Feature Analysis to Estimate Sleep Time Based on Simple Measurement of Biological Information after Awakening

Mahiro Imabeppu[1*], Ren Katsurada[2], Tatsuhito Hasegawa[2]

[1] Department of Electrical, Electronic and Computer Engineering, University of Fukui, Bunkyo, Fukui, Japan.
[2] Graduate School of Engineering, University of Fukui, Bunkyo, Fukui, Japan.

* Corresponding author. Tel.: +81-776-27-8929; email: hb160145@u-fukui.ac.jp
Manuscript submitted January 10, 2020; accepted April 22, 2020.
doi: 10.17706/ijbbb.2020.10.3.144-153

Abstract: Currently, many people wear a wristband type device while sleeping to automatically record how many hours they sleep. Even a system without a wearing device, such as a smartphone application, needs to be set in advance. Therefore, automatic recording of sleep time cannot be realized without advanced measurement preparation. In this study, we propose a method to estimate sleep time without advanced preparation based on a simple measurement of biological information after awakening. We extracted 97 types of features from sensor data that were measured using wearable devices. We analyzed whether significant differences between each feature appear according to the previous sleep time. Furthermore, we evaluated the accuracy when the sleep time is estimated by machine learning using features with a significant difference. We adopted Support Vector Machine (SVM) as a machine learning algorithm and Leave-One-Session-Out Cross Validation (LOSO-CV) as an evaluation method. Consequently, there were seven features with significant differences when the biological information was measured one hour after awakening. By using machine learning, the accuracy of the previous sleep time (three sleep time categories: short, medium, or long) was estimated to be 62.5%.

Key words: Wearable device, sleep time estimation, support vector machine, electrooculogram.

1. Introduction

The problem of lack of sleep is becoming more serious in Japan. According to the National Health and Nutrition Survey conducted by the Japanese Ministry of Health, Labour and Welfare [1], in response to the question, did you get enough rest by sleeping over the past month?, 18.4% of the Japanese respondents answered “Not sleep much” or” No” in 2009; however, in 2017, that increased to 20.2%. Thus, the number of people who cannot get enough sleep has been increasing. According to statistics from the Organization for Economic Co-operation and Development (OECD), the average daily sleeping time is 528 minutes in the United States, 508 minutes in the United Kingdom, 513 minutes in France, 516 minutes in Spain, and 542 minutes in China. While there are many countries in which the sleep time exceeds 500minutes, in some countries the sleep time is shorter; it is 471 minutes in Korea, 479 minutes in Mexico, and only 442 minutes in Japan (Gender Data Portal 2019). It has also been found that the lack of sleep is negatively correlated with the risk of illness [2]. Consequently, some sleep deprivation countermeasures are needed in Japan and around the world.

Some methods have been developed to record sleep time. For example, lee BIZ[1] is a service that records sleep time by measuring body motion vibrations during sleep, the sensor of which is placed under the bed mattress. Wearable devices, such as fitbit[2] and Apple Watch[3], are smart watch devices that can record sleep time by measuring biometric information during daytime activities when the user wears it on his/her wrist arm. Smartphone applications, such as Sleep Cycle[4] and Sleep Meister[5], can record sleep patterns by placing the device beside the user's bed before sleeping; then, the alarm will ring at the optimal wake-up time. Users can view recorded sleep data on the application linked to these services. These devices and services aim to improve people's quality of sleep by enabling users to determine the amount of time they sleep and become aware of their sleep patterns. However, these devices have some disadvantages: (1) the environment in which the device is placed is limited, (2) users need to wear the devices to measure data while sleeping, and (3) users need to start the application before going to bed. Although wearable devices can easily record sleep time, the discomfort of wearing them while sleeping may interfere with the user's ability to sleep. Therefore, in this study, to solve these problems, we conduct a feature analysis for post-estimation of sleep time sleep time by simple measurement of biological information using two wearable devices: JINS MEME[6] and myBeat[7]. Estimating the sleep time from biological information after waking up can prevent the need to start the application before going to bed and the disturbance of sleep by wearing a wearable device while sleeping. When posteriori estimation of sleep time becomes possible with high estimation accuracy, our method can be applied to some scenarios, such as reducing the risk of car accidents by knowing if the driver is short of sleep before getting into the car.

