ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 8, August 2015

Cardiac Abnormality Detection from ECG Using AHMM

Santhosh.J[1], Raji.N[2]

Asst. Professor, Department of Computer Science, Sree Narayana Guru College, Coimbatore, Tamil Nadu, India[1]

M.Phil. Scholar, Department of Computer Science, Sree Narayana Guru College, Coimbatore, Tamil Nadu, India [2]

ABSTRACT: The premature discovery of aberrant health conditions is essential to spot heart related problems and save life. Electrocardiogram which is abbreviated as ECG, ECG data is used to diagnose various heart conditions by generating patterns based on the heart beats. Analyzing the ECG data with the conventional morphological features is very tedious and tough. To analyze the health conditions, there are several devices and different format of data's are gathered. The data collection and monitoring from different devices is very complicated and time consuming. The proposed system begins with the HL7 medical communication standard, and integrates the data into XML file, and then this performs the cardiac abnormality finding process. Our proposed modal improves the accuracy in ECG diagnosis. This includes the detailed measurement of ECG waveform, which is used to diagnose wider range of health issues. This also includes the demographical information such as gender and age for improving the accuracy. Our proposed scheme comprises sub type classification with the use of two more attributes renewed with ontology. Finally the ECG data along with the cardiac conditions have been mapped in an internet browser.

KEYWORDS: Electrocardiogram, XML, cardiac disease, abnormality, HL7 Medical standard, Hidden Markov Model, Classification

I. INTRODUCTION

Electrocardiography simply abbreviated as ECG, which records the electrical activity of the heart and helps to diagnose cardiac abnormality. These electrodes detect the tiny electrical changes on the skin that arise from the heart muscle depolarizing during each heartbeat. The research of ECG signal is extensively used for analysis and diagnosis of heart diseases. It is used to compute the heart rate, heart beat regularity, the presence of any damage to the heart, and the causes of medicines or devices used to normalize the heart can be recognized with the help of ECG signal [1]. Using the ECG signals, the abnormality of heart rhythms has been identified.

To produce ECG output data and then to make suggestion about cardiac abnormalities, which represented from the output. Fig. 1.0 exemplifies the structure of ECG waves with intervals, voltage measures and time.

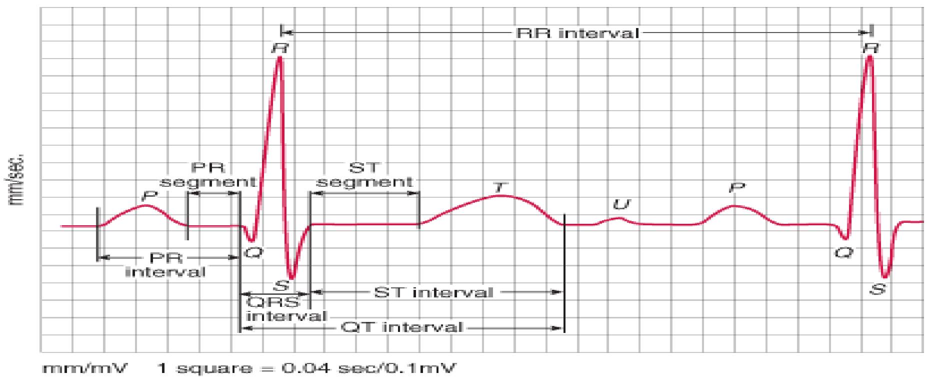

Fig1.0 ECG signal

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 8, August 2015

- P-wave: It is the first positive deflection on the ECG and it represents **atrial depolarization.**
- A Q wave is any negative deflection that *precedes* an R wave
- T-Wave: The T wave is the positive deflection after every QRS complex and it represents ventricular *repolarisation.*

Those ECG data's are stored in different formats and different devices. Due to its dissimilar format, data analysis became vague and time consuming. To overcome this issue, HL7 standard format has been introduced. It transfers the clinical and administrative data between software applications used by various ECG devices and stores the data into an XML format [2].

In order to diagnose diseases from the ECG data, numerous systems have been introduced and developed, those system assist the physicians as a supportive decision system against cardiac condition diagnosis.

