Vital Responder: Real-time Health Monitoring of First-Responders

Ye Can[1,2]

Advisors: *Miguel Tavares Coimbra*[2], *Vijayakumar Bhagavatula*[1]

[1] Department of Electrical & Computer Engineering, Carnegie Mellon University, USA
[2] Instituto de Telecomunicações, Faculdade de Ciências da Universidade do Porto, Portugal

Porto, Mar 31, 2011

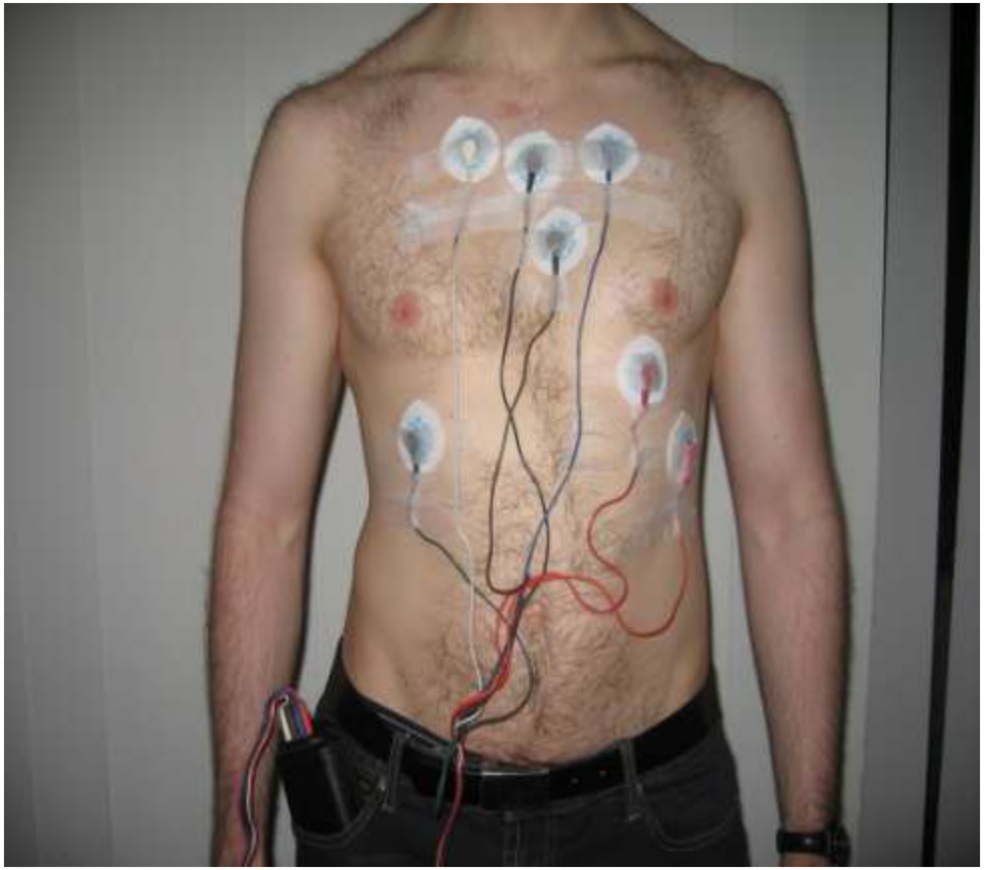
Information and Communication Technologies Institute
Carnegie Mellon | PORTUGAL


