

## Article

# Sleep Stage Classification based on Two-Cycle Sleep Model Recognition using Continuous Heart Rate Variability

Hongbo Ni<sup>1\*</sup>, Mingzhe Guo<sup>2</sup> and Ying Wang<sup>3</sup>

1 Northwestern Polytechnical University, Xi'an, China; nihb@nwpu.edu.cn

2 Northwestern Polytechnical University, Xi'an, China; guomingzhe96@qq.com

3 Northwestern Polytechnical University, Xi'an, China; 2450857608@qq.com

\* Correspondence: nihb@nwpu.edu.cn; Tel.: +86-29-88431531

**Abstract:** Sleep stage on the whole night is not steady. Sleepers generally pass through three to five cycles. In each cycle, there are occur four typical sleep stages, such as wake stage (WS), light stage (LS), deep sleep (DS), rapid eye movement sleep stage (REM). According to the natural routine, in this paper, we investigate the stage transition and analyze the feature of stage transition using the local cluster Algorithm (LCA). Two-cycle sleep model (TCSM) is proposed to automatically classify sleep stages using over-night continuous heart rate variability (HRV) data. The generated model is based on the characteristics of the nested cycle's sleep stage distribution and the transition probabilities of sleep stages. Experiments were conducted using a public data set including 400 healthy subjects (female 239, male 161) and the model's classification accuracy was evaluated for four sleep stages: WS, LS, DS, REM. The experimental results showed that based on the TCSM model, the segmentation classification of pure sleep is 5.2% higher than that of the traditional method, and the accuracy of segmentation classification is 11.2% higher than the traditional sleep staging accuracy. The experimental performance is promising in terms of the accuracy, sensitivity, and specificity rates compared with the ones of the state-of-the-art methods. The study contributes to improve the quality of sleep monitoring in daily life using easy-to-wear HRV sensors.

**Keywords:** sleep stage classification, ECG, nested-cycle sleep pattern, stage transition

## 0. Introduction

Sleep covers almost one-third of the human lifespan, and quality of sleep is very important. Sleep has a very important physiological function, which is of great significance for the recovery of physical strength and energy, hormone secretion rhythm, consolidation and reintegration of memory, and immune function. The quality and quantity of sleep impact the performance of many basic activities, such as learning, memorization, and concentration [1]. Many sleep disorders are related to sleep stage, such as OSA [2], major depressive disorder [3], insomnias, sleep-related breathing disorders, circadian rhythm sleep-wake disorders, sleep movement disorders [4], and so on. In sleep studies for clinical diagnosis and treatment of sleep disturbances, the fundamental step is the identification of an individual's sleep stages.

To evaluate the quality of sleep, it is important to determine how much time was spent in each sleep stage during the night. The sleep stage of the whole night is not a steady-state. Sleepers cyclically passed through five different stages of sleep [6]. The typical length of a complete sleep cycle is about 90 to 110 minutes [7]. In every cycle, there will exist WS, LS, DS,

REM. Meanwhile, stage dynamic transition information reveals significant characteristics of the degree of sleep continuity [9]. The gold standard in this domain is overnight polysomnography (PSG), which include the electrocardiogram (ECG), electroencephalogram (EEG), electromyogram (EMG), electrooculogram (EOG), and respiration signals [10]. Sleepers cyclically passed through five different stages of sleep, i.e., stage 1, stage 2, stage 3, stage 4, and stage 5 or REM sleep. Sleep in stage 1 is light. The eyes move slowly and muscle activity is slow. It is in stage 2 that eyes stop moving and the brain waves become slower. Deep sleep occurs in stages 3 and 4 when no eye movement and muscle activity exist [11]. But the recording of the necessary electrophysiological signals is extensive and complex and the environment of the sleep laboratory, which is unfamiliar to the patient, might lead to distorted results. Meanwhile, manual stage scoring on PSG by domain experts is time-consuming and labor-intensive.

At present, many studies focus on analyzing the HRV features, and some of them are focusing on one stage distribution and transition [12]. Besides, deep learning is also used to classify sleep stage scoring. In this paper, we divided sleep stages into four categories: WS, LS, DS, and REM stage, during which were also called pure sleep stage. Pure sleep stage means in one segment; no other sleep stages are shown. The sleep stage detection algorithm is proposed that used only the heart rate signal, derived from ECG, as a discriminator. This would make it possible for sleep analysis to be performed at home, saving a lot of effort and money. ECG data along with a hypnogram scored by professionals was used from the SHHS database, make it easy to compare the results. For sleep stage classification, we analyzed the sleep stage structure and proposed a TCSM concept and on this foundation, a new framework of sleep stage classification and prediction are presented based on machine learning and pattern recognition technique.

