27M-05

個人差を考慮した高齢化社会における交通事故防止システムの構築

Automated Prediction of Sudden Cardiac Death Risk Using Kolmogorov complexity and Recurrence Quantification Analysis Features Extracted from HRV Signals

藤田ハミド（岩手県立大学、教授）、羽倉淳（岩手県立大学、准教授）、
R. Udyavara Acharya (Ngee Ann Polytech, Singapore)、
Gajo Petrovic(岩手県立大学,大学院後期課程 D2)、田中舘涼（岩手県立大学, H27 年度学部卒業）

＜要旨＞
本研究では、本研究は、体調、精神状態、性格といったドライバー個人の内面に大きく関わった交通事故要因と、ドライバーの年齢、性別、自動車の運行目的といったドライバーの属性、事故現場の天候、事故発生時刻といった環境要因といった客観的に判別可能な交通事故要因を考慮した交通事故のリスク予測を実現する手法の構築を試みるものである。本リスク予測手法は、刻々と変化する状況をリアルタイムで検出し、数十秒後の事故リスクを計算可能にすることを目標とする。本技術が確立すれば、個々のドライバーの状況や特性に応じた危険予測が可能になり、警告システム等に応用することで高齢者を含む交通事故防止が予測できることが期待できる。ここでは、脳波測定と分析に基づく、心臓発作・癲癇発作の予兆を検出手法の構築と、それを利用した交通事故防止手法の構築を行う。さらに、ドライバーの心理状態状態（特に、焦り）と脳波の関係性について調査し、心理状態の検出方法の構築を行う。

1　研究の概要

The driver condition is important parameters to examine that provide safe environment on to the drivers him/her self and surroundings. In this research project we have examined fatal parameters in predicting heart sickness and warn the driver on such related situation before it trigging a threat to the driver and also to the surroundings. SCD(Sudden Cardiac Death：心臓突然死) is an unexpected sudden death due to cardiovascular problems and with or without history of cardiac diseases. Generally SCD occur within one hour post symptoms onset, even though the person has no previous fatal cardiac condition. SCD due to cardiac arrest is the most familiar cause of death worldwide, accounting to more than 50% of all deaths from cardiovascular diseases. SCD is a vital challenge for the clinicians as it occurs in individuals without any previously known cardiac disease. Ventricular fibrillation (VF) is the mechanism underlying most SCD episodes. Survival rate declines approximately by 10% per minute for patients after the VF. VF triggered by VT is a standard mechanism of cardiac arrest leading to SCD. Other triggers of SCD include Coronary Artery Disease (CAD), transient ischemia, myocardial infarction (MI) and heart failure. Multiple strategies have evolved to predict and prevent SCD. One strategy among the classes of risk predictors is clinical marker, such as electrocardiogram (ECG) measures. Experts worldwide are working on these cardiac problems with an intension to predict the SCD before its onset using ECG signals].

Analysis of nonlinear and complex ECG or HRV signals using linear methods might not yield complete information necessary for predicting the SCD onset. In order to unearth the hidden information from these bio-signals nonlinear methods are used. Patients at high risk of SCD exhibit nonlinear heart rate (HR) dynamics like rapid spectral alterations and sustained low frequency HR oscillations. Summary of

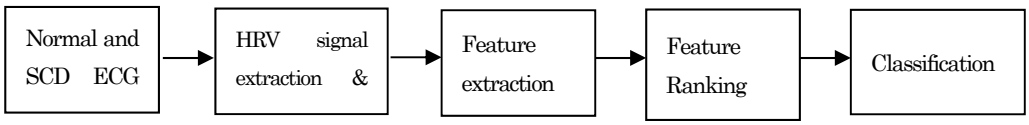

Figure 1 Block diagram of the proposed method.

research work conducted in identifying the SCD before its onset using HRV signals is tabulated in Table 3.

Few researchers examined the ability of different nonlinear methods to find distinctive pattern in HRV signals which cannot be found by conventional linear methods. We have reported[2~4] the characteristics of ECG signals by means of a non-linear short-term fractal scaling exponent as a useful independent risk predicting factor for SCD in patients who had episode of acute myocardial infarction. In addition to the time and frequency domain parameters, Shen et al. (2007)[1] also used nonlinear method to study personal cardiac homecare system for SCD event detection and prediction using ECG signals. They reported an accuracy of 87.5% using wavelet analysis for SCD detection.

