21] Bradley & P. Andrew, “The use of the area under the ROC curve in the evaluation of machine learning algorithms”, *Pattern Recognition*, Vol.30(7), pp. 1145-1159, 1997

[22] Y. Zhiyong, Z. Taohong, L. Jingcheng, “Optimizing area under the ROC curve via extreme learning machines”, *Dorothy Knowledge-Base System*, Vol.130, pp. 74-89, 2017

[23] J.A. Hanley & B.J. McNeil, “The meaning and use of the area under a receiver operating characteristic (ROB) curve”, *Radiology*, 143(1) (1982), pp. 29-36