The rest of this paper is organized as follows. Firstly, we review the related work in section 2. Afterward, we present the datasets, ECG signals preprocessing and sleep stage smoothing algorithm in section 3. The details of our solution are also described in section 3. In section 4, Local Cluster Algorithm (LCA) for smoothing sleep stage results and TCSM model results are presented. Finally, we conclude the paper and discuss future work in section 6.

## **1. Related Works**

Many systems using ECG signals have been proposed in the literature for sleep stage classification. In this section, we briefly review the related work which can be grouped into three categories, different biological signals were adopted to classify sleep stage, sleep stage classification based on stage distribution and transition, feature analysis, and extraction based sleep stage classification.

### *1.1 Signals adopted to classify Sleep stage*

Such methods are divided into two categories, i.e. multi-channel and single-channel processing. In the former approach, the combination of various biological signals such as multi-channel EEG [16] [17] signals, electromyogram (EMG) [18] and electrooculogram (EOG), radio measurements without any attached sensors on subjects [19] capturing the temporal progression of sleep, are utilized to extract informative features. In [29] study, actigraphy and heart rate signals are used to classify the sleep stage. HRV and R-R intervals from EEG signals are used to classify the sleep stage [20]. According to the available evidence [22], EEG signals

are almost sufficient for reliable scoring, but the challenge is getting the EEG signals from the subjects. Conversely, HRV data are easy to access from non-contact sensors.

### *1.2 Stage distribution and transition*

The probabilistic properties of sleep stage sequences and transitions are used to improve the performance of sleep stage detection using cardiorespiratory features [12]. The classifier, based on conditional random fields, achieved an average accuracy of 87.38%. [13] leverages both signal and stage transition features of human sleep for automatic identification of sleep stages. Random forest classifier is trained using thirty EEG signal features. The correction rules with a Markov model were constructed based on stage transition feature, importing the continuity property of sleep and characteristic of a sleep stage transition. In [14], the authors applied a smoothing process on classified results to improve their classifier's performance, which considers the continuity of a subject's sleep. It considered the temporal contextual information and improved the continuity of a subject's sleep stage scoring results. Some correction rules were constructed according to the relationship among the current epoch, prior-epoch, and posterior-epoch. Similarly, correction rules are used in paper [14]. However, their correction rules only contained a few manually defined rules and need more prior knowledge of the sleep study and characteristics of the sleep dataset. [23] previously showed that human sleep-wake stage distributions exhibit multi-exponential dynamics, which are fragmented by obstructive sleep apnea (OSA), suggesting that Markov models may be a useful method to quantify architecture in health and disease. Some works used a hidden Markov model [24] to classify the sleep stage.

### *1.3 Feature analysis and extraction*

In [28], the author proposed a sleep stage classification system based on noise-reduced fractal property of heart rate variability, sensitivity is 77% that distinguish wake stage from the other sleep stages and 72% deep sleep stage from light sleep; In [29] study, actigraphy and heart rate signals were used to classify sleep stage. Body movement indices derived from actigraphy data and autonomic functional indices from heart rate variability is used for discriminating between non-REM sleep and waking/REM sleep at 76.9% sensitivity and 74.5% specificity and between REM sleep and waking at 77.2% sensitivity and 72.3% specificity. In the [30] study, REM sleep was determined with an adaptive threshold applied to the acquired feature. Only the heart rate signal feature: LF/HF ratio, the relative peak frequency power in the HF band, the variability within the HF band, derived from electrocardiogram (ECG), to perform classification [31]. The authors tried to detect sleep and wakefulness using a combination of ECG, actigraphy, and respiratory signals [32]. With an automated Bayesian classifier, they achieved an accuracy of 86.8%. In the study of [33], the author grouped the 6 consecutive windows into one epoch (length, 3 min), whose stage is defined as that of the majority among 6 windows, i.e., if the stage of  $\geq 4$  windows were the same, it was chosen as the stage of an epoch. If there was no majority stage among 6 windows, the epoch was defined as a transitional stage. Moreover, the existing work only focused on analyzing signal features without considering the dynamic stage transition information.