The intention of our study is to develop a morphological model and subsequent software system to diagnosis and promote open exchange and presentation of ECG data. Using the demographical information along with the morphological classification, the system improves the accuracy of the abnormality detection using ECG data. The study also aims to summarize the diagnosed information to the physician with graphical representation of ECG signals.

II. BACKGROUND OR RELATED WORK

A. Image acquisition technique:

With the use of ECG data, abnormal cardiac conditions can be diagnosed. Even though the diagnosis is simple, but diagnose reports was wrong. Some authors [3] in the literature have proposed image acquisition techniques to get data from ECG image. Mean while this uses histogram validation for improving the quality of the image, this has been used to diagnose report automatically. The system [4] renovated with two different models. The first model is an ontological model and subsequent software system to promote open exchange and the second one is presentation of ECG data along with the ontology as an aid for the automatic diagnosis of common cardiac abnormalities. The common method has been used to exchange the visualized result of medical data.

B. Heart rhythm abnormality finding:

Hidden Markov models (HMM) have been used to inspect ECG waveforms to find abnormal heart rhythm [5]. The HMM approach combines structural and statistical knowledge of the ECG signal in a single parametric model. The Model parameters are estimated from training data using an iterative, maximum-likelihood re-estimation algorithm. However, the mechanisms of detecting cardiac abnormality from ECG waveforms by applying MLr-E, is successful and the system doesn't handle P-wave data's. Therefore, to identify all the cardiac abnormalities, the presented system in [6] requires hundreds of complex algorithms to be integrated under one computationally system. Maintaining and updating such a system for every new abnormality is intrinsically complex. This introduces a problem of finding a simple and fast solution toward heart disease recognition from compressed ECG that raises alert to the cardiac specialist as soon as a cardiac disease is recognized. In the previous work [7], only performs cardiac abnormality detection with essentially two clusters which are known as normal cluster and abnormal cluster.

C. ECG feature extraction:

Feature Extraction from ECG data [8] plays a major role in detecting most of the cardiac diseases and its risk. The structural information of the ECG signal holds useful information about the nature of cardiac diseases and its abnormalities. The corresponding details are difficult to analyze visually by the humans, thus computer assisted analysis and classification of cardiac diseases can help physicians to monitor cardiac health easily and accurately. Thus, computer-aided automatic diagnosis and classification of cardiac abnormalities is very helpful in health care, which is more helpful in emergency conditions.

Recently numerous feature extraction techniques have been established to determine the current state of cardiac activity through analysis of heart beats and twists found in ECG data.

Morphological or timing statistical-based features [9] which are usually extracted from ECG signal includes P curve, PQ/PR and QRS widths and QT intervals this also includes P and T amplitudes QRS height and ST level. These

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 8, August 2015

features include both constant and unstable characteristics of the ECG.A large number of techniques exist in the literature for the automatic detection and classification of different ECG signal morphologies.

The above Existing system defeated the difficulty of unfortunate diagnosis and accuracy over inputted ECG image sample. Due to the noisy and different formatted ECG data, ensemble and resulting with sharing become ineffective. This leads to false prediction of illness in the health care domain.

Our proposed system overcomes the hindrance of false prediction and accuracy of cardiac risk identification by validating the input ECG data using data mining techniques that validates the data from XML for further processing.

III. PROPOSED SYSTEM

The course of action in D-ECG for converting ECG waveform data to a common data file with the consideration of demographical features has the following phases.