BlackBerry


BlackBerry


ECG
Watch real-time electrocardiograms on the iPhone!
Instant ECG
App Store

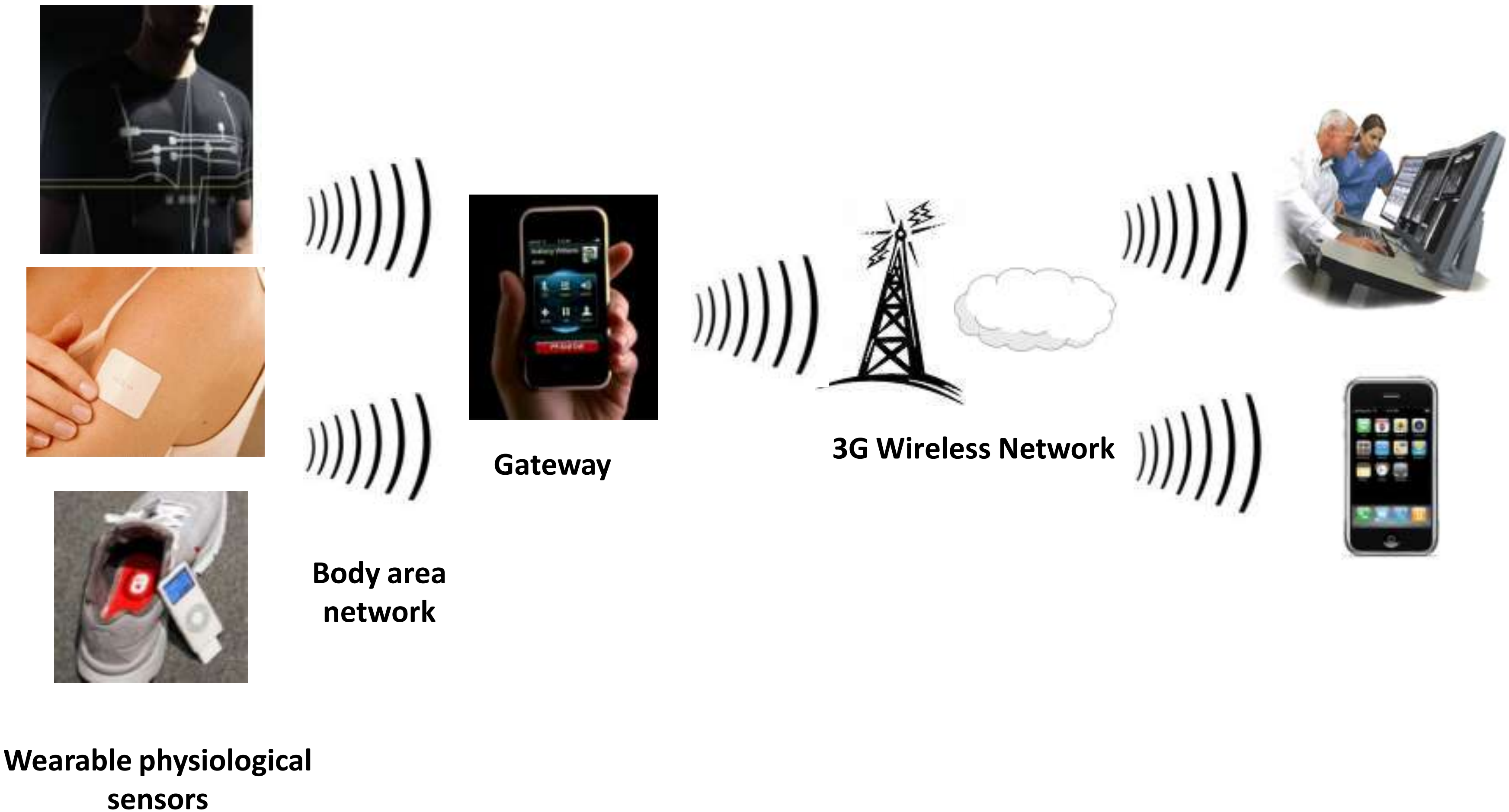
Holter Monitoring



iRhythm

Cutting Edge of Healthcare

Wearable Health Monitoring Systems



Motivating Factors

- Smartphones and >4 Billon Cell Phone Users
- Broadband 3G and Pervasive Connectivity
- Aging Trend and Increasing Healthcare Costs
- Technological Advance, *e.g. Miniature Biosensors, Smart Textiles, Microelectronics etc*.

Wearable Health Monitoring Systems

- **Requirements:**
  - Ubiquitous, e.g. at home
  - Unobtrusive
  - Long-term (24/7/365)
  - Low-cost

- **Potential Applications:**
  - Daily use for common people, *e.g. elderly population*
  - Patients suffering from chronic diseases, *e.g. diabetes patients*
  - People at risk, *e.g. firemen, soldiers*

Vital Responder Project

- *What is Vital Responder?*

"Vital Responder is a project with the main goal to provide secure, reliable and effective first-response systems in critical emergency scenarios."

- *What is the motivation of Vital Responder?*

Critical events can induce stress in first responder professionals (e.g., fire fighters, policemen, paramedics). For example, heat stress can cause heat stroke and heat exhaustion for fire fighters.

Signal Processing of Vital Signals

- **"Silver Bullet"- VitalJacket®**

  A wearable ECG textile shirt that could collect real-time physiological signals and information with a variety of embedded sensors: ECG, Skin Conductance, Temperature, O_2 sensor, GPS, Accelerometer, etc.

- **ECG signals are commonly used to measure stress in human beings.**

- **The goal is to quantify stress and fatigue levels and detect medical emergency events in critical situations from analyzing vital signs (e.g., ECG signals).**

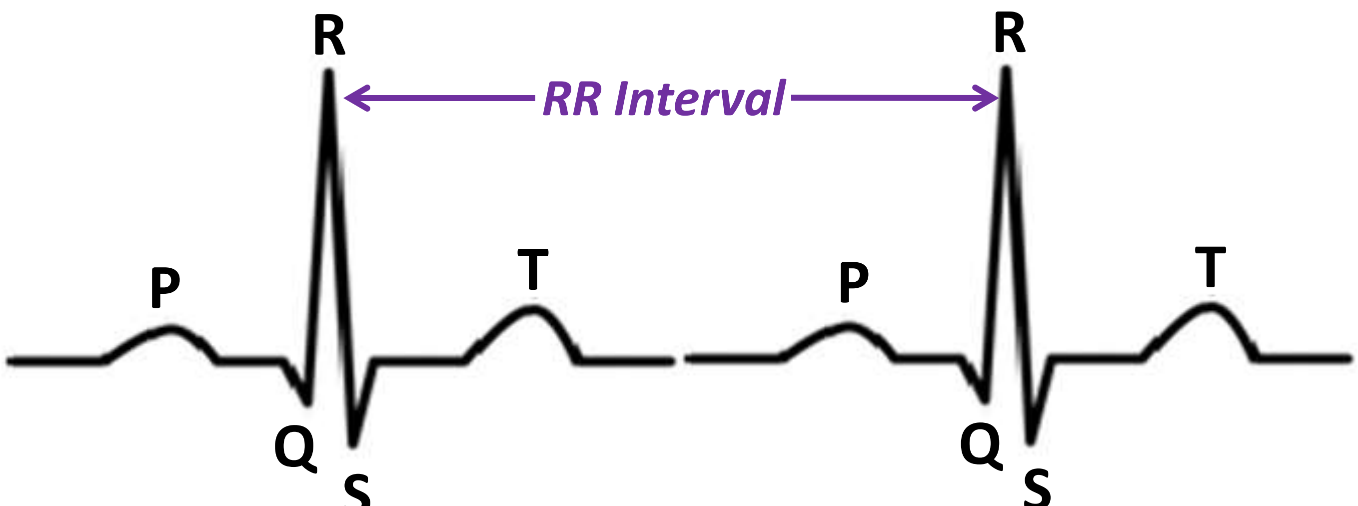
Problem Definition

- Arrhythmias are abnormal heart rhythms, which cause the heart to beat too fast (tachycardia) or too slow (bradycardia) and to pump blood less effectively[1].
- Some types of arrhythmias are life threatening medical emergencies and could trigger cardiac arrest or sudden death.
- Propose an approach to detect and classify arrhythmias from electrocardiogram (ECG) signals of the MIT-BIH Arrhythmias Database with an average accuracy of 99.91% for 15 classes of heartbeats (normal and 14 classes of arrhythmias).

[1] K. Robert *et al.*, "Basis and Treatment of Cardiac Arrhythmias", 1st Ed., Springer, 2006.

