Sleep Stage Estimation based on The Estimated Probability of each Sleep Stage by Learning with Specialized Models

Iko Nakari*,†, Keiki Takadama†

The University of Electro-Communications, 1-5-1 Chofugaoka, Chofu, Tokyo, Japan 182-8585

Abstract

This paper proposed a novel non-contact sleep stage estimation method that ensemble the estimation results of specialized RFs that are trained with each sleep stage, which is unlike the conventional machine learning methods that train each sleep stage at once. Concretely, when ensembling the results of each RF, its estimation probabilities are employed to estimate the most likely sleep stage. Furthermore, the method employs the optimal length for learning with RF at each sleep stage. To deal with individual differences, this paper compares other probabilities of sleep stage estimation by employing hyperparameter X for each sleep stage, which subtracts the largest estimated probability to reduce false positive estimation. Through experiments, the following implications have been revealed: (1) the proposed method contributed to improving the percentage of Accuracy by 70.5% from 62.4% by the conventional machine learning method; (2) the suitable lengths of data for learning with RF were about 2, 4, and 2 min for REM sleep, N12 sleep, and N34 sleep, respectively; (3) hyperparameter X is effective in reducing biased sleep stage estimation caused by individual differences and improving estimation accuracy.

Keywords

machine learning, random forest, sleep stage estimation, mattress sensor,

1. Introduction

Sleep is one of the most important activities of life and people spend a third of their lives sleeping. If sleep deprivation (especially 4-6 hours of sleep) persists, people become in a state of sleep debt. In the state of sleep debt, the ability to think and make decisions are equivalent to staying up all night [1], and it is a factor in the increased risk of industrial and traffic accidents. It also decreases immune function and increases the risk of developing lifestyle-related diseases such as depression and dementia [2, 3]. To avoid becoming a state of sleep debt, it is important to get enough sleep, but in some cases, it is not enough due to sleep diseases. This is because sleep diseases, such as sleep apnea syndrome, worsen sleep quality so that it is the same condition as sleep deprivation even if you spend enough sleep. In order to realize our sleep problems, it is necessary to understand our daily sleep quality. The standard method for measuring sleep quality (sleep stage) is to evaluate biological data (i.e., EEG, EOG and EMG) acquired by Polysomnography (PSG) test based on the Rechtschaffen & Kales (R&K) method [4]. However, the PSG test is a highly restrictive method and requires a person to attach multiple electrodes to his/her head and body, which burdens physical and mental on his/her and prevents obtaining data on sleep as usual. For the facts, it is not suitable for measuring daily sleep quality.

To address the problems, the demand for sleep stage estimation methods by simple sensors (such as mattress sensors) has increased as an alternative to the PSG test. It is challenging to estimate sleep stages without biological data obtained from PSG test. For example, Watanabe developed a mattress sensor and focused on physiological findings which is the relation between heart rate variability and sleep stage [5]. The accuracies of the method are reported as follows: 42.8% in three stages (NREM/REM/WAKE) estimation; 82.6% in NREM estimation; 70.5% in WAKE estimation; and 38.3% in REM estimation. As the results show, the accuracy of the method is not high and it suggests that the method based on findings has limitations.

To tackle the problems, it is necessary to estimate sleep stages from a new perspective other than physiological findings. However, since we do not know what to focus on and it is difficult to find novel physiological findings, machine learning (ML) is a good way to estimate sleep stages from a new perspective. In this study, Random Forests [6] is employed for the ML model because it is easier to analyze what the model learned from the data than deep learning [7], which is widely employed because of its high prediction accuracy. It is essential for analyzing models easier because it leads to the interpretability of the model in the future. The difficult points in learning sleep stages with ML are the followings: (1) it may be different for each sleep stage that the features suitable for classification; (2) ML is not good at learning data with individual differences.

To overcome these problems, this paper aims to im-

AAAI 2023 Spring Symposia, Socially Responsible AI for Well-being, March 27–29, 2023, USA

*Corresponding author.

