Development of Stroke Detection Method by Heart Rate Variability Analysis and Support Vector Machine

Keisuke Kamata[*], Koichi Fujiwara[*], Tomonobu Kodama[*], Manabu Kano[*],
Toshitaka Yamakawa[‡], Norikata Kobayashi[§], and Fuminori Shimizu[§]
[*] Kyoto University, Kyoto, Japan, E-mail: fujiwara.koichi@i.kyoto-u.ac.jp
[†] Kumamoto University, Kumamoto, Japan, E-mail: yamakawa@cs.kumamoto-u.ac.jp
[‡] Shimizu Hospital, Kyoto, Japan, E-mail: kobayashi@shimizu-hospital.or.jp

***Abstract*—It is important to start stroke treatment as early as possible for patient prognosis. In particular, thrombolysis with the tissue plasminogen activator (tPA) that can dissolve blood clots is effective only when it is given within 4.5 hours from the symptom onset. Since it is sometimes difficult for patients to recognize their symptoms, an early stroke detection system is needed. It is possible that a stroke can be detected by monitoring heart rate variability (HRV) because a stroke affects the autonomic nervous system. In the present work, a stroke detection method was proposed by integrating HRV analysis and support vector machine (SVM). The sensitivity and the specificity of the proposed method were 100% and 80%, respectively. The possibility of realizing an HRV-based stroke detection system was shown.**

I. Introduction

Cerebrovascular disease (CVD) is a group of brain dysfunctions related to disease of the blood vessels supplying the brain, including cerebral infarction, cerebral hemorrhage, and subarachnoid hemorrhage. These diseases are better known as stroke. Strokes result in part of the brain not functioning properly, and are a leading cause of disability and the fourth leading cause of death in Japan [1]. The number of stroke patients is increasing with the aging of society because its most significant risk factor is advanced age. 95% of strokes occur in persons age 45 and older, and two-thirds of strokes occur in those over the age of 65 [2].

In order to prevent stroke sequelae, early treatment is required. For example, thrombolysis with the tissue plasminogen activator (tPA) that can dissolve blood clots can be applied to acute stroke, and the greater its benefit is the earlier it is given [3]. However, it is effective within 4.5 hours from the symptom onset. As a result, tPA was given for about 2.4% of patients in US because most patients could not go to a hospital within 4.5 hours [4].

Although patients must get to a hospital as early as possible, it is sometimes difficult for patients to recognize their own stroke because its main symptom is consciousness disorder. In addition, stroke often occurs during sleep. Therefore, a stroke detection system should be developed so that patients with stroke can have the opportunity for early treatment.

The R-R interval (RRI) fluctuation in ECG, called heart rate variability (HRV), is a well-known phenomenon which reflects the autonomic nervous function, and many HRV features have been proposed for HRV analysis [5]. Since a stroke damages the autonomic nervous function and lessens HRV [6], [13], there is the possibility that a stroke can be detected through monitoring HRV.

Although the Holter monitor has been used to measure long-term RRI, its home-use is difficult because the Holter monitor is expensive and requires operation skills. Yamakawa et al. developed a wearable RRI sensor, which can measure RRI without any special skills and be manufactured for less than 100 US dollars [7]. If an HRV-based stroke detection algorithm could be implemented in such a device, a wearable stroke detection system would be realized.

The present work proposes a new HRV-based stroke detection method by integrating HRV analysis and support vector machine (SVM), which is a well-known pattern recognition method. An application result of the proposed method to clinical data is reported.

II. Heart Rate Variability Analysis

Since HRV reflects autonomic nervous activities, HRV analysis has been used for monitoring stress, drowsiness, and cardiovascular disease [5]. In this section, the HRV features used in this work are explained briefly.

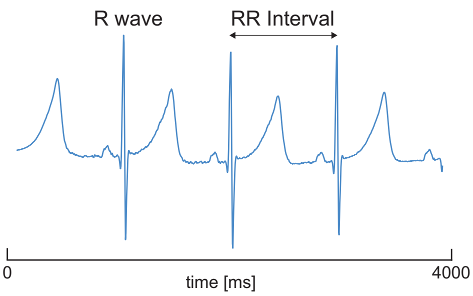

Fig. 1. An example of a typical ECG

A. RR Interval

A typical ECG trace of the cardiac cycle (standard lead II) consists of several peaks as shown in Fig. 1, and the highest peak is called the R wave. The RR interval (RRI) [ms] is defined as the interval between an R wave and the next R wave.