Electrocardiogram (ECG) Basics

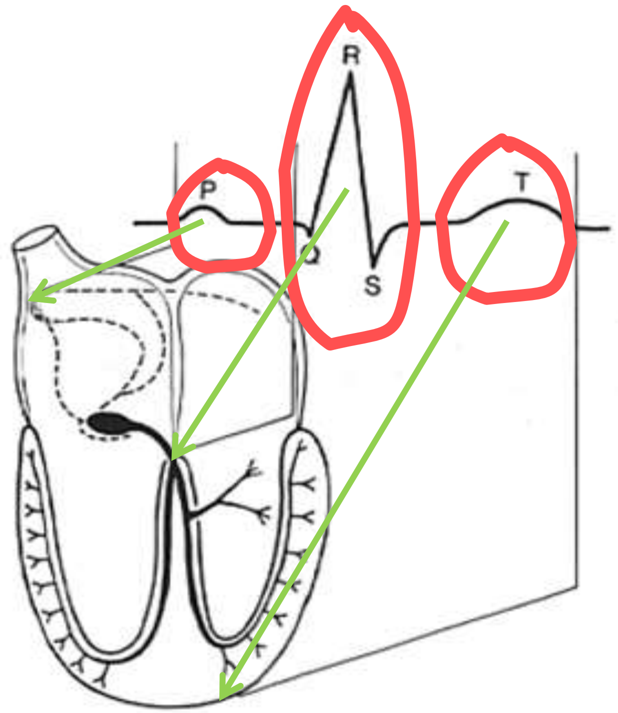
- ECG: the quantitative measure of the electrical activity of heart over time.
- One heart cycle consists of P wave, QRS wave (QRS composite or QRS complex) and T wave.
- RR interval is a measure of instantaneous heart rhythm.



Characteristic ECG waveform of one heart cycle[2]

[2] A.U. Rajendra et al., "Advances in cardiac signal processing", 1st Ed., Springer, 2009

ECG Basics



Physiological meaning behind ECG signal[2]

P wave ***- Atrial depolarization***

QRS complex ***- Ventricular depolarization***

T wave ***– Ventricular repolarization***

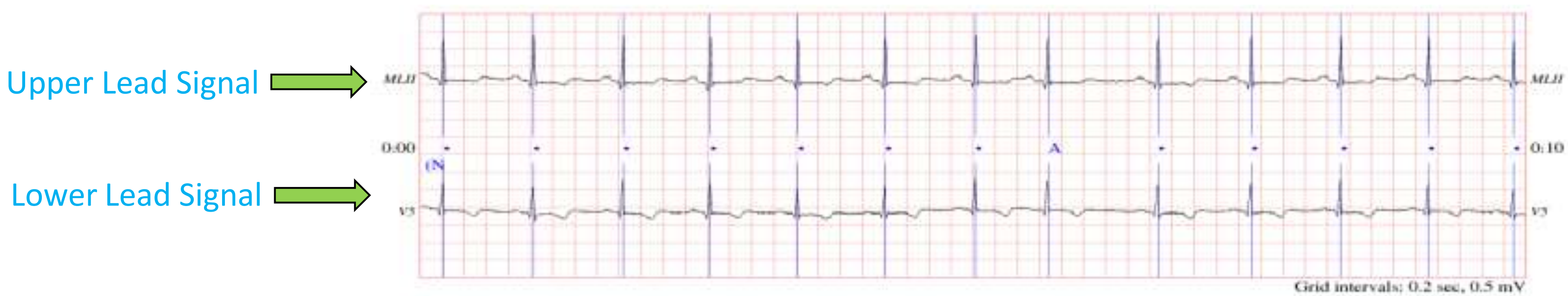
[2] A.U. Rajendra *et al.*, “Advances in cardiac signal processing”, 1st ed., Springer, 2009

Experimental Database

- ***Database***

  MIT-BIH Arrhythmias Database (regarded as the benchmark database in the area of arrhythmias detection and classification)

- ***Records***
  - 48 half-hour records and 15 classes of heart beats
  - Sampled at a rate of 360 Hz
  - Includes two leads signals: upper lead signal and lower lead signal
  - Includes annotation information: the location of R peaks (i.e., the occurrence of the heartbeat) and the type of heartbeat



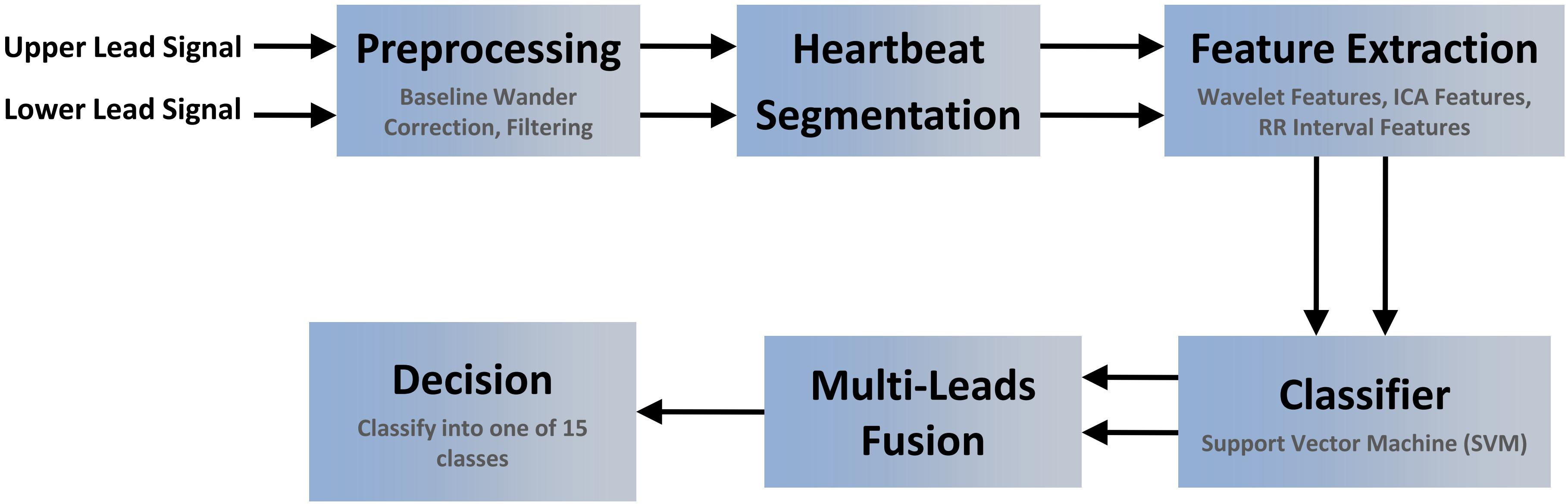
Many thanks to MIT-BIH Arrhythmia Database for the figure