Various research studies are conducted on the prediction of SCD onset using time, frequency and nonlinear methods using HRV signals [Table 3]. The current work proposes an efficient algorithm using nonlinear RQA parameters are extracted from the recurrence plots (RPs) of normal and SCD prone HRV signals and Kolmogorov complexity parameters. Features with significant information are selected using *t*-test and ranked using t-value. These ranked features are fed to classifiers such as k-Nearest Neighbors (k-NN), Decision tree (DT), Support vector machine (SVM) and Probabilistic Neural Network (PNN). The entire process is applied to one minute intervals up to four minutes of HRV signal segments separately. The block diagram of the proposed method is shown in Figure 1.

This paper is organized as follows. Section II discusses the dataset used and different methods. Section III presents the experimental results and Section IV provides discussion on results. The concluding remarks are presented in Section V.

I. Materials and Methods

The block diagram of the proposed method is shown in Figure 1.

A. Dataset Used

The ECG signals (SCD and normal) necessary for the experiment were obtained from MIT-BIH SCD Holter open access database and the Normal Sinus Rhythm database[2]. The above mentioned open source database has 41 ECG signals acquired from 23 SCD patients and 18 normal subjects. The age of SCD patients range from 18-89 years and control subjects from 20-50 years. The sampling rate of SCD and normal ECG signals were 256 Hz and 128Hz respectively. In this work, to maintain uniformity between SCD and normal groups, the SCD ECG signals were sampled at 128 Hz. Only 20 SCD ECG signals were used for analysis and the remaining three ECG signals were not used, as they did not have the VF episodes.

B. Pre-Processing

In this paper, the ECG signals four minutes before the SCD occurrence (onset) were acquired from the 24 hours of ECG recordings of SCD patients. From the ECG recordings, HRV signals of 10 minutes were

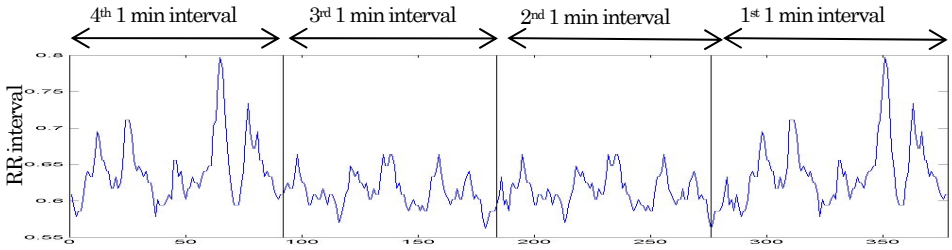

Figure 2: The four one minute intervals of normal HRV signal.

[1] [T.W. Shen, H.P. Shen, C. Lin, Y. Ou. Detection and prediction of Sudden Cardiac Death (SCD) for personal healthcare. 29th Annual International Conference of the IEEE, Buenos Aires, 22-26 August, 2575-2578, 2007.]

[2] A.L. Goldberger, L.A. Amaral, L. Glass, J.M. Hausdorff, P.C. Ivanov et al., PhysioBank, PhysioToolKit, and PhysioNet: components of a new research resource for complex physiologic signals. Circulation, 101, e215-220, 2000.

extracted using Pan-Tompkins algorithm[3] and only the four minutes duration HRV is used for further analysis. The obtained HRV signals of four minutes duration were further divided into four one minute intervals. The details of this process used for ECG signal is discussed in our previous paper [3]. The normal HRV signal of four one minutes' intervals is shown in Figure 2.

C. Recurrence Plot (RP)

Recurrence plots[4] (RPs) are useful for evaluating the geometry of system by utilizing non-linear behavior in non-stationary time-series. It is useful to study the behavior of physiological data. RP plots reveal the distance relationships between points on a dynamical system. This is a graphical tool helps to identify the shift in the dynamical system and invisible periodicities which are unnoticeable in the time evolution. A brief explanation about the construction of RPs is described below.

Consider a_i as the i^{th} dot on the range in m-dimensional space. The RP is an array of points in $y \times y$ matrix, in which a point is placed at (i,j) if a_j is adequately near a_i. To obtain the RP from time-series a_y, m-dimensional orbit of a_i is constructed where an embedding dimension (m) is chosen by method of delays. Further, radius r is selected such that the ball of radius r is centered at a_i in $\mathbb{R}^m$ holds adequate number of other dots a_j. In the end, a point is marked for every (i,j) where a_j falls in the projectile of radius r and center a_i. The plot acquired is the RP and in this work, the delay = 1 and embedding dimension = 10 are used for generating the RPs]. RP is calculated for pre-processed HRV signals.