A part of the raw RRI data collected from a healthy person is shown in Fig. 2 (a). Since the raw RRI data are not sampled at equal intervals, frequency domain features cannot be extracted. Thus the raw RRI data is interpolated by using spline and resampled at equal intervals. Figure 2 (b) shows the resampled RRI data whose sampling interval is one second.

B. Time Domain Features

The following time domain features can be calculated from the original RRI data [5].

- **meanNN**: Mean of RRI.
- **SDNN**: Standard deviation of RRI.
- **RMSSD**: Root mean square of difference of adjacent RRI.
- **Total Power (TP)**: Variance of RRI.
- **NN50**: The number of pairs of adjacent RRI whose difference is more than 50 ms within a given length of measurement time.

C. Frequency Domain Features

The following frequency domain features can be obtained from the power spectrum density (PSD) of the resampled RRI data, and the PSD can be calculated by using Fourier analysis or an autoregressive (AR) model [5].

- **LF**: Power of the low frequency band (0.04Hz - 0.15Hz) in a PSD. LF reflects both the sympathetic and parasympathetic nervous systems activity.
- **HF**: Power of the high frequency band (0.15Hz - 0.4Hz) in a PSD. HF reflects the parasympathetic nervous system activity.
- **LF/HF**: Ratio of LF to HF. LF/HF expresses the balance of the sympathetic nervous system activity with the parasympathetic nervous system activity.
- **LFnu**: Normalized LF defined as LF/TP [9].
- **HFnu**: Normalized HF defined as HF/TP [9].

Figure 2 (c) shows the PSD and its LF/HF of the resampled RRI data shown in Fig. 2 (b). According to the HRV analysis guideline, the RRI data should be measured for at least three minutes for precise frequency analysis [5].

III. Stroke Detection Algorithm

The present work proposes a stroke detection algorithm that discriminates stroke/healthy (S/H) by using HRV data. In the proposed algorithm, an S/H discriminant model is constructed from HRV data of both stroke patients and healthy persons by using SVM.

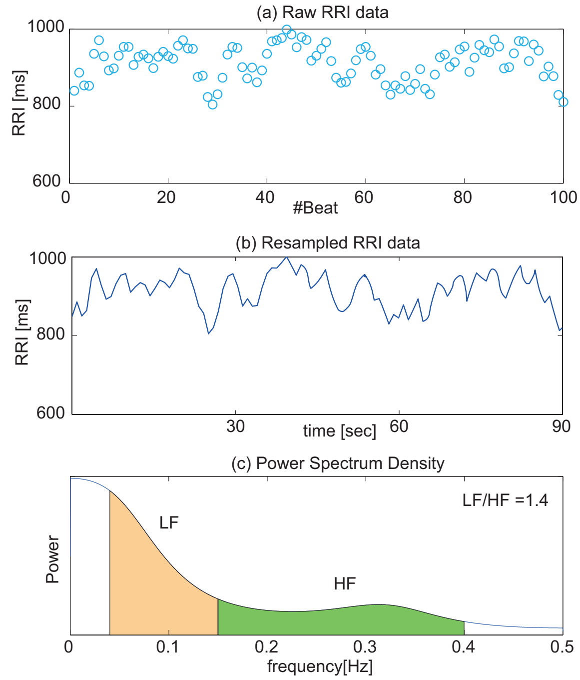

Fig. 2. An example of RRI data analysis: (a) the raw RRI data, (b) the resampled RRI data and (c) the PSD and its LF/HF

A. Support Vector Machine

SVM is a nonlinear classification technique, that was developed for classifying data into two classes [10]. Given the modeling data that consist of two classes, SVM constructs an optimal separation hyperplane which has the maximum margin. The margin is defined as the distance between the separation hyperplane and its closest samples. If the modeling data are not linearly separable, the original data are mapped into a high dimensional space by a kernel function and a separation hyperplane is constructed in the high dimensional space. This is equivalent to nonlinear model construction in the original space.