- Data and Feature extraction and Preprocessing
- Classification and prediction

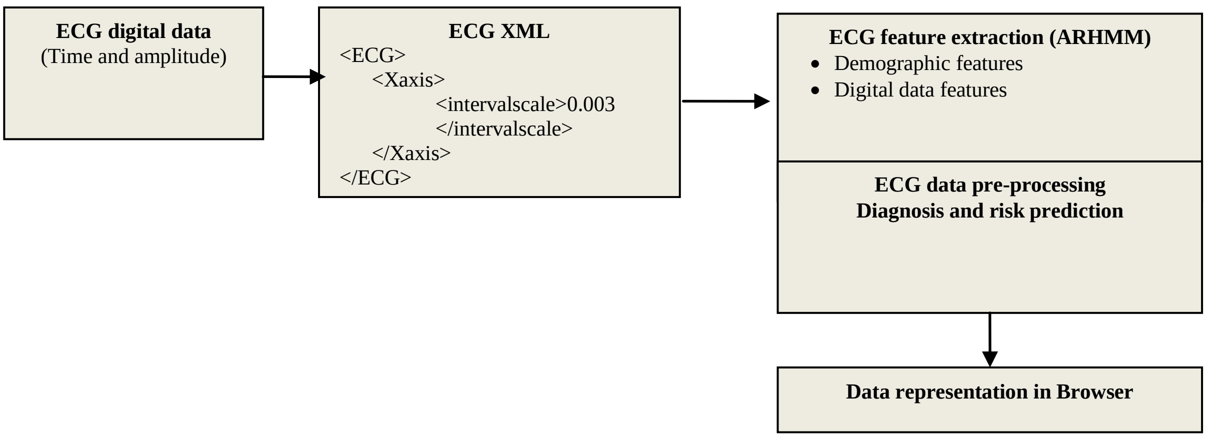

Fig: 2.0 Process flow of D-ECG framework

The above fig 2.0 represents the overall process performed in the HARS, where ECG digital dataset are converted into XML, and dater this performs the feature extraction and diagnosis processes.

The first phase of our proposed work involves the feature extraction process from ECG dataset. The system proposes a new feature extraction method with demographical feature finding. For effective feature extraction the system performs ARHMM model. The following figure 3.0 represents the sample ECG report, which has been taken for experiment. In the ECG report, there are demographical features specified with wave details, using the sample report, the feature has been extracted. The figure 3.0 shows the sample ECG report format, which includes age, gender etc., and the sample ECG digital report has temporal data as x-axis and amplitude as y-axis along with the start, end and peak values.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 8, August 2015

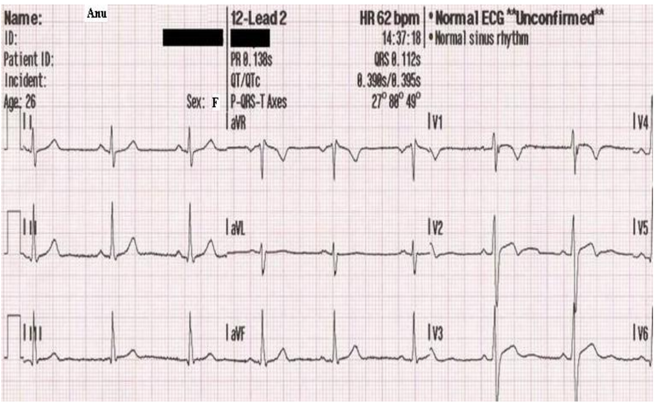

Fig 3.0 ECG sample report for feature extraction

The above ECG data has been converted into XML documents, where those XML documents will take into the preprocessing stage. This job completed with the use of the *Matlab tool*. The initial ECG digital data sample has displayed below (see fig 4.0).

Sample #	Time (second) *x*-axis	Amplitude (mV) *y*-axis
1	0.000	-0.270
2	0.003	-0.255
3	0.006	-0.245
:	:	:
3333	20	-0.345

Fig 4.0 ECG digital report extracted from the ECG device

The XML Schema consists of structures for the components of the waveform drawn from the *x* and *y* coordinates of the graphical waveform which showed in fig 5.0. The *x*-axis represents the time interval in seconds while the *y*-axis represents amplitude in millivolts.

```
  <ECG  xmlns:xsi="http://www.w3.org/2001/XMLSchema
instance">
    <Lead>MLII</Lead>
    <GraphInfo>
      <XAxis>
        <IntervalScale>0.003</IntervalScale>
        <StartTime>10</StartTime>
        <EndTime>29.997</EndTime>
        <TimeUnit>Second</TimeUnit>
      </XAxis>
      <YAxis>
        <DigitalData>-0.27,-0.255,-0.245,-0.24,
  -0.24,-0.245,-0.25,-0.235,-0.225,-0.235,-0.23,
  -0.245,-0.245,-0.24,-0.23,-0.225,[..]
        </DigitalData>
        <AmpUnit>mV</AmpUnit>
```