D. Feature extraction

The feature extraction process is significantly crucial step in developing an automated algorithm. The extraction of features is challenging for HRV signals as it is complex, and non-stationery in nature. In order to overcome this, a new-Recurrence Quantification Analysis (RQA) [17] and Kolmogorov complexity[5] are used in this paper. RQA is used to extract significant parameters from the $RP_{i,j}$. The parameters obtained from RP describe the underlying dynamic system. RQA quantifies and provides valuable information about the nonlinearity and complexity of the non-stationary data. RQA parameters are extracted from the Recurrence Plot (RP) depicts the nonlinear behavior of the HRV signals.

The entropy is a nonlinear feature that reflects the degree of chaos within a system. It is often used to analyze the bio-signals for detecting the presence of abnormalities. Therefore, in addition to RQA parameters, Kolmogorov complexity is also extracted from the HRV signals to improve the classification accuracy. Kolmogorov complexity [19, 20] is used to quantify the complexity and irregularity of HRV signals for normal and SCD subjects.

RQA parameters (Recurrence Rate (RR), Determinant (DET), Mean diagonal line length (meanLen), entropy (Ent), transitivity (Trans), Recurrence time entropy (RTE), Laminarity (LAM), Longest vertical line (Vmax), Longest diagonal line (Lmax) and Recurrence Times (T1, T2)) are extracted from RP of HRV signals. These RQA parameters together with Kolmogorov complexity feature are subjected to ranking followed by classification.

E. Feature Ranking

In order to select the features with significant information the feature ranking method is used. The feature ranking methods rank the extracted features according to their statistical significance. In this work, the features are ranked based on their t- value [23].

F. Classification

In order to classify normal and SCD subjects using HRV signals k-Nearest Neighbour (KNN), Decision Tree (DT), Support Vector Machine (SVM) and Probabilistic Neural Network (PNN) classifiers are used. In this work, the ten times ten-fold cross-validation method is used to test the performance of all classifiers.

II. Results

Figure 3 and 4 shows typical RPs for (a) first one

[3] J. Pan, W.J. Tompkins. A real time QRS detection algorithm. IEEE Trans on Biomed Eng, 3, 230-236, 1985.

[4] J.P. Eckmann, S.O. Kamphorst, D. Ruelle D. Recurrence plots of dynamical systems. Europhys Lett, 4, 973-977, 1987.

[5] A. Kolmogorov. On tables of random numbers. Sankhya: The Indian journal of statistics, Series A, 25, 369-376, 1963.

minute, (b) second one minute, (c) third one minute, and (d) fourth one minute of normal and SCD HRV segments respectively. Figure 3 (a) - (d) RP of normal HRV segments exhibits the continuous variation, and the dots are scattered throughout the plot; thus indicating the variation in normal HRV signals.

Figure RP of SCD cases has diagonal line and small squares indicating less variation in the signals which in turn indicates more uniformity in the HRV signal. It can be seen from the figure that HRV exhibits periodicity in some places and looks more rhythmic; the dots are concentrated in few places in the plots. In addition, Figure 4 shows continuous periodicity (regularity) in the plot from (a) - (d); thus indicating the high rhythmicity in the signal.

Results (mean and SD) of all the features extracted and ranked four minutes before SCD onset are shown in Table 1. It can be seen from this table that the first three features are clinically significant ($p < 0.05$) for four minutes before SCD onset. The mean values of Kolmogorov complexity, Lmax, Ent, RTE, meanLen, Vmax, RR are more for SCD as compared to normal HRV signals. The HRV signal varies more rapidly but rhythmically for SCD as compared to normal subjects.

Table 2 shows the classification results of normal and SCD HRV signals using different classifiers for four minute before SCD onset. The k-NN and PNN classifiers performed better than rest of the classifiers for four minutes before the SCD. PNN classifier yielded the highest average classification accuracy of 86.8%, sensitivity of 80% and specificity of 94.4%. The k-NN classifier yielded 86.8%, sensitivity of 85% and specificity of 88.8%.

Table 1: Range of features extracted from normal and four minute prior SCD HRV signals.

Features (RQA and entropies)	Normal		SCD		*p*-value	*t*-value
	Mean	SD	Mean	SD		
K-Complex	3.270	0.472	4.184	0.646	< 0.0001	4.925
Trans	0.839	0.085	0.754	0.109	0.0120	2.645
Ent	0.265	0.189	0.754	0.109	0.0208	2.417
Lmax	4.222	2.556	6.75	5.045	0.0635	1.913
RTE	0.606	0.180	0.695	0.114	0.0751	1.832
meanLen	2.256	0.367	2.532	0.541	0.0773	1.818
Vmax	4.055	3.152	6.65	5.412	0.0837	1.778
RR	0.086	0.048	0.153	0.151	0.0837	1.778
DET	0.296	0.166	0.394	0.236	0.1552	1.451
T2	10.52	3.521	9.131	2.864	0.1861	1.347
T1	8.233	3.438	6.675	3.837	0.1977	1.312
LAM	0.361	0.162	0.447	0.263	0.2417	1.190

Table 2: Results of automated classification for four minutes before SCD HRV signals.