Experimental Setup

- *Segmentation*
  - Annotation information regarding the location of the R peaks is used to segment the data.
  - Each segment is 300 samples long (0.83s), with 100 samples before the R peaks and 200 samples after.
- *Dataset includes entire database (all 110,076 heartbeat instances for 15 classes)*

Beat Type	Type Annotation	Total Number	Training Number	Training Percent
Normal Beat	N	75017	9753	13%
Left Bundle Branch Block	L	8072	3229	40%
Right Bundle Branch Block	R	7255	2902	40%
Aterial Premature Contraction	A	2546	1019	40%
Premature Ventricular Contraction	V	7129	2852	40%
Paced Beat	P	7024	2810	40%
Aberrated Atrial Premature Beat	a	150	75	50%
Ventricular Flutter Wave	!	472	284	60%
Fusion of Ventricular and Normal Beat	F	802	401	50%
Blocked Atrial Premature Beat	x	193	97	50%
Nodal (junctional) Escape Beat	j	229	115	50%
Fusion of Paced and Normal Beat	f	982	491	50%
Ventricular Escape Beat	E	106	53	50%
Nodal (junctional) Premature Beat	J	83	42	50%
Atrial Escape Beat	e	16	8	50%
Total	15	110076	24131	22%

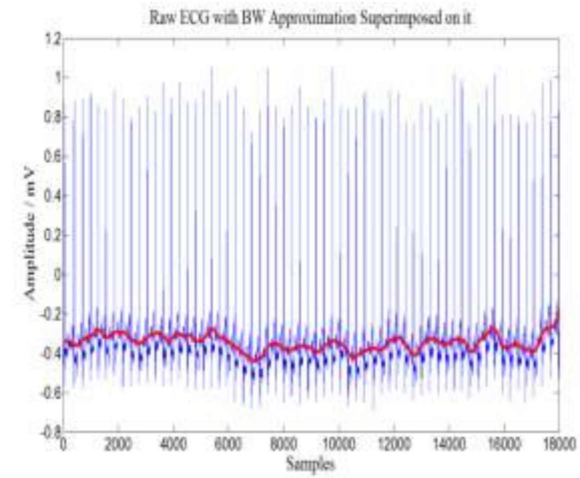
Methodology Overview



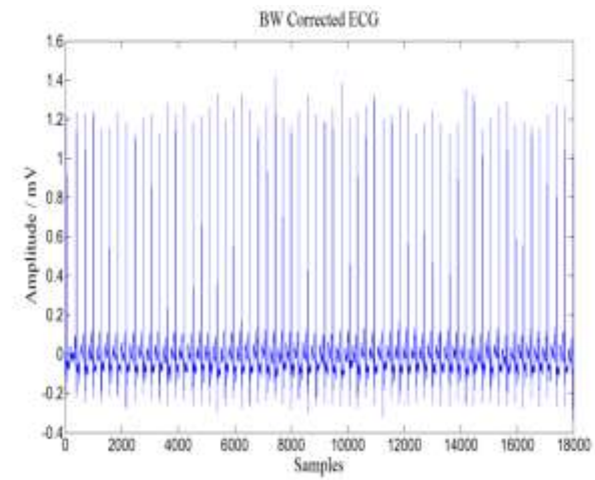
Information and Communication Technologies Institute
Carnegie Mellon | PORTUGAL

Preprocessing

- ***Baseline Wander (BW) Correction***
  - Baseline Wander is one common type of low-frequency artifact in biomedical signals due to human respiration



- ***Band-Pass Filtering:***
  - The energy of the QRS Complex wave is concentrated from 0.5 Hz to 40 Hz[3]
  - Band-Pass filter to emphasize the QRS Complex

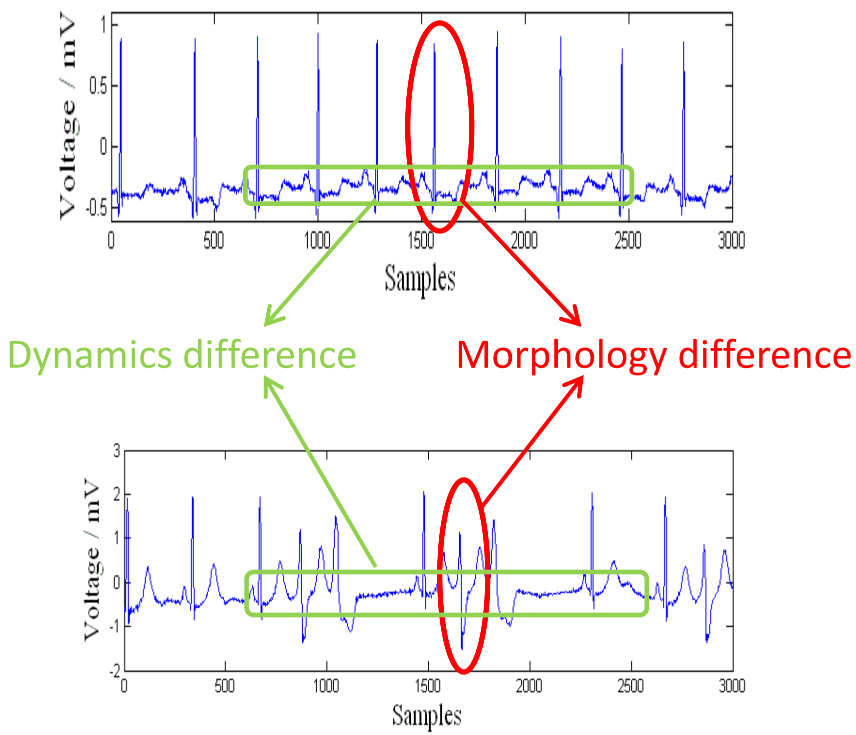


[3] N. V. Thakor, J. G. Webster, W. J. Tompkins, "Estimation of QRS Complex Power Spectra for Design of a QRS Filter," *IEEE Trans. Biomed. Eng.*, vol. 31, no. 11, pp. 702-706, 1984.