2. Related Studies

In this section, we describe related studies on the estimation and measurement of sleep quality and sleep time. We describe the prediction of sleep quality using wearable device measurement data, the algorithm using the status log of the smartphone, and the method used to analyze the sleep state from voice data. Finally, we describe the objectives of our research study.

Sathyanarayana *et al*. [3] proposed a method to predict sleep quality through deep learning using biological data obtained from wearable devices. In that study, an experiment was conducted to obtain data for seven days from subjects that were always wearing a wristband type three-axis accelerometer ActiGraph GT3X +. Their method detected the user's awaking status, that is whether or not the user was sleeping, based on the sensor data measured throughout the day. Furthermore, by detecting the activity data from the sensor data while the user was awaking, their method estimated the user's sleep quality. Thus, this method predicts the quality of sleep at the time the user wakes up from biological information measured by the wearable device. It differs from our method in that the sensor device is always worn when the user is awake.

Niu *et al*. [4] proposed an algorithm, called Bedtime Prediction (BTP), that predicts the user's bedtime and wake-up time from the user's past smartphone status logs (screen status, motion, power, etc.). Many people can easily use this method because more people use smartphones than wearable devices due to the increased availability and use of smartphones in modern society. However, with BTP, it is necessary to always record the user's smartphone log data.

1 NeuroSpace: leeBIZ, https://www.neurospace.jp/leebiz
2 Fitbit: fitbit, https://www.fitbit.com
3 Apple: Apple Watch, https://www.apple.com/watch/
4 Sleep Cycle: smart alarm clock, https://apps.apple.com/am/app/sleep-cycle-smart-alarm-clock/id320606217
5 Sleep Meister, https://apps.apple.com/am/app/sleep-meister/id599456126
6 JINS: JINS MEME, https://jins-meme.com
7 UNION TOOL: myBeat, http://www.uniontool.co.jp/product/sensor/index.html

Wu *et al*. [5] conducted a study to visualize individual sleep quality by clustering sound data during sleep. Their method analyzes the quality of sleep by using voice data acquired from a microphone placed near the user's head. There is no burden on the sleeping user; however, to record audio data, it is necessary to operate an audio data recording device, such as a microphone, before going to sleep, which users may find burdensome.

The contributions of our research in comparison to the related research mentioned above is presented below.

1. By obtaining a simple measurement of biological information using a wearable device after waking up, we clarify the number of features that can influence the posteriori estimation. Moreover, to date, research on the posteriori estimation of sleep time using wearable devices has not been conducted.
2. We verify whether the number of features extracted from biological information after N hours of waking up changes significantly due to the differences in sleep time.
3. We discuss the feasibility of the estimation by performing a posteriori estimation of sleep time by machine learning using the analysis results of features.

The related research studies used sensor or voice data that were constantly measured during sleep as well as sensor data and smartphone log information measured during the waking hours of the day. In contrast, the originality of our study is that it uses experiments to verify whether it is possible to predict sleep time length simply by measuring biological information one time, N hours after the user gets up.

3. Methods

3.1. Proposed Method

Fig. 1 shows the outline of the proposed method. First, the users sleep, as usual. Then, N hours after waking up, they wear the wearable devices to measure biological information for one minute. In this study, the user solves a simple calculation task to control their measurement environment. It is expected that, as N decreases, the sleep quality is more reflected in the biological information. In this study, the value of N was one hour and three hours, in consideration of practical use in the real world. Extracting some feature values from the measured biological information, our method, finally, estimates sleep time by machine learning using selected, effective features. Our method estimates three sleep time categories: short, medium, and long.