Classifiers	No of Features	Accuracy	Sensitivity	Specificity
DT	2	76.3%	75.0%	77.7%
k-NN	3	86.8%	80.0%	94.4%
PNN-0.03	3	86.8%	85.0%	88.8%
SVM RBF -0.3	3	84.2%	85.0%	83.3%
SVM Poly 1	2	84.2%	80.0%	88.8%
SVM Poly 2	2	78.9%	75.0%	83.3%
SVM Poly 3	2	78.9%	75.0%	83.3%

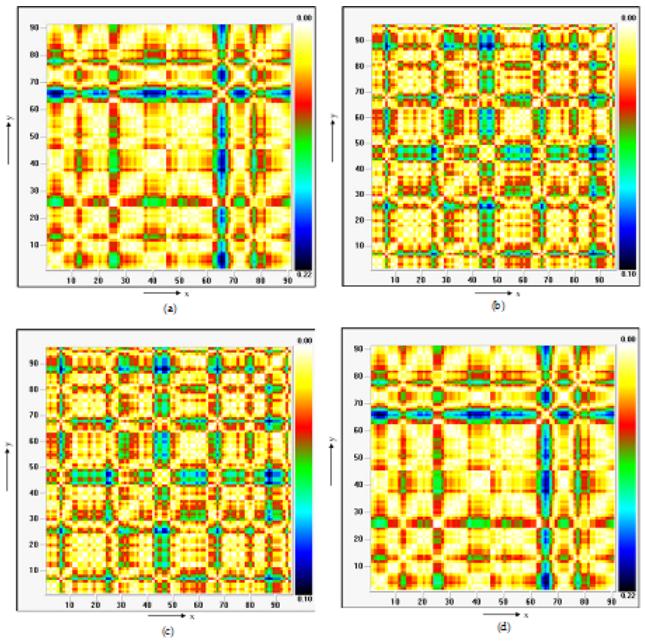

Figure 3 RPs for (a) 1st one minute, (b) 2nd one minute, (c) 3rd one minute, and (d) 4th one minute of normal HRV segments.

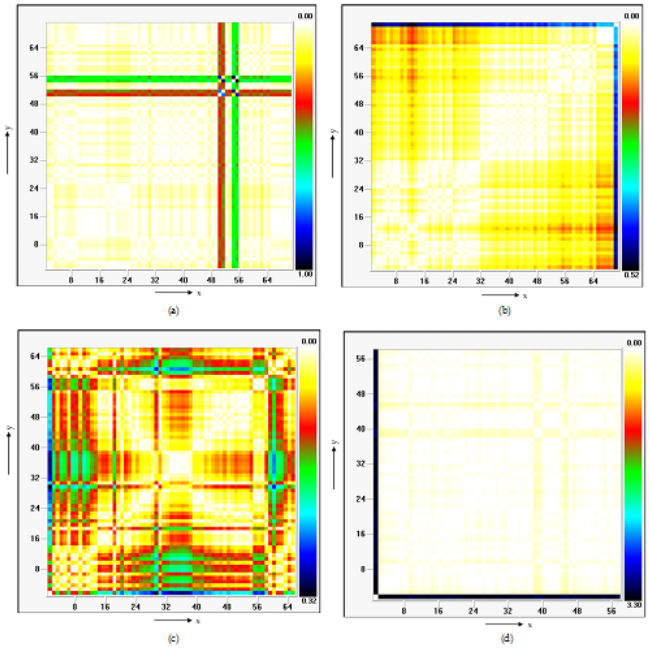

Figure 4 RPs for HRV signals of (a) one minute, (b) two minutes, (c) three minutes, and (d) four minutes before SCD onset.

III. Discussion

Many researchers have developed algorithm for SCD prediction before its onset using HRV signals. The unique feature of our work is the improved efficiency for the prediction of SCD risk four minutes before its onset using novel nonlinear features namely Kolmogorov complexity and RQA features extracted from the RPs.