Morphological Features and Dynamic Features

- **Motivation:**
  - Arrhythmias are different from normal heartbeats in terms of both morphology and dynamics
- **Morphological Features:**
  - Wavelet features and ICA features
  - Provides a representation of a single heartbeat
- **Dynamic Features:**
  - RR Interval features
  - Characterizes the ‘rhythm’ around the heartbeat

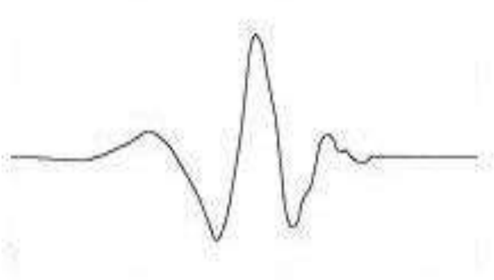
Normal ECG Sequence



Arrhythmia ECG Sequence

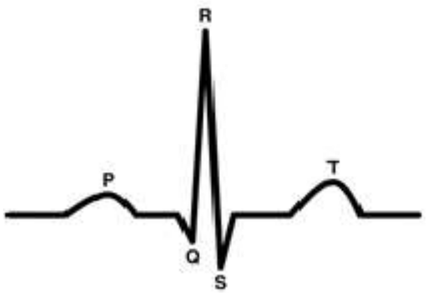
Wavelet Features

- ***Analyze the ECG signals using wavelets:***
  - The Discrete Wavelet Transform (DWT) is suitable to characterize the energy distribution of non-stationary signals, such as ECG signals.
- ***Choice of Wavelets***
  - Select Daubechies Wavelets of order 8 due to their similarity with the most characteristic ECG waveforms.



Daubechies wavelets of order 8

Look similar?

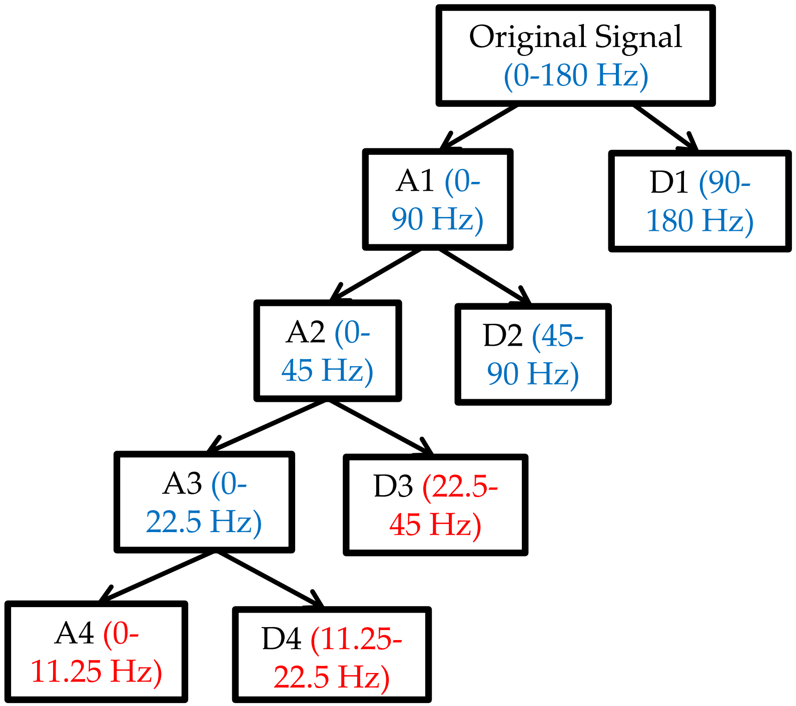


Characteristic ECG waveform

Many thanks to Mindy Schockling and Rajendra for the figures

Wavelet Features

- ## *Choice of Wavelet Features*
  - Given the sampling frequency 360 Hz, the highest frequency in the ECG signals is 180 Hz
  - The energy of QRS Complex is concentrated from 0.5 Hz to 40 Hz[3]
  - Applying 4-level wavelet decomposition, the spectrum is covered by the detail coefficients at level 3 (D3) and 4 (D4) as well as the approximation coefficients at level 4 (A4)
  - D3, D4, A4 coefficients are selected as our wavelet features (118 wavelet features for each heartbeat)



4-Level Wavelet Decomposition

[3] N. V. Thakor, J. G. Webster, W. J. Tompkins, "Estimation of QRS Complex Power Spectra for Design of a QRS Filter," *IEEE Trans. Biomed. Eng.*, vol. 31, no. 11, pp. 702-706, 1984.