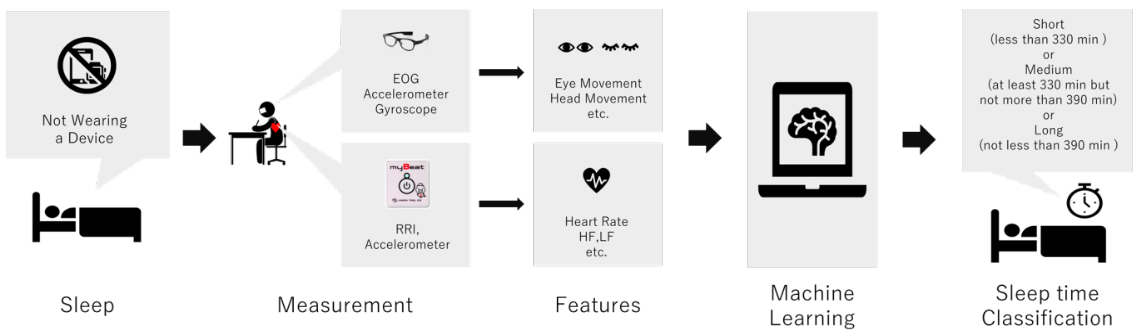


Fig. 1. The sleep time estimation process using posteriori measurement of biological information.

3.2. Devices

In this study, we used two wearable devices: myBeat and JINS MEME (Fig. 2) to measure biological information after the users woke up. myBeat is a small and lightweight wearable heart rate sensor that captures the heart cycle with high accuracy by directly picking up the electrical signals of the heart; it analyzes the autonomic balance from the acquired data. Our method uses the device type that measures heart rate from a belt that is worn near the solar plexus and touches the skin directly. Because myBeat is connected wirelessly, the device is not wired connected to a personal computer so it does not greatly obstruct the user; therefore, the stress of measurement can be decreased. myBeat can measure the cardiac cycle and the R-R interval (RRI), low frequency (LF)/high frequency (HF), and three-axis acceleration.

Fig. 2. The devices used to measure biological information simply: myBeat (Left) and JINS MEME (Right).

JINS MEME is a glasses-type wearable device that measures eye movement, acceleration, and angular velocity. JINS MEME uses three electrooculogram (EOG) measurement points to measure eye movement; it also has a six-axis motion sensor to measure head acceleration and angular velocity. There are two types of JINS MEME: JINS MEME ES, the application programming interface (API) of which acquires extracted eye-movement data from EOG by the JINS original algorithm, and JINS MEME ESR, which acquires raw data (four-dimensional EOG and six-dimensional motion sensor data). In this study, we used JINS MEME ES because the device can acquire calculated feature vectors from raw EOG data that is understandable for analysis. For example, it is possible to obtain data on the strength and speed of blinking and the direction of the eyes.

3.3. Simple Calculation Task

During the measurement of biological information, users conduct a simple calculation task, as illustrated in Fig. 3, which is based on the Kraepelin test[8]. In this test, users add adjacent numbers together and write the unit digit of the result between the numbers. This test is used to control the user's measurement environment. In this study, the correct answer rate and the number of answers of test were not considered.

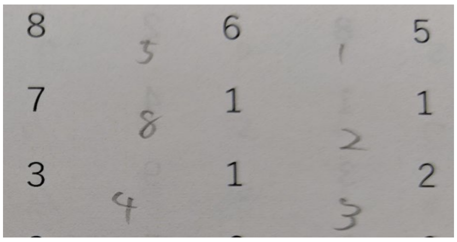


Fig. 3. Simple calculation task used to control the user's environment.

3.4. Feature Values

myBeat can acquire data on three-axis acceleration and the heart rate cycle (RRI). JINS MEME can obtain feature values on EOG, three-axis acceleration and three-axis angular velocity data of the head. These are converted into features aggregated into units of one minute of measurement. Table 1 shows that extracted feature values. There are 97 feature values; 28 feature values were extracted from the data measured by

[8]Japan・ Mental Technology Institute Uchida Kraepelin test, https://www.nsgk.co.jp/uk

myBeat and 69 feature values were extracted from the data measured by JINS MEME. The activity in myBeat is the wearer's body movement; HR is the user's heart rate and RRI is the heart rate interval. HF, LF, LF/HF, and LF_{ratio} are characteristic indexes extracted from RRI to understand the user's condition. HF and LF change depending on the balance of the sympathetic and parasympathetic tone. HF is a component in the power spectrum ranging from 0.14 Hz to 0.4 Hz; it represents the activity of the parasympathetic nervous system. LF is a component in the range of 0.05 Hz to 0.15 Hz in the power spectrum; it represents the activity of the sympathetic nervous system. LF/HF is a stress index based on the ratio of LF to HF. LF_{ratio} is calculated by $LF_{ratio} = LF/(LF + HF) \times 100$.