The main contributions of this paper are listed as follows.

(1) Sleep structure characterization and modeling. Sleep is a gradual evolutionary process. Sleeping overnight usually takes three to five sleep cycles, and when transitioning from one sleep stage to another, there must be a transition, usually called sleep stage transition; Accordingly, if there is no stage transition, called pure sleep in this paper, and the local cluster algorithm (LCA) method is used to smooth the sleep stage to obtain pure sleep and stage transition.

(2) HRV feature extraction and selection. This paper extracts ECG data from PSG data, and then uses the heartbeat interval detection method to detect the R wave to calculate the RR interval and perform false and missed detection. Based on the HRV data and sleep structure, for the sleep staging requirements and the TCSM model, the time domain, frequency domain and nonlinear domain characteristics of the HRV are utilized respectively, and the feature selection is based on the max-relevance min-redundancy criteria.

(3) Verification of Sleep staging method based on the TCSM model. Based on the actual sensory data of the larger experimental data (400 people), the existing sleep staging method is used as a reference, and TCSM is used as the comparison verification.

This paper studies the duration of stage transitions at different time scales and the optimal windowing criteria for pure sleep. In the experiment, the recognition of pure sleep and stage transition, the multi-classification of pure sleep, and the verification analysis of the coarse-fine grained classification of stage transition are carried out in multi-angle. The Hidden Markov Model is used to describe the objective law of natural stage transition in the sleep stage, improving the accuracy of sleep stage classification. The comparison experiments of large-scale datasets verified the accuracy and effectiveness of proposed method.

## 2. Material and Methods

In this section, we introduce the dataset, ECG and sleep stage preprocessing, the system flowchart and classification. The method consists of three phases: processing, feature selection, and classification. The flowchart of the whole architecture is shown in Fig.1.

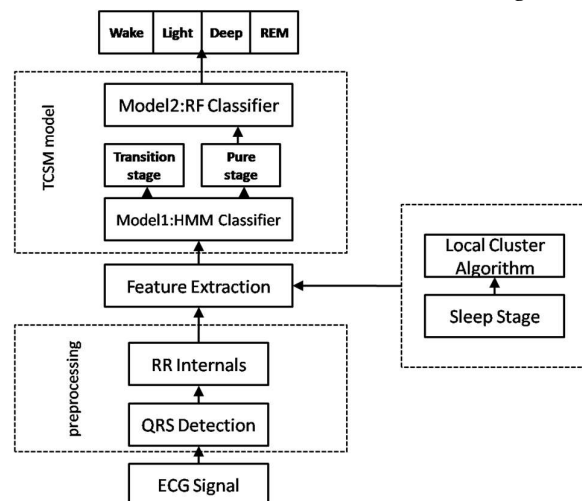


Figure.1 The flowchart of the proposed method. Consisting of three sections: data preprocessing, sleep stage LCA algorithm and TCSM model.

### 2.1 Data Collection

There are two standards commonly used to define sleep stages: The Rechtschaffen and Kales (R&K) [34], and that developed by the American Academy of Sleep Medicine (AASM) [35]. The AASM standard adopted in this paper classifies sleep into 4 different stages. To develop and validate the proposed method, a set of PSG recordings from real subjects are used. The recordings are taken from the Sleep Heart Health Study (SHHS). This database emerged from a multi-center cohort study to determine cardiovascular and other consequences of sleep-disordered breathing. Details about the design of the SHHS study can be found in [36]. Each recording comes with the annotations of physicians' off-line scorings following the R&K procedure [37]. The Sleep Heart Health Study (SHHS) is a multi-center cohort study [36], initiated by the American National Heart Lung and Blood Institute to determine whether sleep-disordered breathing is associated with a higher risk of various cardiovascular diseases. The study includes two rounds of PSG recordings. We use only the first round (SHHS-1) because all records have the same sampling rate (125 Hz), contrary to the second round where records can be sampled at 125 or 128 Hz. Dataset SHHS-1 contains 5793 PSG records.

Here we adopted 400 subjects (male 161 subjects, female 239 subjects) and all of them were healthy. And the age is between 40 and 59, the mean age of all the subjects is 53.44. In the PSG records, for most subjects, along the 'wake' period before the subject goes to sleep and another after he or she wakes up is observed [38]. To preserve all the sleep data and HRV signals, the WS is not trimmed at the beginning of sleep. The hypnogram, which reports the different sleep stages across the sleep time, was obtained in TABLE I.