Recently, we have developed a novel SCD index for the prediction of SCD four minutes earlier by using a single number derived from nonlinear features extracted from DWT coefficients obtained from ECG signals [3]. Our study used nonlinear features namely fractal dimension, Hurst's exponent, detrended fluctuation analysis, correlation dimension, approximate entropy, sample entropy extracted from DWT coefficients. These features combined with SVM classifier are able to differentiate normal and SCD subjects with an accuracy of 92.11% four minutes before SCD onset. The highlight of our previous work is the formulation of a novel Sudden Cardiac Death Index (SCDI), based on the combination of selected features to enable the clinician to easily and effectively determine the SCD risk of a patient by using a single number.

It can be seen from Table 3 that, the efficiency of SCD risk prediction is improved significantly using RQA and Kolmogorov complexity features extracted from HRV signals. This paper proposes unique recurrence plots for 1min, 2min, 3min and 4 min before SCD onset. This algorithm is able to effectively predict the SCD even four minutes prior to the onset with an accuracy of 86.84% using only *three* features obtained from HRV signals. This result can be used in the efficient prediction of SCD occurrence risk four minutes prior to its onset using HRV signals. This system will help to provide immediate medical treatment and save life as it is able to predict SCD accurately 4 minutes early.

Our method obtained the highest accuracy using only 3 features with HRV signals. Hence our method computationally less complex and fast, and can be installed in vehicle for elderly SCD prediction. The system is robust as we have performed ten-fold cross validation.

The novelty of this work is the SCD prediction four minutes before its onset using just *three* HRV features extracted from HRV signals and proposed unique RPs for each class. In future, authors will be exploring the possibility of predicting SCD still earlier (earlier than 4 minutes) using more subjects. In addition, the integration of ECG and HRV features for the prediction of SCD risks.

IV. Conclusion

In spite of the advances in signal processing and machine learning in modern technology, accurate identification of SCD is the biggest challenge in cardiology. In this paper, an efficient algorithm is proposed for an automated prediction of SCD four minutes prior to its onset using HRV signals. We have also proposed unique recurrence plots for normal and SCD classes to differentiate them visually. Our algorithm has achieved the highest performance in the prediction of SCD compared to all the reported previous works (Table 3). The proposed method can be tested using huge diverse database and employed in real world scenario.

5 論文・学会発表等の実績

研究成果を、タイヤ A ジャーナル（トムソン・ロイターの Science Citation Index (SCI) の高インパクトファクタークラスジャーナル）で公開した[2] [3] [4] [5]

[1] U. Rajendra Acharya; Hamido Fujita, et.al "Automated Prediction of Sudden Cardiac Death Risk Using Kolmogorov Complexity and Recurrence Quantification Analysis Features Extracted from HRV Signals" Systems, Man, and Cybernetics (SMC), pp. 1110-1115, 2015 IEEE International Conference on, DOI:10.1109/SMC.2015.199

[2] Hamido Fujita, U Rajendra Acharya, et.al "Sudden cardiac death (SCD) prediction based on nonlinear heart rate variability features and SCD index", Applied Soft Computing Journal (Elsevier), Volume 43, June 2016, Pages 510-519.
http://www.sciencedirect.com/science/article/pii/S1568494616301016

[3] U. Rajendra Acharya, Hamido Fujita, et.al "An integrated index for detection of Sudden Cardiac Death using Discrete Wavelet Transform and nonlinear features, Knowledge-Based Systems
Volume 83, July 2015, Pages 149-158,
http://www.sciencedirect.com/science/article/pii/S0950705115001094

[4] YING-TSANG LO, HAMIDO FUJITA, et.al "PREDICTION OF CORONARY ARTERY DISEASE BASED ON ENSEMBLE LEARNING APPROACHES AND CO-EXPRESSED OBSERVATIONS", Journal of Mechanics in Medicine and Biology, Vol. 16, No. 1 (2016) 1640010 (10 pages)
doi:10.1142/S0219519416400108
http://www.worldscientific.com/doi/abs/10.1142/S0219519416400108

[5] U. Rajendra Acharya, Hamido Fujita, et.al " Decision support system for fatty liver disease using GIST descriptors extracted from ultrasound images," Information Fusion, Volume 29, May 2016, Pages 32-39,
http://www.sciencedirect.com/science/article/pii/S1566253515000895

[7] #SMC2016, IEEE SMC 2016, Budapest, Hungary, October 9, 2016,
http://smc2016.org/sites/default/files/program/Tentative%20program%20_%203%20_%20REV.pdf

The below to accepted papers are to be presented in IEEE SMC conference in Budapest, October 2016. These could modify the results of [1]. However, these are supported by Kakenhi-ippan(C).

(1) #1142 Automated Characterization of Arrhythmias Using Nonlinear Features from Tachycardia ECG Beats

(2) #1144 Automated Diagnosis of Coronary Artery Disease using Nonlinear Features Extracted from ECG Signals