Independent Component Analysis (ICA)

- ***Independent Component Analysis (ICA):***
  - Proposed to solve the blind source separation (BSS) problem
  - ICA separates the observed signals into a set of underlying independent sources

- ***Mathematical Formulation:***

$$x = As$$

x: Matrix of N observations (110076*300)

A: Mixture matrix (N x M) (110076*18)

s: Matrix of M independent sources (18*300)

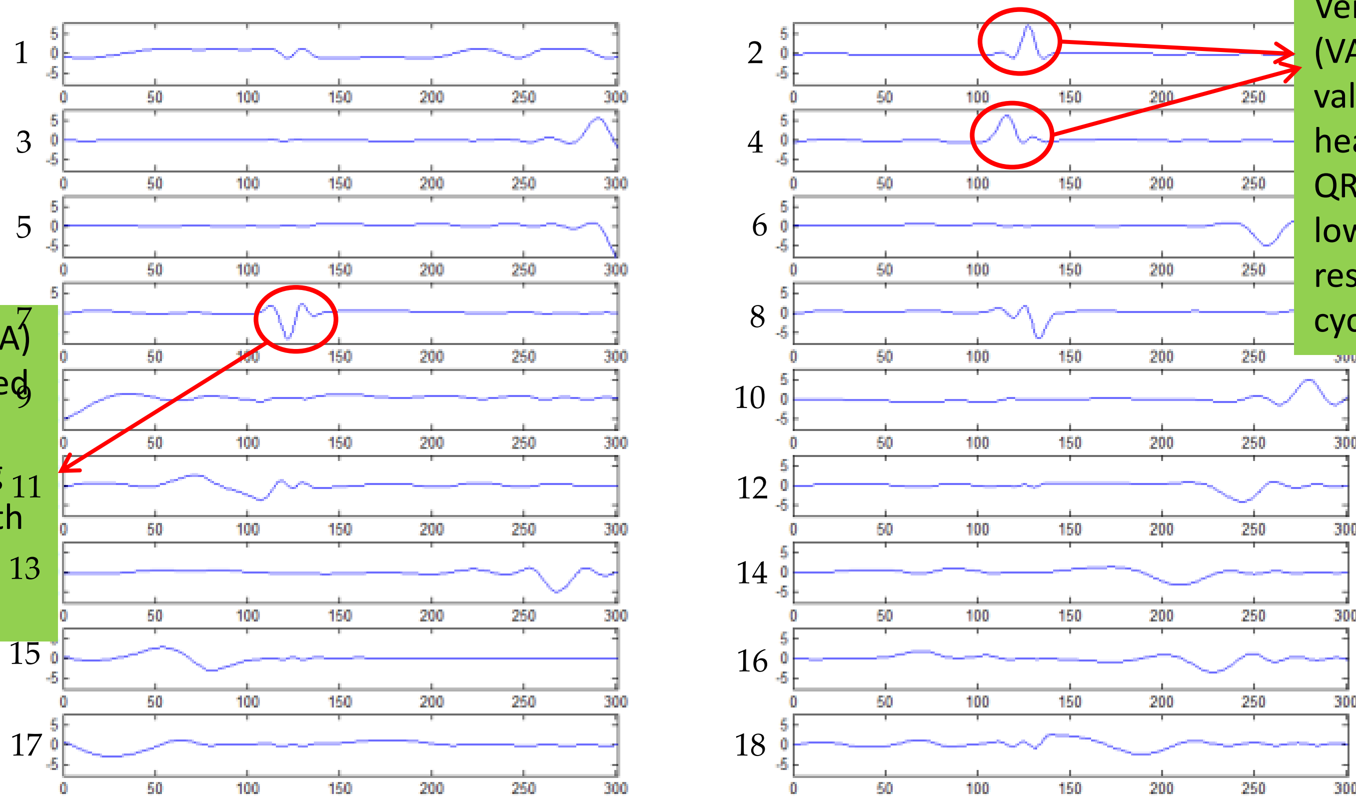
- ***Justification of the use of ICA for ECG signals[4]***
  - Atrial activity (AA) and ventricular activity (VA) are generated by independent physiological sources
  - Both of AA and VA present non-Gaussian distributions

[4] J.J. Rieta *et al.*, "Atrial activity extraction for atrial fibrillation analysis using blind source separation," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 7, pp. 1176-1186, 2004.

ICA Features

- ***Extracting the ICA Features:***
  - Utilized approximately 10,000 normal beats to train 18 ICA bases
  - Obtained 18 ICA coefficients for each heartbeat
- ***Physiological meaning behind ICA components:***
  - Three ICA components seem to be well matched to ventricular activity (VA) and atrial activity (AA)
  - These three components are dominant in representing one heartbeat since the coefficients of the other components are close to zero