TABLE 1. DESCRIPTION OF THE DATASET

	Sleep Stage				
	<i>Wake</i>	<i>LS</i>	<i>DS</i>	<i>REM</i>	<i>Total</i>
Total epoch(30s)	93995	186226	55136	64307	399664
Equivalent Hour	783.3	1551.9	459.5	535.9	3330.5

## 2.2 Data preprocessing

### 2.2.1 The preprocessing of ECG signals

We use the ECG deal package [39] to deal with the ECG signal and get the R wave index. The R wave detection results are as the Fig.2. According to the definition of HRV, the RR interval is calculated by the subtraction of later R wave and its adjacent front one. The sequences of the RR interval were HRV. Before the extraction of the frequency measures, a preprocessing was carried out on the RR interval based on the fact that the spectral analysis cannot be affected only on regularly sampled signals [40]. This preprocessing is presented by resampling the data at a frequency of 4 Hz using an interpolation method developed by Berger et al [41].

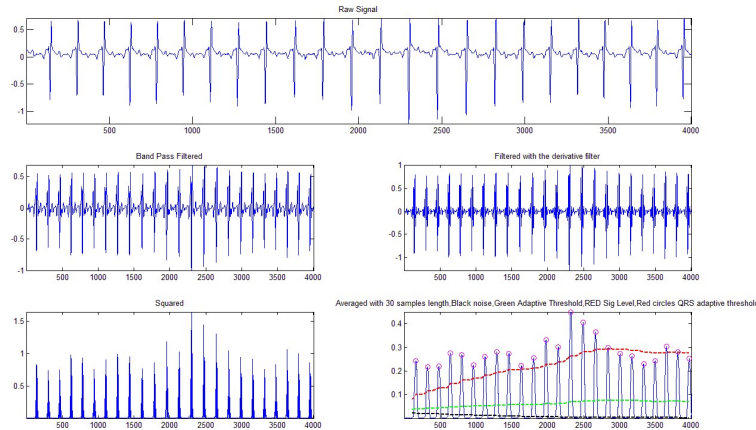


Figure.2 R wave detection results.

## 2.2.2 Sleep Stage Structure Characterization

### Pure Sleep Stage Analysis:

Fig.3 is W, LS, DS, REM percent and duration time. When people sleep at night, LS accounts for the largest proportion, about 45%, and sleep less than 4 hours. The proportion of the awakening period is relatively large before a person falls into sleep. If the waking up before going to sleep is removed, there will be very few awake samples, so in this paper, W is not removed; REM duration time is about 15%, and the duration time is 1.2 hours.

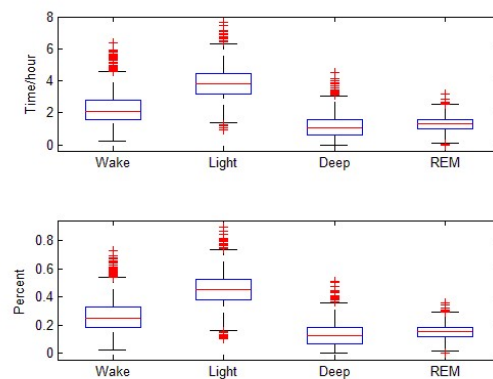


Figure.3 different pure sleep stage duration and percent in all pure sleep stage

Fig.4 shows the sleep stage of 400 people. The sleep time of one night is set to 9 hours, and the proportion of different pure periods of sleep is counted in each hour. From the beginning of sleep, as sleep deepens, the proportion of each sleep stage is also different. The proportion of waking or awakening period is higher in the first hour and the last hour; the proportion of LS in the whole night's sleep is high. Deep sleep accounted for a higher proportion in the first half of the night, that is, deep sleep was mainly in the first half of the night; while the REM period was difficult to occur in the first half of the night and improved in the latter half of the night. It can be seen from the proportion of different sleep stages that the probability of occurrence of

the sleep cycle was different in different sleep stages, and the sideshow that sleep has its natural laws and structures. This figure sets the duration to 9 hours and covers 90% of the experimenter's data. Some people sleep for more than 9 hours, and the length of time is not within the statistical range. The default data for more than 9 hours is invalid. In this figure, there is no correlation between the sleep structure and the number of sleep hours according to the proportion of different sleep stages in each hour. Therefore, in the next section, the sleep cycle analysis and the proportion of pure sleep in different cycles were introduced.