Table 1. The Features Extracted from the Sensor Data Measured by myBeat and JINS MEME

Devices	Features	Statistics	Dimension
myBeat	Activity	max,mean,min	3
	HF	max,mean,min,var	4
	HR	max,mean,min,var	4
	LF/HF	max,mean,min,var	4
	LF	max,mean,min,var	4
	LFratio	max,mean,min,var	4
	RRI	max,mean,min,std,var	5
JINS MEME	accX	max,mean,min,var	4
	accY	max,mean,min,var	4
	accZ	max,mean,min,var	4
	blinkStrength	max,mean,min,var	4
	blinkSpeed	max,mean,min,var	4
	blink_cnt	-	1
	eyeMoveDown	max,mean,min,var	4
	eyeMoveUp	max,mean,min,var	4
	eyeMoveLeft	max,mean,min,var	4
	eyeMoveRIght	max,mean,min,var	4
	pitch	max,mean,min,var	4
	roll	max,mean,min,var	4
	yaw	max,mean,min,var	4
	eyeMoveDown_cnt	-	1
	eyeMoveDown_cnt_0 ~eyeMoveDown_cnt_3	-	4
	eyeMoveUp_cnt	-	1
	eyeMoveUp_cnt_0 ~eyeMoveUp_cnt_3	-	4
	eyeMoveLeft_cnt	-	1
	eyeMoveLeft_cnt_0 ~eyeMoveLeft_cnt_3	-	4
	eyeMoveRight_cnt	-	1
	eyeMoveRight_cnt_0 ~eyeMoveRight_cnt_3	-	4
Total			97

In JINS MEME, accX, accY, and accZ are the head accelerations, and pitch, roll, and yaw are the angular velocities. The blinkStrength, blinkSpeed, and blink_cnt are the strength, speed, and the number of blinks, respectively. The eyeMoveUp, eyeMoveLeft, eyeMoveLeft, and eyeMoveRight are the features of the strength of the eye movement for each direction. The eyeMoveUp_cnt, eyeMoveDown_cnt, eyeMoveLeft_cnt, and eyeMoveRight_cnt are the number of eye movements for each direction. The number of blinks is counted when the strength of the blink is greater than zero. The number of up, down, left, and right movements is counted when the eyes move in each direction. eyeMoveUp_cnt_0 ~ eyeMoveUp_cnt_3, eyeMoveDown_cnt_0 ~ eyeMoveDown_cnt_3, eyeMoveLeft_cnt_0 ~ eyeMoveLeft_cnt_3, eyeMoveRight_cnt_0 ~ eyeMoveRight_cnt_3 is the index of the number of occurrences of each eye with the strength of the count of 0.

3.5. Machine Learning Algorithm

We used Support Vector Machine (SVM) [6] as the machine learning algorithm. SVM is one of supervised machine learning methods used for pattern recognition. The maximum margin with each data point is calculated from the training sample, and it learns the parameters of the linear classifier. The margin is the distance from the boundary to the nearest sample for each class. In our method, the kernel function was the Gaussian kernel, the learning parameters were cost = 0.9 and gamma = 0.8, and the others were default values.