† These authors contributed equally.

ikko0528@cas.lab.uec.ac.jp (I. Nakari); keiki@inf.uec.ac.jp (K. Takadama)

prove the accuracy of sleep stage estimation and proposes a novel non-contact sleep stage estimation method that ensemble the estimation results of specialized RFs that are trained with each sleep stage. In addition, this paper tries to discover a way to adjust sleep stage estimation for individual differences. Concretely, this paper employs TANITA sleep scan SL511 (Japan) as the mattress sensor for acquiring bio-vibration data, and prepares several RFs for each sleep stage. When ensembling the results of each RF, its estimation probabilities are employed to estimate the most likely sleep stage. Furthermore, since some characteristics of each sleep stage appear for a short time and others appear intermittently for a long time, multiple data lengths are employed for training RF. For example, WAKE and N1 sleep suddenly occur with large body movements and other sleep stages (especially REM sleep) intermittently occur for a long time. This paper investigates the optimal data length for each sleep stage and employs it to train RF. For individual differences, this paper focuses on the other probabilities of sleep stage estimation to improve the accuracy of the estimation.

This paper is organized as follows. The next section describes the sleep mechanism. Section 3 describes the related works of non-contact sleep stage estimation and RF which is the main ML method in our proposed method and Section 4 proposes our sleep stage estimation based on an ensemble of models trained for each sleep stage. The experiment is conducted in Section 5 and the results are analyzed in Section 6. Finally, our conclusion is given in Section 7.

2. Sleep Stage

The sleep stage is an indicator of the depth of sleep defined by the R&K method. The depth of sleep is classified into six stages in each epoch (30 seconds), i.e., WAKE, REM, Non-REM1 (N1), N2, N3, and N4. N4 is often included in N3 and is denoted as N34 in this paper. The proportion of each sleep stage in healthy young adults per night is as follows: WAKE is 1-5%; REM is 15-25%; N1 is 5-20%; N2 is 45-75%; and N34 is 10-22% (depending on age and physical condition on that day). In order to determine the sleep stage, the R&K method needs biological data such as electroencephalography (EEG), electrooculogram (EOG), and electromyogram (EMG) acquired by the PSG test. Figure 1 shows the example of the overnight sleep stages, where the vertical axis indicates the sleep stage and the horizontal axis indicates the time. As shown in Figure 1, the structure of the sleep stage in a healthy person repeats deep sleep (N3 sleep) and shallow sleep (above N3 sleep) alternately, and the regular sleep repeats this cycle (about 90 to 120 minutes) three to five times a night. Each cycle is connected by about 20 to 30 minutes of REM sleep, and this cycle is called the ultradian rhythm.

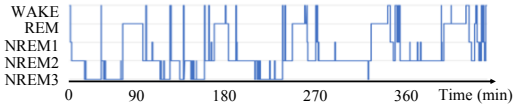


Figure 1: Example of the overnight sleep stage.

The characteristics of each sleep stage are described below. The WAKE stage and N1 sleep occur suddenly in a short time, and these sleep stages often occur with large body movements. The size of body movements tends to be larger in the WAKE stage than in the N1 stage. In addition, some WAKE stages are self-aware while others are not. In this paper, WAKE stages longer than 2 minutes are excluded because their characteristics are different from WAKE stages that occur suddenly. REM sleep is determined by focusing on "rapid eye movement" and "decreased skeletal muscle activity" in the R&K method. Other characteristics of REM sleep are increased or unstable heart rate and respiratory rate. The mattress sensors only can measure the characteristics related to heart rate and respiratory rate, but these characteristics occur intermittently, rather than continuously. The characteristics of N2 and N34 are that the heart rate and respiratory rate tend to decrease with deeper sleep (N34), and the sleep stage does not change suddenly but remains in the same sleep stage continuously.

3. Related Works

3.1. Sleep Stage Estimation by mattress sensor

Watanabe et al. tried to extract the relation between the change in the heart rate and sleep stages through the frequency band containing the multiple biological rhythms of a human to build a foundation of sleep stage estimation from heart rate variability (HRV) [8]. They focused on two biological rhythms, the ultradian rhythm and the circadian rhythm, which is an approximate 25 hours cycle. From their study, the relations between the frequency of HRV and sleep stage have been revealed, and they built a sleep stage estimation method based on the heart rate data acquired from the air mattress sensor [5].