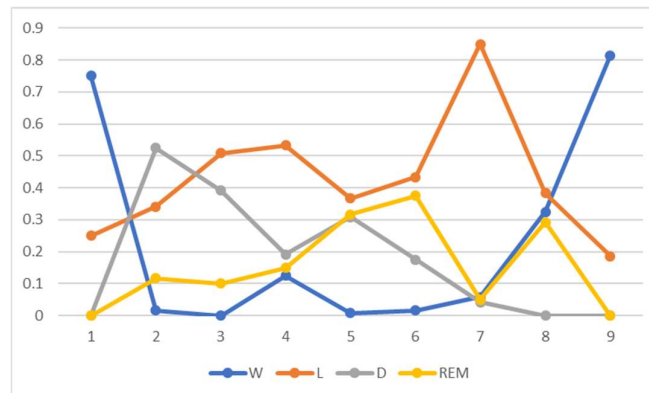


Figure.4 Sleep stage percent in different hour in whole night

#### Sleep stage cycle analysis:

The sleep cycle referred to the existence of a biological rhythm in a person's sleep state, that is, from the beginning of sleep to the end of sleep, there is a cycle in the middle. AASM divides sleep into five cycles, each of which lasts for about 90-110 minutes. The sleep cycle is defined as the end of a cycle with the REM period. NREM and REM alternately appear, alternating once called a sleep cycle. The number of occurrences of the cycle from awake or light sleep to REM is the number of cycles in sleep.

From the analysis, 80% of people have 4-5 cycles of sleep at night, some have three sleep cycles, and a few have more than 5. In this paper, the sleep cycle of most people is 4 sleep cycles, so the sleep cycle is set 4 and is analyzed, and the data is shown in Fig.5. From the Fig.6, we can see that in different sleep cycles, the sleep stage duration is different. After entered the sleep state, the proportion of the W period is less than 10%. As LS deepens, the proportion increases; while LS occupied a major part of sleep. At the same time, it can be found that deep sleep mainly occurs in the first half of the night, and deep sleep decreases in the latter half of the night; the REM phase mainly occurs in the latter half of the night. As can be seen, the awakening period is 10 minutes in the period after falling asleep, LS is 50 minutes, DS is 10-30 minutes, and REM is 15-25 minutes.