SVM has been widely used for many applications such as spam mail filtering and bioinformatics [11], [12].

B. Modeling

The S/H discriminant model construction procedure is shown in Algorithm 1. First, the RRI data are collected from both patients with stroke and healthy persons, and the collected data are converted to HRV data. Next, the S/H discriminant model is constructed from the converted HRV data by using SVM.

In the stroke detection procedure described in Algorithm 2, the RRI data of subjects are collected, and they are converted to HRV data. Subject condition at every RRI collection is diagnosed as stroke or healthy from the HRV data by using the constructed S/H discriminant model. In order to determine whether the subject has stroke or not, the stroke ratio A is used: $A = 100\ t_s/t$ [%], where t_s is the sum of stroke duration estimated by the model, and t is the total analyzed period.

Algorithm 1 S/H discriminant model construction

Collect RRI data from both patients with stroke and healthy persons and label the RRI data as stroke or healthy.
Derive HRV features from the collected RRI data.
Combine the derived HRV features into one modeling data matrix.
Preprocess the modeling data matrix appropriately.
Construct the S/H discriminant model from the preprocessed matrix by using SVM.

Algorithm 2 Stroke Detection

Derive HRV features from RRI data.
Preprocess the derived HRV features.
Classify subject condition into stroke or normal at every RRI collection by using the constructed model.
Calculate the stroke ratio A.
if $A \geq \bar{A}$ **then**
    Output stroke.
else
    Output healthy.
end if

Finally, the subject is diagnosed as stroke when the stroke ratio is more than the predefined threshold $\bar{A}$.

IV. Application to Clinical Data

This section reports actual application results of the proposed stroke detection method to clinical data, and discusses the future possibility of realizing an HRV-based stroke detection system.

A. Data Collection

The clinical data of patients with stroke were collected at the stroke care unit (SCU) of the Shimizu hospital, and a wearable RRI sensor [7] was used for RRI data collection. Meanwhile, the ECG data of healthy persons were collected at the Shiga University of Medical Science (SUMS) hospital when they were in Polysomnography (PSG). The R waves in the ECG data recorded in PSG were detected and each RRI was calculated. In both sets of collected data, the RRI data containing strong artifacts were eliminated and the RRI data collected during rest were clipped for analysis.

These data collection and analysis were approved by the Research Ethics Committee of the Shimuzu hospital and the SUMS hospital. The numbers of collected stroke patients and healthy persons were six and ten, respectively. The subject profiles are shown in Tables I and II. Measurement start time [A] and [B] in Table I denote the elapsed time from estimated stroke onset time, and stroke surgery start time to RRI measurement start time, respectively. Empty fields in Measurement start time [B] mean patients did not have surgery.

B. RRI Data and HRV Features

The obtained raw RRI data of patient 2 and healthy person 3 recorded during rest are shown in Figs. 3 and 4. These figures show that fluctuation of the RRI data of the stroke patient was smaller than that of the healthy person and that the patient experienced tachycardia while the healthy person did not; this result is consistent with the previous research [6]. Other patients except patients 1 and 6 showed the same tendency. As shown in Fig. 5, patient 1 also experienced tachycardia, although fluctuation of their RRI data was similar to the healthy person.

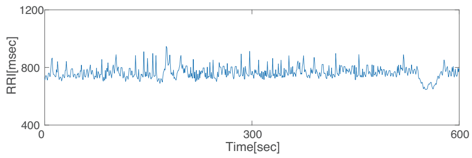

Fig. 3. Obtained RRI date of patient2

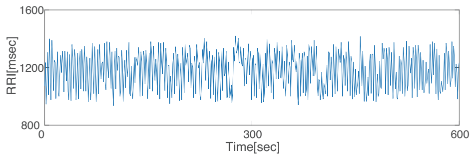

Fig. 4. Obtained RRI date of healthy person 3

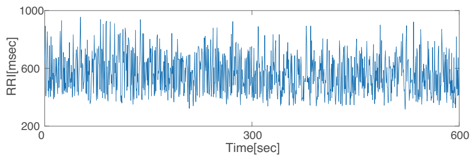

Fig. 5. Obtained RRI date of patient1

For HRV analysis, a rectangular window function was used with a window size of three minutes. The time domain features were extracted directly from the raw RRI data. On the other hand, the RRI data was resampled so that its sampling points are arranged at equal intervals for frequency analysis. In this work, the third-order spline was used for RRI interpolation, and the sampling interval was one second. An AR model of order ten was used to calculate frequency domain features. The total number of HRV features in this analysis was ten.