3.2. Random Forests

This study employs Random Forests (RF) [6]. The RF model repeats random sampling from training data, randomly constructs decision trees with different conditional

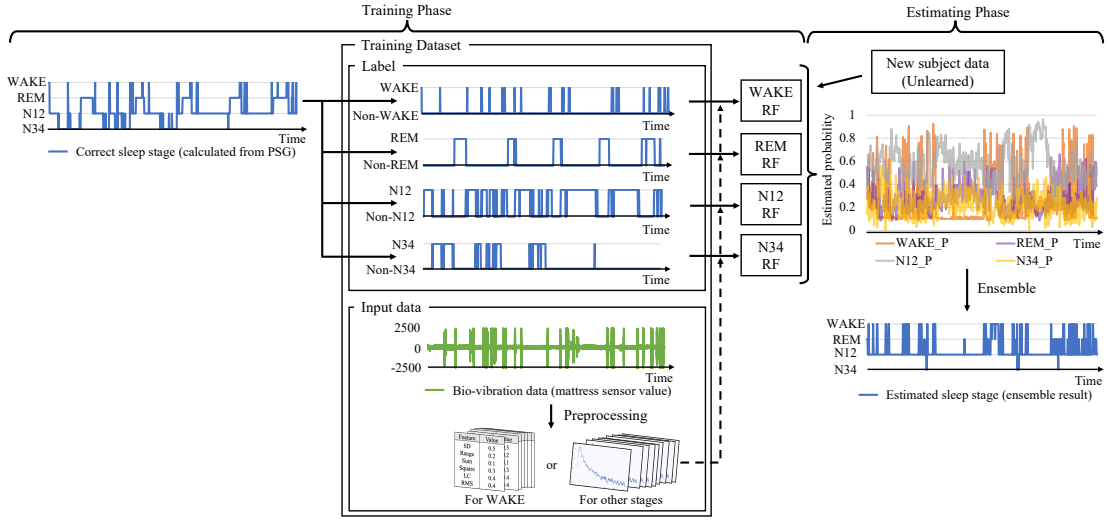


Figure 2: Overview of Multi-Timescale REM Estimation.

branches, and classifies them by the majority rule of those results. In this research, Gini impurity is the splitting condition, it becomes low when all the samples contained in each node of the decision tree are the same. RF processing is as follows:

1. Generate bootstrapped sample S_j from training data set S.
2. One-third of the original data is called Out-Of-Bug (OOB), and it is used for constructing a decision tree. Each node processing is as follows:
   a) Extract m_{try} features randomly with not allowing duplicate value.
   b) Choose the feature that minimizes Gini impurity, and divides nodes.
3. Repeat 1. to 2. N_{tree} times.

where N_{tree} is the number of decision trees. In the classification problem, it is recommended to use the square root of the total number of features for the variable m_{try}, which is employed to divide the nodes of decision trees.

4. Ensemble of Models Trained for each Sleep Stage

Figure 2 is the overview of the proposed method. The proposed sleep stage estimation method, Ensemble of Models Trained for each Sleep Stage, starts from training four RFs (i.e., trained with "WAKE and others," "REM sleep and others," "N12 sleep and others" and "N34 sleep and others," respectively). Then the method estimates sleep stages by an ensemble of four RFs based on their estimation probabilities. All graphs in Figure 2, the horizontal axis represents "time", and the vertical axis represents sleep stages, bio-vibration data or estimated probability of RF.

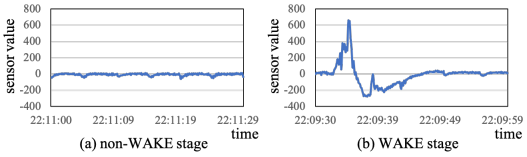


Figure 3: The example of bio-vibration data in (a) non-WAKE stage and (b) WAKE stage.

4.1. Input data

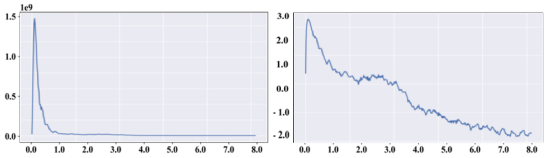
For training each sleep stage, this paper designs two types of input data. The first type is the features related to body movement for the WAKE stage, and the other type is the logarithmic power spectrum for the non-WAKE stage (i.e., REM sleep, N12 sleep and N34 sleep).