4. Experiments

4.1. Experimental Procedure

In this experiment, the subjects were five healthy men ranging in age from 21 to 24; each subject measured for five days. They measured sensor data for one minute as one session. Data measurement was conducted twice each day: one hour and three hours after getting up. A total of 50 sessions of data were collected. The five seconds before and after the measurements were taken were removed for each session because they included extra motion, such as the noise created by preparing to take the measurement. In this study, users worn a fitbit to measure the correct label for sleep time while sleeping. fitbit can record the sleep depth and the sleep time, as shown in the sleep status graph in Fig. 4. Although our method does not require uses to wear a device while sleeping, we used fitbit to record correct sleep time data; however, the quality of sleep is not used in this study. While sleep time can be recorded through self-reporting, that method is not very accurate. As shown in Table 1, in the experiment, we extracted 97 features from the data of 44 sessions, except for six sessions where abnormal values were confirmed.

We predict sleep time by classifying it into three categories: short (Q1, less than 330 minutes), medium (Q2, 330 minutes to less than 390 minutes), and long (Q3, more than 390 minutes). The sleep time ranged from Q1 to Q3 so that the amount of data in the three categories would be as uniform as possible, and so that each category would include two or more subjects.

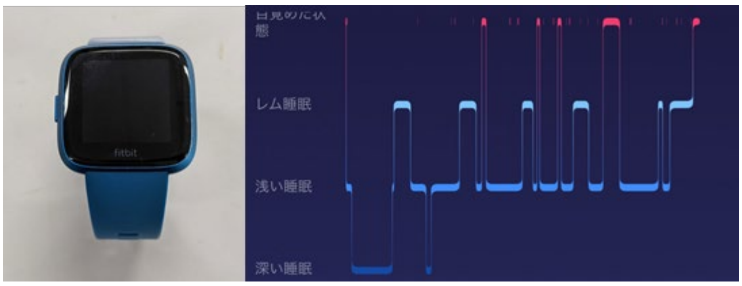


Fig. 4. The used device Fitbit (left) and the sleep status graph (right) acquired while sleeping.

4.2. Experimental Results and Discussion

Table 2. The Results of the Significant Difference Tests

Features	Significant difference after one hour			Significant difference after three hours		
	Q1-Q2	Q1-Q3	Q2-Q3	Q1-Q2	Q1-Q3	Q2-Q3
HR_min	**	**	-	-	-	-
RRI_max	*	*	-	-	-	-
accX_max	*	-	-	-	-	-
accX_mean	**	-	-	-	-	-
roll_mean	**	-	-	-	-	-
roll_min	*	-	-	-	-	-
yaw_max	*	-	-	-	-	-
eyeMoveLeft_cnt_1	-	-	-	*	-	-

Table 2 shows the results of the significant difference tests using a t-test (significance levels 5% and 1%) between Q1, Q2, and Q3. A significant difference at 5% significant levels are marked as *; differences at 1% levels are marked as **. Looking at the items obtained one hour after waking up, there are some features with a significant difference: HR_min and RRI_max obtained from the heart rate and heart rate frequency. This is because when the sleep time is short, the time it takes to reach wakefulness is long due to the lack of sleep and the heartbeat decreases to a level similar to what occurs when sleeping. In contrast, when the sleep time is long, the body is awakened and the heartbeat increases. Moreover, this result shows that there are some features related to head movement that have a significant difference: x-axis angular velocities, such as roll_mean and roll_min. We considered that the reason for this is that the posture of a subject is more stable in the awake state than in drowsiness state. Because the lack of sleep affects posture control [7], the significant difference appeared on the head-movement features. However, we did not consider individual differences in this study; therefore, there may be individual differences. We could not confirm a significant difference between Q2 and Q3 in the data after one hour. There may be no effect on the biological information measured in this study when sleep time is longer than330 minutes(Q2). Therefore, in the future, if we classify sleep time in detail, we will investigate the measurement of biological information differently, or we will add features and explore the possibility of a detailed classification.

As seen in Table 2, after three hours of sleep, there is have almost no significant difference among the items. Because the body awakens even if the sleep time is short, when the user wakes up, over time, the condition of the body is not significantly different from when the user got up.

Fig. 5 shows the box plots of three features with significant differences, HR_min, accX_mean, and roll_mean, in descending order from the top. HR_min is lower when the sleep time is shorter. If the user only slept for a short amount of time, it took too much time to wake up; therefore, the body condition was still similar to the condition while sleeping, such as decreasing HR. Since roll was the angular velocity around the head, if the sleep time was short, the posture became unstable and the angular velocity of the head increased. Moreover, when we compared the data, three hours later and one hour later, the shorter the sleeping time, the same tendency was observed as after one hour, but no significant difference was observed after three hours.