Table III summarizes the derived HRV features of each patients, and Table IV shows the mean and the standard deviation of HRV features derived from healthy persons and stroke patients, respectively. In addition, Figures 6 - 8 are the derived HRV features from the RRI data shown in Figs. 3 - 5. They show that HRV features of patients 2-5 were smaller than healthy persons while those of patients 1 and 6 were larger than other patients.

By comparing HRV features between healthy persons and stroke patients, only meanNN has significant difference ($p < 0.01$) as shown in Table IV. However, this result is not consistent with the previous research, which showed meanNN, SDNN, RMSSD, LF, and HF of stroke patients are significantly smaller than healthy persons [6].

According to Tables I and III, the damaged brain areas of patients 1 and 6 were extensive while those of other patients

TABLE I
PATIENTS DEMOGRAPHIC AND CLINICAL CHARACTERISTICS

Patient	Sex	Age	Stroke type	Measurement start time [A]	Measurement start time [B]
p1	female	74	CI (extensive), ICO	3h 41m	
p2	male	66	CI	7h 43m	5h 13m
p3	male	66	ICH	4h 00m	7h 00m
p4	male	85	perimesencephalic SAH	18h 30m	
p5	female	89	SAH	6h 00m	5h 45m
p6	female	81	right MCAO (extensive)	8h 00m	

CI:Cerebral Infraction, ICO:Internal carotid artery confinement, ICH:Intercerebaral Hemorrhage, SAH:Subaracnoid Hemorrihage, MCAO:Middle Cerebral Artery Occlusion

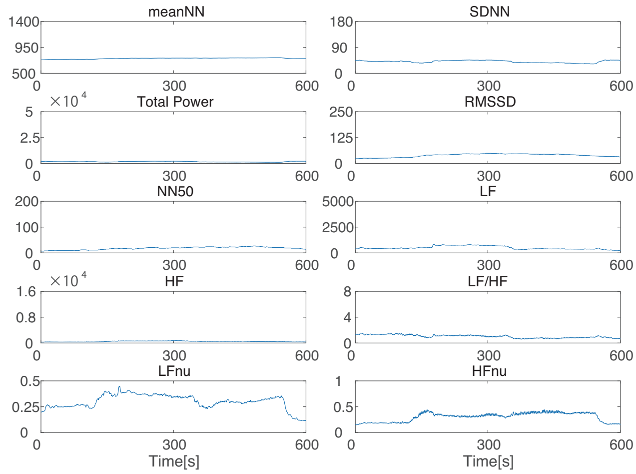

Fig. 6. HRV features derived from patient 2

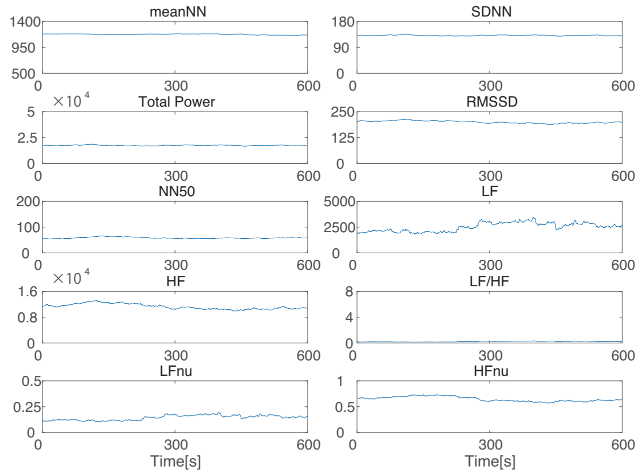

Fig. 7. HRV features derived from healthy person 3

were localized, and SDNN, Total Power, RMSSD, NN50, LF, and HF of patients 1 and 6 were larger than those of other patients.