4.1.1. Features for WAKE Stage

Fig. 3 shows the examples of the bio-vibration data in one epoch in (a) non-WAKE stage and (b) WAKE stage, where the vertical axis indicates the sensor value while the horizontal axis indicates time. As shown in Fig. 3, the amplitude of bio-vibration data is small in non-WAKE

Table 1
Six features based on bio-vibration data in one epoch.

feature	Formula
SD	$\sigma(x)$
Range	$max(x) - min(x)$
SUM	$\sum_{n=1}^{N}(x_n)$
Square	$\sum_{n=1}^{N}(x_n^2)$
LC	$\frac{1}{N-1}\sum_{n=1}^{N-1}(x_{n+1} - x_n)$
RMS	$\sqrt{\frac{1}{N}\sum(x^2)}$



(a) power spectrum (b) logarithmic spectrum

Figure 4: Examples of (a) power spectrum and (b) logarithmic spectrum ($L = 64$).

stage, while that is large in the WAKE stage. From this fact, it is suitable to estimate the WAKE stage by designing the features such as the variance of the bio-vibration data in one epoch from the viewpoint of body movement. Our previous work designed the following six kinds of features from the bio-vibration data in one epoch are designed as shown in (Table 1): standard deviation (SD), the difference between maximum and minimum (Range), the summation (SUM), the sum of squares (Square), the average of variation (Level Change: LC) and the Root-Mean-Square (RMS) [9].

4.1.2. Features except for WAKE stage

To make it easier for ML to learn the characteristics of bio-vibration data for each sleep stage, the frequency analysis is applied for decomposing each vibration (i.e., heartbeats, respiration and body movement) to frequencies. This process is conducted as follows.

1. Applying the Fast Fourier Transform (FFT) [10] to the bio-vibration data in a L-second window to convert the data to a power spectrum (note that the sampling frequency of the mattress sensor is 16Hz, and data size is $L \times 16$). In this study, window size (L) is set as next for capturing several scales of REM. $L = \{32, 64, 128, 256\}$. According to the sampling theorem [11], the frequency that can be analyzed by FFT is up to 8Hz, so that the data size of the power spectrum is $L \times 8$, and the frequency resolution is $1/L$ Hz. Figure 4(a) shows the example of the power spectrum ($L = 64$) calculated from bio-vibration data, where the vertical axis indicates the density of the power spectrum and the horizontal axis indicates the frequency. In particular, the frequency band between 0.1Hz and 0.3Hz is related to respiration, and the frequency band between 0.6Hz and 1.5Hz is related to heartbeats. Regarding the BM, the larger/smaller BM, the higher/lower density of the power spectrum. However, as shown in Figure 4(a), it is difficult to understand the shape of the power spectrum above 1Hz because of the high density of frequencies below 1Hz.
2. In order to make it easier to understand above 1Hz and for RF to learn, the power spectrum is converted into a logarithmic spectrum (log10). Figure 4(b) shows the example of the logarithmic spectrum converted from the power spectrum of Figure 4(a), where the vertical axis indicates the density (logarithmic value) of the spectrum and the horizontal axis indicates the frequency. Furthermore, the density of each frequency in the logarithmic spectrum is normalized to 0, 1 based on the value of the density of the overall frequency.
3. This logarithmic spectrum is calculated per 30 seconds (stride size is 30 seconds) and labeled with the correct sleep stage (one sleep stage/others) determined by R&K method for RF to learn. Figure **??** shows the example of strides (window size is 128 seconds) and how to label the sleep stage to the spectrum. When labeling the sleep stage to the spectrum, bio-vibration data often have multiple sleep stages, so that, in this study, the sleep stage which is labeled to the spectrum is determined by a majority vote of the proportions occupied by those sleep stages. Note that, when using RF for the sleep stage estimation (not the learning phase), the logarithmic spectrum is not labeled with the correct sleep stage, and the output of the estimation is for the first epoch.

4.2. Ensemble based on Estimated Probability and Hyperparameter X

Each sleep stage is trained by a separate RF, and each RF selects and trains features specific to each sleep stage. When data is input to each RF, the probability of each sleep stage is output by each RF. Then, the sleep stage with the highest probability is determined as a result of the ensemble. However, it is also possible that the probability of both two sleep stages increases (i.e., the estimated probabilities of WAKE, REM, N12 and N34

Table 2
Information of healthy subjects.