Some Insights into the 18 ICA Components

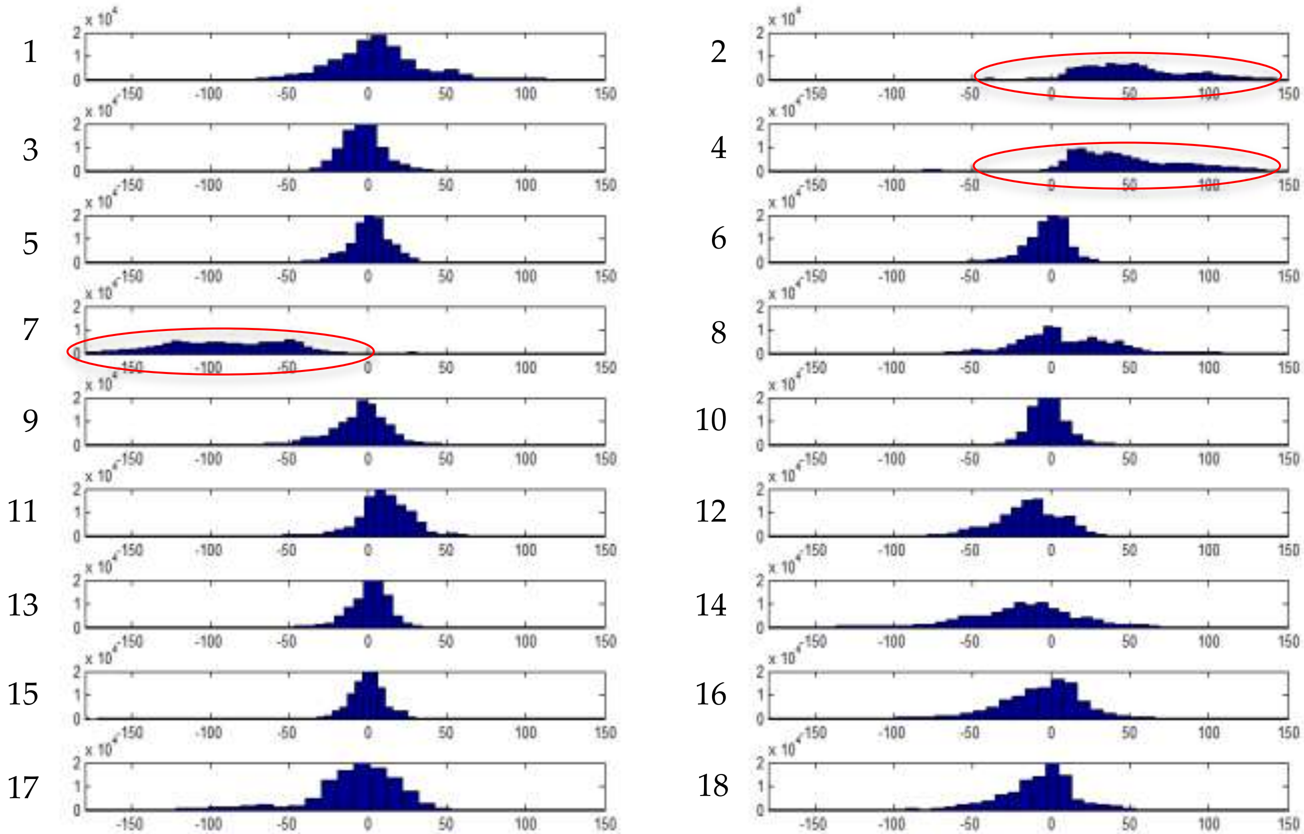


Ventricular activity (VA) presents high values within the heartbeat (i.e., the QRS complex) and low values in the rest of cardiac cycle[4]

Atrial activity (AA) could be modeled as a sawtooth signal consisting of a sinusoid with several harmonics[4]

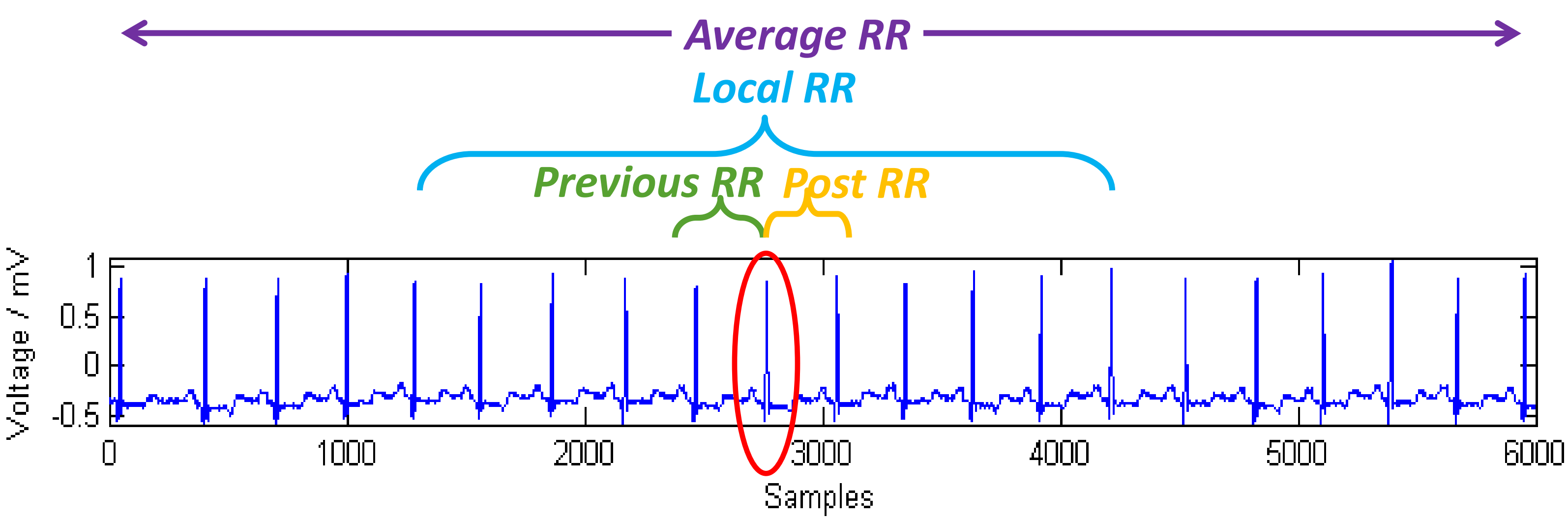
[4] J.J. Rieta *et al.*, "Atrial activity extraction for atrial fibrillation analysis using blind source separation," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 7, pp. 1176-1186, 2004.

Histograms of the ICA Coefficients



Dynamic Features

- *Dynamic Features*
  - Provides a characterization of the 'dynamics' around the corresponding heartbeat.
  - Four RR interval features: previous RR interval, post RR interval, local RR interval, average RR interval.



Combining Morphological and Dynamic Features

- **Morphological Features: 118 wavelet features and 18 ICA features**
  - Apply Principal Component Analysis to reduce the feature dimensionality to 26 that accounts for 99.3% variance
- **Dynamic Features: four RR-interval features**
- **Concatenate morphological features and dynamic features to obtain the final feature vector with a dimensionality of 30**

Classification: SVM

- Select Support Vector Machine (SVM) and Gaussian RBF Kernel for classification
- Use 10-Fold Cross Validation to select model parameters based on the training data
- Train a SVM classifier using the training dataset
- Evaluate the testing dataset with the trained SVM classifier

Two-Lead Fusion

- ***Motivation:***
  - Previous work typically use only upper lead signal.
  - Normal heartbeats are usually more prominent in the upper signal while the abnormal beats are more visible in the lower one.
- ***How?***
  - Reject inconsistently classified samples and reserve them for physicians' examination or for further classification.
  - Using both lead signals boosts our classification confidence.