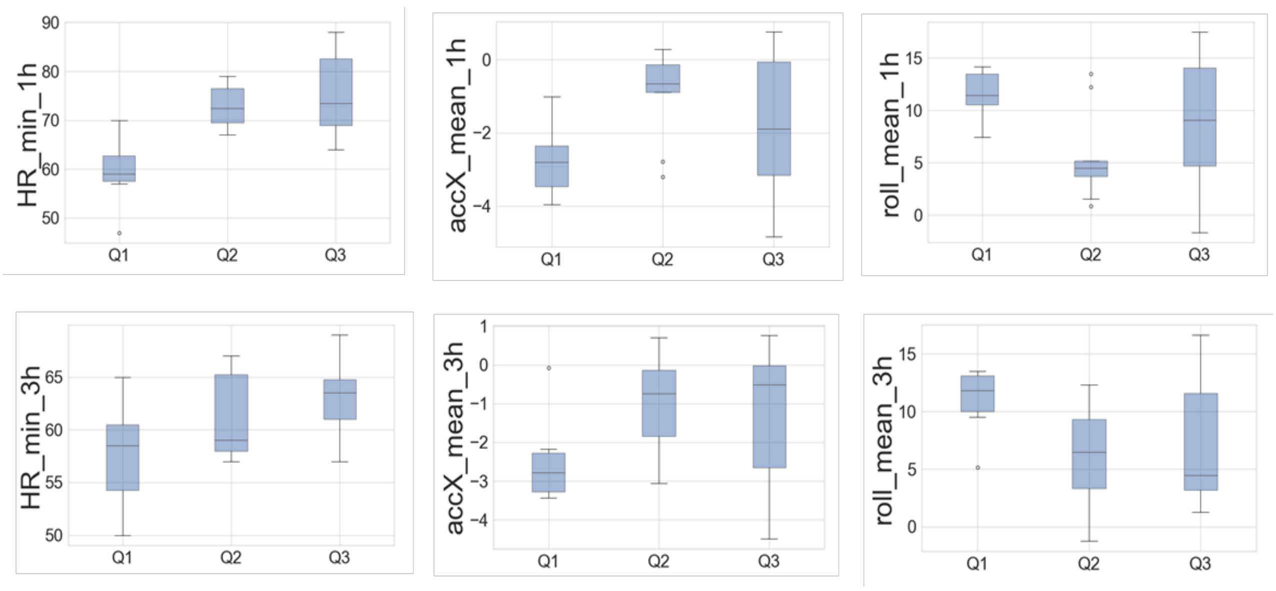


Fig. 5. One hour after the feature value with significant difference (top); Box-and-whisker diagram after three hours (bottom).

4.3. Estimation Accuracy Verification

Based on the results presented in the previous section, we evaluated the accuracy of the posteriori estimation of sleep time by machine learning. We used seven types of features from one hour after significant difference shown in Table 1 without eyeMoveLeft_cnt_1, which does not have a significant difference when measured one hour after waking up. We performed Leave-One-Session-Out Cross Validation (LOSO-CV) to verify the generalization ability of the proposed method. In LOSO-CV, one session is used as the test data and the remaining sessions are used as the training data. Moreover, since there is almost no significant difference in the feature value three hours after getting up, we used 24 sessions' data which were measured after one hour after waking up.

Table 3 shows the confusion matrix with the correct data on the vertical axis and the estimation results on the horizontal axis. Precision is calculated on the vertical axis and Recall is calculated on the horizontal axis. Our method could estimate three categories of sleep time with an overall accuracy of 62.5%. According to the confusion matrix, Q1 (short) and others (normal) can be accurately classified relatively. This is because there were no features with a significant difference between Q2 and Q3, as mentioned in the previous section. However, our method could estimate two categories of sleep time with 87.5% accuracy when estimating whether sleep time is short (Q1), in other words, whether sleep is insufficient (Table 4). This indicates that it is difficult to precisely classify sleep time, but it is possible to estimate the binary classification of whether sleep is insufficient.