ID(Age)	WAKE	REM	N12	N34	Total
A (20's)	46(28)	176	462	164	848(830)
B (20's)	53(34)	113	416	2	584(565)
C (30's)	103(22)	184	417	0	704(623)
D (40's)	75(49)	65	465	2	607(581)
E (40's)	44(28)	85	443	23	595(579)
F (40's)	34(18)	121	265	0	420(404)
G (40's)	53(30)	110	487	1	651(628)
H (50's)	35(23)	249	576	0	860(848)
I (60's)	98(29)	159	463	0	720(651)

are 90%, 50%, 80%, 30% respectively) due to individual differences. In particular, the probability of the WAKE stage increases with more body movement, but some subjects may not be in the WAKE stage.

To deal with the problem, the method subtracts the hyperparameter $X\%$ from the probability of the sleep stage with the highest probability before comparing it to the other sleep stages. If subtracting the probability changes the relationship between the probabilities of each sleep stage, the sleep stage with the highest probability at that time will be the result of the ensemble. For example, the probabilities of WAKE, REM, N12 and N34 are 90%, 50%, 80%, 30% respectively, and the hyperparameter for WAKE stage (X_W) is set as 20%, the probabilities are changed to 70%, 50%, 80%, 30% respectively. In this case, the ensemble result is changed to N12 sleep from the WAKE stage.

5. Experiments and Results

To investigate the effectiveness of the proposed sleep stage estimation based on an ensemble of models trained for each sleep stage, this paper conducted the human subject experiments of nine healthy subjects. The experiments are divided into three parts. The purposes of each experiment are the followings:

1. To determine a suitable window length of input data for each sleep stage.;
2. To determine the hyperparameter X_W, X_R, X_{N12} and X_{N34} for the highest estimated probability of ensembling each of the RFs corresponding to the four sleep stages.
3. Sleep stage estimation with the hyperparameters selected in Experiment 1.

The information on the subjects is summarized in Table 2. The column "ID (Age)" indicates the ID and the age of the subject. The columns from "WAKE" to "N34" indicate the number of epochs in each sleep stage (WAKE, REM, NREM1, NREM2 and NREM34), and the column "Total" indicates the total number of epochs in one night. Note that, the numbers in parentheses for WAKE and Total represent the number of epochs excluding consecutive awakenings of more than 2 minutes (sleep disruption) and are excluded from the calculation of Accuracy rates in this paper. The average number of epochs (30 seconds) of sleep without sleep disruption is 634 ± 136.

The performance of the proposed method is compared with single RF trained with four sleep stages at once. As evaluation criteria, this study employs the Accuracy rate between PSG-based sleep stages and estimated sleep stages. In addition to Ex.1, Precision, Recall, F-measure and Specificity of each sleep stage estimation are employed.

5.1. Experimental Setup

The electrodes were attached to the body and head of each subject to acquire EEG, EOG and EMG, and the mattress sensor was placed under the mattress in the bed to acquire bio-vibration data in one night. After sleep, the correct sleep stages for each subject were determined according to the R&K method based on the data measured by PSG (helped by a medical specialist), and the bio-vibration data measured by mattress sensor is converted to "Features for the WAKE stage" or "Features except for WAKE stage" that logarithmic spectrums of several scales (i.e., window sizes are $L = \{32, 64, 128, 256, 512\}$). Each feature is labeled with the correct sleep stage in each epoch. The parameters of RF are set as follows: (i) the maximum depth of the decision tree is 10; (ii) the number of the decision tree is 50; (iii) the number of features employed to construct the decision tree is 16, 23, 32, 46 and 64 for window sizes 32, 64, 128, 256 and 512 respectively.

5.2. Ex.1 Comparison of Estimation Accuracy by Window Length of Input Data for each Sleep Stage

Ex.1 is conducted for determining a suitable window length of input data for each sleep stage except for the WAKE stage. As evaluation criteria, Ex.1 employs the Accuracy, Precision, Recall, F-measure and Specificity of REM sleep, N12 sleep and N34 sleep respectively.

Figure 5, Figure 6 and Figure 7 show the averaged results of each evaluation criteria for each sleep stage. The vertical and horizontal axes indicate the percentage and each criterion, respectively. The blue, orange, gray and yellow bars indicate the length of the window (input data), i.e., 32 seconds, 64 seconds, 128 seconds and 256 seconds. In REM sleep, "128 seconds size of window length" is the most suitable for REM sleep estimation because Precision, Recall and F-measure are higher than any other window length. In N12 sleep, "256 seconds

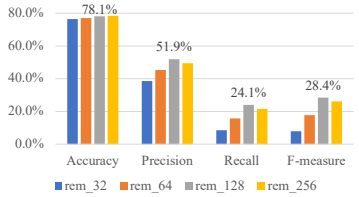


Figure 5: Ex.1 Averaged results of the REM sleep.