The mean and the standard deviation of HRV features when HRV data of patients 1 and 6 were eliminated are shown in the bottom row in Table IV, and there were significant differences in all HRV features ($p < 0.05$) except LF/HF, LFnu and HFnu as well as meanNN ($p < 0.01$). This result agrees with the previous research [6], and indicates that the area of the brain damaged by the stroke affects HRV change. In fact, Barron *et al.* and Naver *et al.* have suggested that effect on HRV differs depending on the cerebral hemisphere damaged by the stroke [8], [13].

These results indicate that it is difficult to detect stroke by monitoring changes in respective HRV features, and that relationships between multiple features should be monitored together.

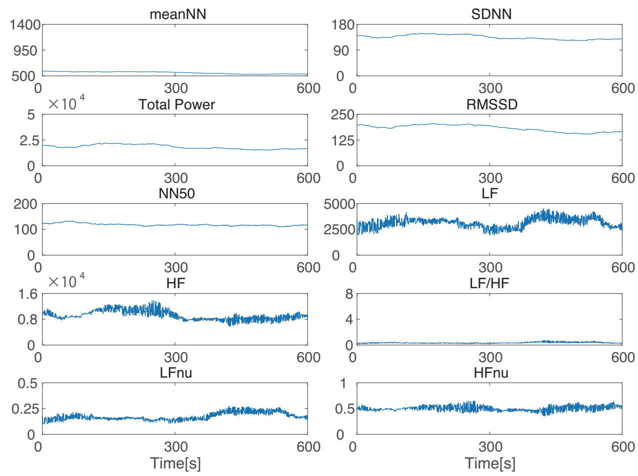

Fig. 8. HRV features derived from patient 1

TABLE II
HEALTH PERSONS DEMOGRAPHIC CHARACTERISTICS

subject	Sex	Age	subject	Sex	Age
h1	female	19	h6	male	19
h2	female	19	h7	female	21
h3	female	20	h8	male	23
h4	male	20	h9	male	19
h5	male	22	h10	male	19

C. Stroke Detection

The stroke HRV data for ten minutes were clipped from three stroke patients p1 – p3, while the healthy HRV data

TABLE III
HRV FEATURES OF EACH PATIENT

Patient	meanNN	SDNN	Total Power	RMSSD	NN50	LF	HF	LF/HF	LFnu	HFnu
1	555±19	135±8	1824±2092	181±16	118±5	3127±511	9151±1489	0.35±0.10	0.17±0.04	0.50±0.05
2	759±10	41±4	1672±335	39±8	18±5	500±160	492±132	1.03±0.24	0.30±0.07	0.30±0.08
3	755±7	29±3	853±160	29±4	2.92±0.86	154±74	319±82	0.49±0.26	0.17±0.06	0.38±0.09
4	915±5	15±3	220±84	7.5±3.4	0.07±0.25	50±14	21±21	3.13±1.17	0.24±0.05	0.11±0.11
5	670±12	19±3	359±114	21±3	3.84±4.36	50±13	163±47	0.31±0.05	0.14±0.03	0.46±0.05
6	899±15	61±9	3811±1154	74±13	19±3	1167±531	1502±543	0.81±0.31	0.30±0.08	0.39±0.06

TABLE IV
HRV FEATURES OF HEALTHY PERSONS AND STROKE PATIENTS

subject	meanNN	SDNN	Total Power	RMSSD	NN50	LF	HF	LF/HF	LFnu	HFnu
healthy persons	1072±160	75±39	7416±6473	77±55	30±22	1162±1016	2708±3560	0.70±0.47	0.81±0.06	0.45±0.16
patients 1–6	759±137**	50±45	4192±7006	59±64	27±45	841±1120	1942±3571	1.02±1.07	0.22±0.07	0.36±0.14
patients 2–5	775±102**	26±12*	776±656*	234±13*	6.2±8.0*	189± 214*	249± 203*	1.24±1.30	0.21±0.07	0.31±0.15

$^{*}p < 0.05,^{**}\,p < 0.01$ different from healthy persons by the Mann-Whitney two sample test

for ten minutes were clipped from five healthy persons h1 – h5, and they were combined into one modeling data matrix. In order to determine the threshold of the stroke ratio, the stroke and healthy HRV data that were not used for modeling were clipped from the same subjects as the modeling data, and they were combined into one threshold tuning data matrix. On the other hand, the validation data were constructed from three stroke patients p4 – p6 and five healthy persons h6 – h10. Each HRV feature was standardized as zero-mean and unit variance. The S/H discriminant model was constructed following Algorithm 1 by SVM with the Gaussian kernel. The threshold of the stroke ratio was determined so that the stroke and healthy HRV in the threshold tuning data could be discriminated perfectly, and it was 0.21.