Table 3. Three-Category Confusion Matrix of LOSO-CV

		Actual			Precision
		Q1: short	Q2: medium	Q3: long	
Predict	Q1: short	4	0	1	80.0%
	Q2: medium	0	7	3	70.0%
	Q3: long	2	3	4	44.4%
Recall		66.7%	70.0%	50.0%	62.5%

Table 4. Two-Category Confusion Matrix of LOSO-CV

		Actual		Precision
		Q1: short	Q2 & Q3: normal	
Predict	Q1: short	4	1	80.0%
	Q2 & Q3: normal	2	17	89.4%
Recall		66.7%	94.4%	87.5%

5. Conclusions

In this study, we performed feature analysis to estimate sleep time based on the simple measurement of biological information and post-estimation using machine learning. Based on the results of the analysis, significant differences were found from seven types of features. These features were extracted from heartbeat, heart frequency, head acceleration, and angular velocity one hour after getting up. We also used these results to evaluate the estimation accuracy for sleep time estimation by machine learning. Our method successfully estimated three categories of sleep time with 62.5% accuracy. In the future, we will conduct an additional analysis to improve the accuracy of the Q2 and Q3 sleep times. Furthermore, because the amount of sufficient sleep differs between users, we will estimate sleep efficiency, calculated by sleep status, from simple measurement of biological information.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

All authors approved the manuscript to be published, and agree to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

References

[1] Ministry of health, labour and welfare (in Japanese). (2017). *The National Health and Nutrition Survey in Japan*, Japan.

[2] Ibarra-Coronado, G. E., Pantaleón-Martínez, M. A., Velazquéz-Moctezuma, J., *et al*. (2015). The bidirectional relationship between sleep and immunity against infections. *Journal of Immunology Research*, *2015*, 1-14.

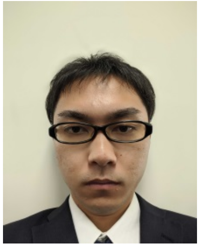
[3] Sathyanarayana, A., Joty, S., Fwrnandez-Luque, L., *et al*. (2016). Sleep quality prediction from wearable data using deep learning. *JMIR Mhealth Uhealth*, *4(4)*, e125.

[4] Niu, K., Zhang, S., Jiao, H., *et al*. (2018). BTP: A bedtime predicting algorithm via smartphone screen status. *Wireless Communications and Mobile Computing*, *2018*, 1-11.

[5] Wu, H., Kato, T., Numao, M., *et al*. (2017). Personal pleep pattern visualization via clustering on sound data. *Proceedings of the AAAI 2017 Joint Workshop on Health Intelligence* (pp. 4-9). W3PHIAI 2017. USA, California, San Francisco.

[6] Scikit-learn: Machine learning in python– scikit-learn 0.21.3 documentation. Retrieved from https://scikit-learn.org/stable/modules/generated/sklearn.svm.SVC.html

[7] Patel, M., Gomez, S., Berg, S., *et al*. (2008). Effects of 24-h and 36-h sleep deprivation on human postural control and adaption. *Exp. Brain Res., 185(2)*, 165-173.



Mahiro Imabeppu is an undergraduate student at Department of Electrical, Electronic and Computer Engineering, University of Fukui. His research interests include machine learning, wearable device.

Ren Katsurada is a student at Graduate School of Engineering, University of Fukui. His research interests include human activity recognition, applying deep learning.

Tatsuhito Hasegawa received Ph.D. degree in engineering from Kanazawa University, Ishikawa, in 2015. From 2011 to 2013, he was a system engineer with the Fujitsu Hokuriku Systems Limited. From 2014 to 2017, he was an assistant with the Tokyo Healthcare University. Since 2017, he has been a senior lecturer with the Graduate School of Engineering, University of Fukui. His research interests include human activity recognition, applying deep learning, and intelligent learning support system. He is a member of IEEE, IPSJ, HIS, and JASAG.