Fig 5.0 ECG digital report converted to XML

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 8, August 2015

A. AHMM:

AutoRegressive Hidden Markov Model (ARHMM) is a combination of autoregressive time series and *Hidden Markov Chains (HMC).* The Interpretation is generated by a few autoregressive time series while the switches between each autoregressive time series are controlled by a HMC. A time series may sometimes consist of observations generated by different mechanisms at different times. When this happens, the time series observations would act like switching back and forward between couple of distinct states. When changing into a different state, time series may have a significant change in their means or in their frequencies or breadths of their fluctuations. The *Autoregressive Hidden Markov model* (*ARHMM*) is often being used to deal with this kind of time series. As indicated by the name, an ARHMM is the combination of an autoregressive time series model and a hidden Markov model. The autoregressive structure admits the existence of dependency amongst time series observations while the hidden Markov chain could capture the probability characteristics of the transitions amongst the underlying states. Actually, ARHMM is also referred as *time series with change in regime* (or *states*) by the econometricians.

To be more specific, let us see an example of ARHMM. As usual, Y = *{Y1, Y2... YT}* denote the observation sequence. Each Y_t is a observation vector with k component $Y_t = \{y1, y2\ldots y_k\}'$.

X = $\{X_1, X2\ldots X_T\}$ is a hidden state sequence with *N* possible states. X is assumed to be a Markov chain with transition matrix *A* = [*aij]* and initial distribution vector $\pi = \pi_i$.

But it should be mentioned that the ARHMM with demographical for distinct state *Xt* could also be developed with more complexity. In such cases, the error term *"t* will usually be replaced by *"Xt* which depended on the value of current state *Xt*. E-M algorithm or segmental K-mean algorithms could only lead to a local maximum of the HMM likelihood function. For ARHMM, this is also true.

To get the parameter estimates with a global maximum likelihood, a grid search approach might be used. In grid search approach, the parameter space is seen as a grid with many small cells and all the vertices are used as the initial values of the parameters. Because the parameter space is so big in the case of ARHMM, the grid search method requires considerable computational power which is intractable for practical purposes.

Another notable feature of ARHMM estimation is the high autoregressive coefficients. This is exactly the reason why ARHMM are superior to conventional HMM in this application. Conventional HMM assumes there are independency relations between the observations. But this is rarely the case for time series observations. As in this system, ECG data are collected on a day-by-day bases and apparently the independency assumption is inappropriate. Comparatively, the autoregressive structure contributes the superiority of ARHMM in a way it prevents the frequent fluctuations of state path. Conventional HMM are very sensitive to the numerical swings of the current ECG and hence mistakes several fluctuations of ECG as the switches of states. While for the same data, ARHMM state path are more stable and close to reality. Using the ARHMM, the ECG data can be diagnosed and reported as graphical/text formats on web browser.

The diagnosis task involved validating accuracy of the system in correctly detecting abnormal cardiac conditions as well as normal ECGs.

IV. EXPERIMENTS AND RESULTS

A. Dataset:

For experiment, our system used dataset from physionet domain and the dataset named as "MIT-BIH Arrhythmia Database" which contains 234 records. The followings are the sample digitized data for each option (see fig 6.0).