Figure 9 shows the estimated stroke ratio of the validation data. By applying the threshold of the stroke ratio, almost all subjects except h6 could be discriminated correctly although the reason why h6 was discriminate as a stroke patient was unidentified. The sensitivity and the specificity of the proposed method were 100% and 80%, respectively.

This result showed the possibility of realizing a stroke detection system based on HRV analysis.

V. Conclusion and Future Works

In the present work, a new stroke detection algorithm based on HRV analysis was proposed. Although the number of subjects used for this analysis was small, the application result to the clinical data showed that the proposed algorithm functioned well, and its sensitivity and specificity were 100% and 80%, respectively. In the future work, large scale clinical data collection is needed for stroke detection performance improvment. In addition, the proposed method will be implemented into a wearable RRI sensor to realize a stroke detection system.

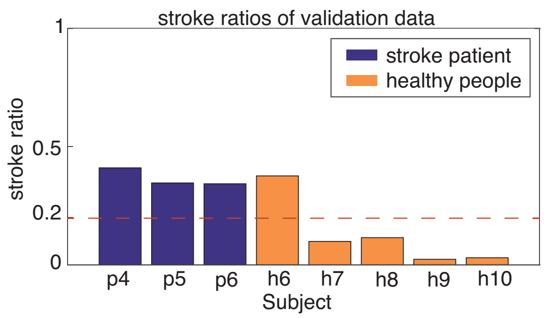

Fig. 9. Stroke detection result

References

[1] K. Toyoda, Epidemiology and Registry Studies of Stroke in Japan, Journal of Stroke, vol. 15, pp. 21–26 (2013)
[2] V. L. Feigin *et al.*, Global and regional burden of stroke during 1990-2010: findings from the Global Burden of Disease Study 2010, 2014, Lancet, vol. 383, pp. 245–254.
[3] J. M. Wardlaw *et al.*, Thrombolysis for acute ischaemic stroke, Cochrane Database Syst. Rev., 2014, vol. 7, PMID 25072528.
[4] M. C. Fang *et al.*, Trends in Thrombolytic Use for Ischemic Stroke in the United States, J. Hosp. Med., 2010, vol. 5, 406–409.
[5] M. Malik *et al.*, Heart rate variability standards of measurement, physiological interpretation, and clinical use, Eur. Heart. J., 1996. vol. 17, pp. 354–381.
[6] J. T. Korpelainen *et al.*, Abnormal heart rate variability reflecting autonomic dysfunction in brainstem infraction, Acta Neurol. Scand., 1996, vol. 94, pp. 337–342.
[7] T. Yamakawa *et al.*, Real-Time Heart Rate Variability Monitoring Employing a Wearable Telemeter and a Smartphone, APSIPA ASC 2014, Dec. 9-12, Siem Reap, Cambodia.
[8] H. K. Naver *et al.*, Reduced heart rate variability after right-sided stroke, 1996, Stroke, vol. 27, 247–251.
[9] R. L. Burr, Interpretation of normalized spectral heart rate variability indices in sleep research: a critical review. Sleep, 2007, Vol. 30, pp. 913-919.
[10] C. M. Bishop, Pattern recognition and machine learning. New York: Springer-Verlag, 2006, pp. 325–336.
[11] O. Amayri, and N. Bouguila, A study of spam filtering using support vector machines, Artif. Intell. Rev., vol. 34, no. 1, pp. 73–108, June 2010.
[12] J. R. Bock, and D. A. Gough, Predicting protein-protein interactions from primary structure, Artif. Intell. Rev., vol. 17, no. 5, pp. 455–460, May 2001.
[13] S. A. Barron *et al.*, Autonomic consequences of cerebral hemisphere infarction, Stroke, 1994, vol. 25, pp.113–116.

Acknowledgement

The authors thank the SUMS hospital for providing the healthy person data.