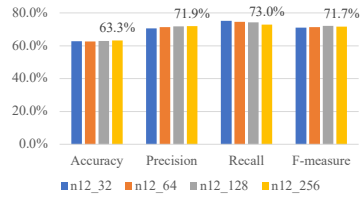


Figure 6: Ex.1 Averaged results of the N12 sleep.

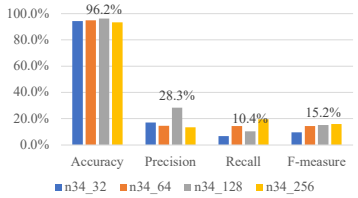


Figure 7: Ex.1 Averaged results of the N34 sleep.

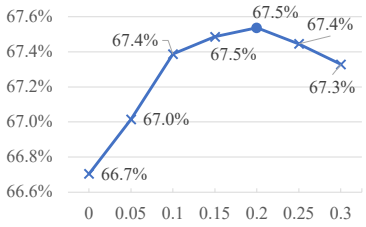


Figure 8: Ex.2 Averaged accuracy of sleep stage estimation in sensitivity analysis of X_W.

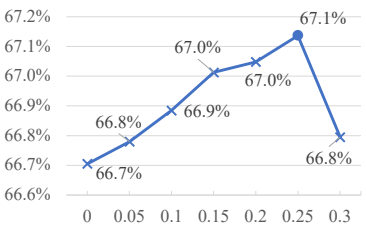


Figure 9: Ex.2 Averaged accuracy of sleep stage estimation in sensitivity analysis of X_R.

size of window length" is the most suitable for N12 sleep estimation because there is no significant difference in all F-measures, and Accuracy is higher than the other window lengths.

5.3. Ex.2 Sensitivity Analysis of Hyperparameter X

Ex.2 is conducted for determining the hyperparameter X_W, X_R, X_{N12} and X_{N34} for the highest estimated probability of ensembling each of the RFs corresponding to the four sleep stages. In Ex.2 all sleep stages are estimated by the flow of the proposed method, which ensembles RFs specialized for estimating each sleep stage. Since the adjustment of the hyperparameter X for each sleep stage changes the overall result of sleep stage estimation, the evaluation criterion is the percentage of Accuracy between the sleep stage from PSG and the predicted sleep stage. Hyperparameter X for each sleep stage has experimented with values of 0.00, 0.05, 0.10, 0.15, 0.20, 0.25 and 0.30.

Figure 8, Figure 9, Figure 10 and Figure 11 show the results of sensitivity analysis for Hyperparameter X for each sleep stage. The vertical and horizontal axes indicate the accuracy of sleep stage estimation and the value of hyperparameter X. From the results, the most suitable hyperparameter X is determined as the followings: $X_W = 0.20$, $X_R = 0.25$, $X_{N12} = 0.00$ and $X_{N34} = 0.30$.

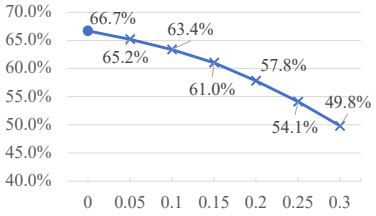


Figure 10: Ex.2 Averaged accuracy of sleep stage estimation in sensitivity analysis of X_{N12}.

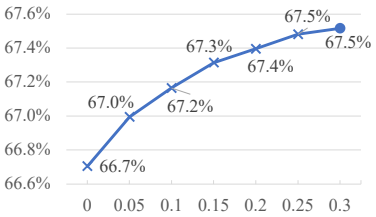


Figure 11: Ex.2 Averaged accuracy of sleep stage estimation in sensitivity analysis of X_{N34}.