Upper Lead Signal

Lower Lead Signal

Many thanks to MIT-BIH Arrhythmia Database for the figure

Results

Summary of the results on both leads

	Upper Lead	Lower Lead
Total Number of Instances	85945	85945
Correctly Classified Instances	85300	85158
Average Classification Accuracy (%)	99.25	99.08
Mean Squared Error	0.1377	0.1537
Squared Correlation Coefficient	0.9590	0.9542

A measure regarding the goodness of fit of a model that provides information how well future data will be predicted

Results

	N	L	R	A	V	P	a	!	F	x	j	f	E	J	e	∑
N	**64844**	8	12	73	307	4	3	6	4	0	1	1	0	1	0	**65264**
L	14	**4822**	0	0	6	0	1	0	0	0	0	0	0	0	0	**4843**
R	2	0	**4347**	1	3	0	0	0	0	0	0	0	0	0	0	**4353**
A	64	0	3	**1454**	6	0	0	0	0	0	0	0	0	0	0	**1527**
V	38	1	1	2	**4231**	1	0	0	2	0	0	0	1	0	0	**4277**
P	1	0	0	0	0	**4212**	0	1	0	0	0	0	0	0	0	**4214**
a	7	0	0	0	0	0	**65**	3	0	0	0	0	0	0	0	**75**
!	2	0	0	0	5	0	6	**172**	1	2	0	0	0	0	0	**188**
F	8	0	0	0	5	1	0	0	**385**	0	0	0	0	2	0	**401**
x	2	0	0	0	5	0	0	2	1	**86**	0	0	0	0	0	**96**
j	1	0	0	0	0	0	0	0	5	0	**107**	0	0	1	0	**114**
f	4	0	0	0	1	0	0	0	2	0	3	**477**	2	1	1	**491**
E	1	0	0	0	0	0	0	0	0	0	0	0	**52**	0	0	**53**
J	6	0	0	0	0	0	0	0	0	0	0	1	0	**34**	0	**41**
e	2	0	0	0	0	0	0	0	0	0	0	0	0	0	**6**	**8**
∑	**64996**	**4831**	**4363**	**1530**	**4569**	**4218**	**75**	**184**	**400**	**88**	**111**	**479**	**55**	**39**	**7**	**85945**

Confusion matrix of the result from upper lead
signal (Average classification accuracy: 99.25%)

Results

- By excluding 1.4% (1238) beats that are inconsistently classified between two lead signals, we obtain a final arrhythmia detection accuracy of 99.93% (84644/84707) and a heartbeat classification accuracy of 99.91% (84630/84707)

	N	L	R	A	V	P	a	!	F	x	j	f	E	J	e	∑
N	**64447**	0	0	5	23	0	0	0	0	0	0	0	0	0	0	**64475**
L	0	**4819**	0	0	1	0	0	0	0	0	0	0	0	0	0	**4820**
R	0	0	**4341**	0	0	0	0	0	0	0	0	0	0	0	0	**4341**
A	18	0	1	**1412**	3	0	0	0	0	0	0	0	0	0	0	**1434**
V	6	0	0	0	**4043**	0	0	0	0	0	0	0	0	0	0	**4049**
P	1	0	0	0	0	**4209**	0	0	0	0	0	0	0	0	0	**4210**
a	2	0	0	0	0	0	**65**	0	0	0	0	0	0	0	0	**67**
!	0	0	0	0	0	0	5	**162**	0	0	0	0	0	0	0	**167**
F	2	0	0	0	0	0	0	0	**376**	0	0	0	0	1	0	**379**
x	0	0	0	0	3	0	0	0	1	**86**	0	0	0	0	0	**90**
j	0	0	0	0	0	0	0	0	0	0	**101**	0	0	0	0	**101**
f	1	0	0	0	0	0	0	0	0	0	0	**478**	0	0	0	**479**
E	1	0	0	0	0	0	0	0	0	0	0	0	**51**	0	0	**52**
J	4	0	0	0	0	0	0	0	0	0	0	0	0	**33**	0	**37**
e	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**6**	**6**
∑	**64482**	**4819**	**4342**	**1417**	**4073**	**4209**	**70**	**162**	**377**	**86**	**101**	**478**	**51**	**34**	**6**	**84707**

Confusion matrix of the final result of incorporating results from two leads

Results

Summary of results on each class

Beat Type	# of rejected beats	Final Sensitivity	Final Specificity
N	789	99.95%	99.86%
L	23	100%	99.99%
R	12	99.99%	100%
A	93	99.65%	99.97%
V	228	99.26%	99.99%
P	4	100%	99.99%
a	8	92.86%	99.99%
!	21	100%	99.99%
F	22	99.99%	99.99%
x	6	100%	99.99%
j	13	100%	100%
f	12	100%	99.99%
E	1	100%	99.99%
J	4	97.06%	99.99%
e	2	100%	100%

Comparison with the Benchmark

	Proposed Method	Benchmark Paper[5]
Class Number	15	14
Instances Number	110,076	103,898
Preprocessing	BW Correction + Filtering	Filtering
Features	WT (D3, D4, A4) + ICA + RR	WT (A4) + ICA
Feature Dimensionality Reduction	PCA	Mutual Information Selection
Classifier	SVM	SVM
Lead Information Fusion	Yes	No
Classification Accuracy (%)	99.9%	98.6%
Mean Squared Error	0.1377	0.4378
Squared Correlation Coefficient	0.9590	0.8224

[5] X. Jiang *et al.*, “ECG Arrhythmias Recognition System Based on Independent Component Analysis Feature Extraction,” *IEEE TENCON*, 2006.

Comparison with the Benchmark

- Applied the algorithm of the benchmark on my dataset: 98.28% (84466/85945)

- Five random splits of the dataset results in the following average classification accuracies respectively: 99.91%, 99.86%, 99.92%, 99.94%, 99.83%.

Conclusion

- Proposed a method that can reliably discriminate between 15 classes of heartbeats based on wavelet features, ICA features, and RR interval features
- Validated our algorithm over the entire MIT-BIH Arrhythmias Database (none of the previous works seem to have used the whole database)
- Obtained an average classification accuracy of 99.91% (the highest in the literature, to the best of our knowledge)

Acknowledgements

Sincere thanks are due to

-- Professor Miguel Tavares Coimbra

-- Professor Vijayakumar Bhagavatula (Co-advisor at Carnegie Mellon)

-- All of the colleagues within the group for providing the most stimulating working environment and selfless help