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 8, August 2015

Record	N							V	F	O	N		E	P	F	O	Q
	.	L	R	A	a	J	S	V	F	!	e	j	E	P	f	p	Q
100	2239	-	-	33	-	-	-	1	-	-	-	-	-	-	-	-	-
101	1860	-	-	3	-	-	-	-	-	-	-	-	-	-	-	-	2

Symbol	*Meaning*
N	Normal beat
L	Left bundle branch block beat
R	Right bundle branch block beat
A	Atrial premature beat
a	Aberrated atrial premature beat
J	Nodal (junctional) premature beat
S	Supraventricular premature beat
V	Premature ventricular contraction
F	Fusion of ventricular and normal beat
[	Start of ventricular flutter/fibrillation
!	Ventricular flutter wave
]	End of ventricular flutter/fibrillation
e	Atrial escape beat
j	Nodal (junctional) escape beat
E	Ventricular escape beat
/	Paced beat
f	Fusion of paced and normal beat
x	Non-conducted P-wave (blocked APB)
Q	Unclassifiable beat
|	Isolated QRS-like artifact

Fig 6.0 Beat types and its symbol representation

From the above fig 6.0, the beat type has been displayed and that will be extracted as an XML file and applied into the diagnosis process. Then the system finds the abnormality based on the diagnosis rules listed in the HL7 medical device communication standard. The results are converted to output in a browser. By this feature the users can able to see output with abnormality detection in a textual and graphical form without the use of any software.

B. Implementation:

The D-ECG framework has been developed in C#.net with feature extraction and diagnosis, later this has been presented in a web browser using ASP.net framework.

Fig: 7.0 the result from D-ECG framework

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 8, August 2015

The above fig 7.0 shows the output of the proposed framework, where the graphical and textual data are summarized on the web page. From the above figure, the chart is represented for risk assessment based on the ECG result. From the above experiment, the abnormality has been found in term of percentage. The risk has been calculated for every time interval and it also displays the overall finding with demographical features.

VI.CONCLUSION

The developed D-ECG framework provides a graphical construction for the demonstration and sharing framework of ECG data so that it can be made readily accessible for presentation or screening on a large amount of computing environments. Using the ECG digital data as input the user can diagnosis abnormalities of cardiac conditions. These resources are based on the HL7 standard and AHMM model. The improved ontology structure used for data representation and thus provides an easy presentable text format for the user after successful risk analysis via web browser.

REFERENCES

1. de Chazal, Philip, and Richard B. Reilly. "Automatic classification of ECG beats using waveform shape and heart beat interval features." *Acoustics, Speech, and Signal Processing, 2003. Proceedings.(ICASSP'03). 2003 IEEE International Conference on*. Vol. 2. IEEE, 2003..
2. Lu, Xudong, Huilong Duan, and Huiying Zheng. "XML-ECG: An XML-based ECG presentation for data exchanging." *Bioinformatics and Biomedical Engineering, 2007. ICBBE 2007. The 1st International Conference on*. IEEE, 2007.
3. Schindewolf, Thomas, et al. "Evaluation of different image acquisition techniques for a computer vision system in the diagnosis of malignant melanoma." *Journal of the American Academy of Dermatology* 31.1 (1994): 33-41.
4. Centers for Disease Control and Prevention. "National diabetes statistics report: estimates of diabetes and its burden in the United States, 2014." Atlanta, ga: US Department of health and human services (2014).
5. Andreão, Rodrigo Varejão, and Jérôme Boudy. "A comparison of wavelet transforms through an HMM based ECG segmentation and classification system." *The Proceeding of IASTED Conference on Biomedical Engineering Innsbruck, Austria February*. 2006..
6. Andreão, Rodrigo Varejão, Bernadette Dorizzi, and Jérôme Boudy. "ECG signal analysis through hidden Markov models." *Biomedical Engineering, IEEE Transactions on* 53.8 (2006): 1541-1549..
7. Sufi, Fahim, and Ibrahim Khalil. "Diagnosis of cardiovascular abnormalities from compressed ECG: a data mining-based approach." *Information Technology in Biomedicine, IEEE Transactions on* 15.1 (2011): 33-39..
8. Alexakis, C., et al. "Feature extraction and classification of electrocardiogram (ECG) signals related to hypoglycaemia." *Computers in Cardiology, 2003*. IEEE, 2003.
9. De Chazal, Philip, Maria O. Dwyer, and Richard B. Reilly. "Automatic classification of heartbeats using ECG morphology and heartbeat interval features." *Biomedical Engineering, IEEE Transactions on* 51.7 (2004): 1196-1206.