5.4. Ex.3 Single RF vs. Proposed Method

Ex.3 is conducted for validating the effectiveness of the proposed method and the method is compared with the result of RF trained with all sleep stages at once. For the proposed method, hyperparameter X has experimented

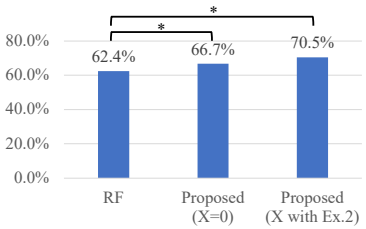


Figure 12: Ex.3 Averaged accuracy of sleep stage estimation by RF and the proposed method with $X = 0$ and that with X with Ex.2.

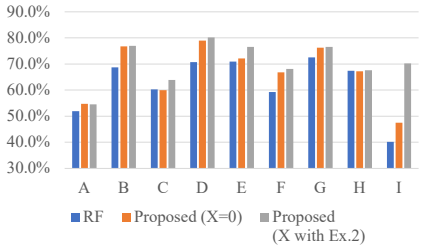


Figure 13: The results of each subject in Ex.3.

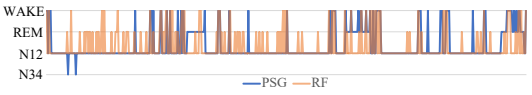


Figure 14: The result of sleep stage estimation by single RF for subject D.

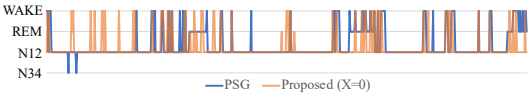


Figure 15: The result of sleep stage estimation by proposed method ($X = 0$) for subject D.

with two patterns, $X = 0$ and X with the results of Ex.2. As evaluation criteria, Ex.3 employs the percentage of Accuracy between the sleep stage from PSG and the predicted sleep stage. The vertical and horizontal axes indicate the accuracy of sleep stage estimation and estimation methods respectively.

Figure 12 shows the result of sleep stage estimation by single RF, proposed with $X = 0$ and proposed with the results of Ex.2. The vertical and horizontal axes indicate the accuracy of sleep stage estimation and sleep estimation methods. As shown in Figure 12, the estimation accuracies of RF, proposed method ($X = 0$) and the proposed method (X with the results of Ex.2) are 62.4%, 66.7% and 70.5% respectively. The proposed method is 8.1% higher than the conventional sleep stage estimation method that classifies 4 sleep stages by s single RF. This paper conducted a paired t-test for the results and found a significant difference between the single RF and the proposed method. This result suggests, the proposed method significantly increases the Accuracy of sleep stage estimation more than conventional single RF.

6. Discussion

6.1. Single RF vs. Proposed Method

To discuss the differences in sleep stage estimation results between single RF and the proposed method, Figure 13 shows the results of each subject. The vertical and horizontal axes indicate the accuracy of sleep stage estimation and subject IDs. The blue, orange and gray bars indicate the result of single RF, proposed with $X = 0$ and proposed with the results of Ex.2. The accuracies of the estimation for each subject are improved by the proposed method. Figure 14 and Figure 15 show the estimation result of subject D by single RF and the proposed method ($X = 0$). The vertical and horizontal axes indicate the sleep stage and time. The blue and orange lines indicate the correct sleep stage (obtained from PSG) and estimated sleep stage. The estimation by single RF has more false REM estimations than the estimation by the proposed method. Such false REM estimations by single RF tend to occur in other subjects, and some subjects also have more false WAKE stage estimations than estimations by the proposed method. Since the proposed method trains a dedicated RF for each sleep stage, it can learn the characteristics of each sleep stage from multiple perspectives. On the other hand, a single RF must learn the characteristics of four sleep stages, which limits the number of characteristics that can be learned. This is the reason why the proposed method has higher estimation accuracy than the single RF.

6.2. Effectiveness of Hyperparameter X.

As shown in Figure 13, even with the proposed method, the estimation may not improve the Accuracy as in Subject I. The lack of improvement in the Accuracy rate is caused by the over-estimating REM sleep as shown in Figure 16. This means that subject I, who is elderly, has over-estimating REM sleep because the bio-vibration data in normal conditions are similar to the bio-vibration data during REM sleep in the training data (i.e., the estimated probability of REM sleep is often greater than 50%). Even if the estimated probability of REM sleep is high, the estimation probabilities of other sleep stages (espe-

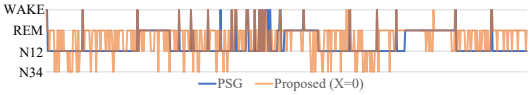


Figure 16: The result of sleep stage estimation by proposed method ($X = 0$) for subject I.

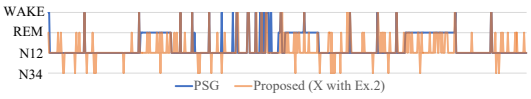


Figure 17: The result of sleep stage estimation by proposed method (X with Ex.2) for subject I.

cially N12 sleep) may also be high. Based on this, when only the estimated probability of REM sleep is high, REM sleep is employed as the estimation result, and when the estimated probability of other REM sleep is also high, estimation of other sleep stages can be employed to reduce over-estimating REM sleep estimation. Hyperparameter X, which is set for each sleep stage, can deal with this problem because it subtracts X% from the probability of the sleep stage with the largest estimated probability.

Then, Subject I was able to reduce over-estimating REM sleep estimation by setting the hyperparameter X obtained in Ex.2.

This suggests that hyperparameter X is effective in reducing biased sleep stage estimation caused by individual differences and improving estimation accuracy.

7. Conclusion

This paper proposed a novel non-contact sleep stage estimation method that ensemble the estimation results of specialized RFs that are trained with each sleep stage. Concretely, when ensembling the results of each RF, its estimation probabilities are employed to estimate the most likely sleep stage. Furthermore, the method employs the optimal length for learning with RF at each sleep stage. To deal with individual differences, this paper compares other probabilities of sleep stage estimation by employing hyperparameter X for each sleep stage, which subtracts the largest estimated probability to reduce false positive estimation. Through experiments, the following implications have been revealed: (1) the proposed method contributed to improving the percentage of Accuracy by 70.5% from 62.4% by the conventional machine learning method; (2) the suitable lengths of data for learning with RF were about 2, 4, and 2 min for REM sleep, N12 sleep, and N34 sleep, respectively; (3) hyperparameter X is effective in reducing biased sleep stage estimation caused by individual differences and improving estimation accuracy.

The future tasks are the followings: (1) expansion of subject data to improve the accuracy of N34 estimation due to the small amount of N34 sleep data; (2) to improve the estimation accuracy by classifying the individual difference trends from subject sleep data and learning for each similar subject.

References

[1] H. Van Dongen, G. Maislin, J. M. Mullington, D. F. Dinges, The cumulative cost of additional wakefulness: dose-response effects on neurobehavioral functions and sleep physiology from chronic sleep restriction and total sleep deprivation, Sleep 26 (2003) 117–126.

[2] J. M. Mullington, M. Haack, M. Toth, J. M. Serrador, H. K. Meier-Ewert, Cardiovascular, inflammatory, and metabolic consequences of sleep deprivation, Progress in cardiovascular diseases 51 (2009) 294–302.

[3] C. Holingue, A. Wennberg, S. Berger, V. Y. Polotsky, A. P. Spira, Disturbed sleep and diabetes: A potential nexus of dementia risk, Metabolism 84 (2018) 85–93.

[4] A. Rechtschaffen, A. Kales, A Manual of Standardized Terminology, Techniques and Scoring System for Sleep Stages of Human Subjects, Washington DC, 1968.

[5] T. Watanabe, K. Watanabe, Noncontact method for sleep stage estimation, IEEE Transactions on biomedical engineering 51 (2004) 1735–1748.

[6] L. Breiman, Random forests, Machine learning 45 (2001) 5–32.

[7] I. Goodfellow, Y. Bengio, A. Courville, Y. Bengio, Deep learning, volume 1, MIT press Cambridge, 2016.

[8] T. Watanabe, K. Watanabe, Estimation of the sleep stages by the non-restrictive air mattress sensor relation between the change in the heart rate and sleep stages, Transactions of the Society of Instrument and Control Engineers 37 (2001) 821–828.

[9] I. Nakari, Y. Tajima, R. Takano, A. Toboru, K. Takadama, WAKE Detection During Sleep using Random Forest for Sleep Apnea Syndrome Patient, AAAI Spring Symposium, 2019.

[10] J. W. Cooley, J. W. Tukey, An algorithm for the machine calculation of complex fourier series, Mathematics of computation 19 (1965) 297–301.

[11] C. E. Shannon, Communication in the presence of noise, Proceedings of the IRE 37 (1949) 10–21.