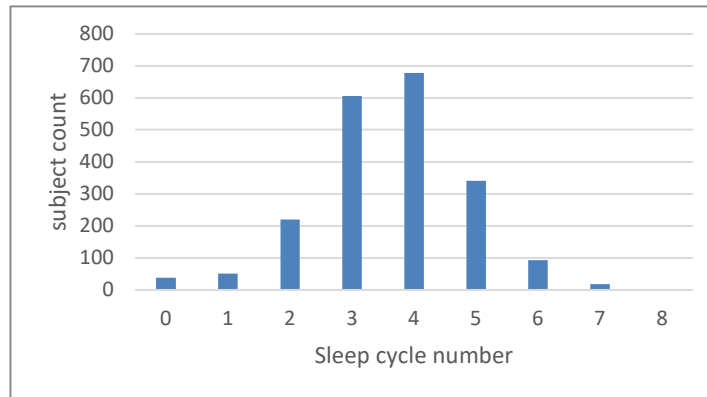


Figure.5 Sleep cycles analysis

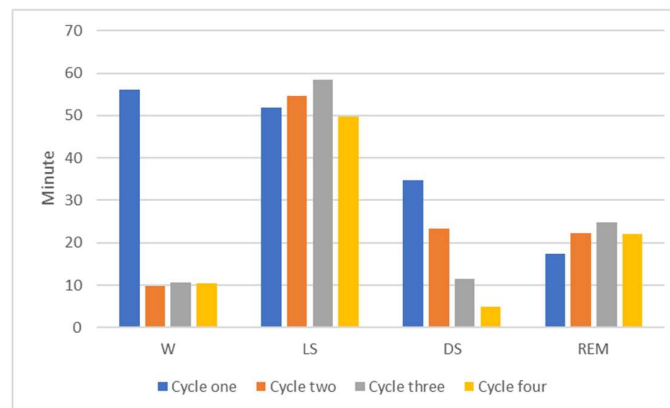


Figure.6 Time spent in different sleep cycles

### Sleep Stage transition analysis:

It can be seen from the Fig.7 that the transition probability from the sleep stage P1 to the sleep stage P1 is the largest, indicating that the person continues for a while during sleep and does not change all the time. From L to other sleep stages, L has a high probability of transition to L, and some people have a higher probability of reaching D, indicating that the probability of going from shallow sleep to deep sleep is also higher. In the deep sleep transition, it can be seen that part of it is from deep sleep to deep sleep, and come from deep sleep to light sleep. In actual sleep, the transition probability from deep sleep to light sleep is also relatively large. At the same time, the probability of going from deep sleep to rapid eye movement is extremely small, that is, the previous sleep entering the REM phase, and the minimum may be deep. In the stage transition of REM, the stage transition of RR is the largest, and at the same time, part of it is from REM to the awakening period. From the figure, the stage transition of a person under a normal sleep state can be generally seen, which lays a theoretical foundation for the use of stage transition for sleep stage classification.



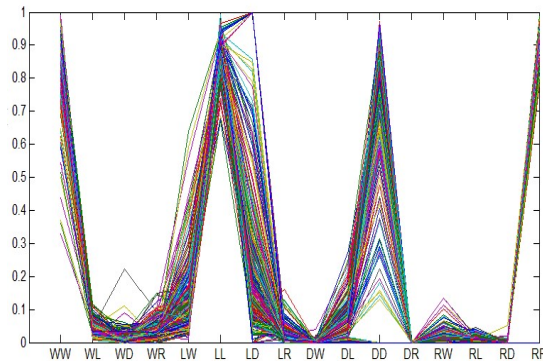


Figure.7 Sleep stage transition probability

TABLE I. Sleep stage transition probability in whole night

Sleep stage	W	LS	DS	REM
W	0.9042	0.0434	0.0143	0.0381
LS	0.1016	0.6920	0.1880	0.0183
DS	0.0013	0.0783	0.9203	0.0001
REM	0.0143	0.0151	0.0013	0.9693

In this paper, the sleep cycle is set to four, to facilitate the description of the sleep structure, and in line with the sleep distribution of most people. TABLE.3 shows the probability distribution of the six-stage transitions in different cycle periods, and the data is visualized as shown in Fig.8. We can see that the mode of stage transition changes from the first to the fourth cycle. In the first and second cycles, the model did not change significantly. There was no big change in WL, LW, LD, DL after falling asleep to the second sleep cycle, because the probability of REM occurring in the first half of the night is small. Therefore, the proportion of RL and WR is small, and there is no large change; in the third and fourth cycles, LD is reduced, that is, the transition probability from light sleep to deep sleep is reduced, and DL is reduced, that is, deep sleep to light sleep. The transition probability is reduced because the proportion of deep sleep is reduced in the third cycle. In the fourth cycle, WL increased because the proportion of deep sleep decreased, and the proportion of awakening and shallow sleep increased. It can be seen throughout the night's sleep that as sleep progresses, RL and WR increased. This is because as the sleep progresses, the proportion of REM increases gradually, so the probability of stage transition increases. WL is the first drop in sleep throughout the night, gradually rising, just starting to fall asleep, waking up to occupy the main, with sleep deeper, W decreases, other pure sleep increases. At the end of the sleep, the awakening period increases. LW did not change much during the whole night's sleep. LW is from asleep to awakening. Here, W is the awakening period, and it is the state of waking without consciousness. The probability of stage transition in the different sleep cycles is different. If only the sleep transition probability of the whole night is used for sleep staging, the accuracy of sleep staging will be reduced. Therefore, in the paper, sleep structure is used in sleep staging, and different nights of sleep are used.

TABLE II. Sleep stage transition probability in the whole night in different sleep cycles

Cycle No.	WL	LW	LD	DL	RL	WR
Cycle 1	0.2455	0.1432	0.2972	0.2659	0.02317	0.02496
Cycle 2	0.2144	0.1536	0.2999	0.2717	0.02715	0.02579
Cycle 3	0.2856	0.2052	0.2123	0.1959	0.0415	0.04604
Cycle 4	0.3562	0.2611	0.1433	0.1316	0.0415	0.05429

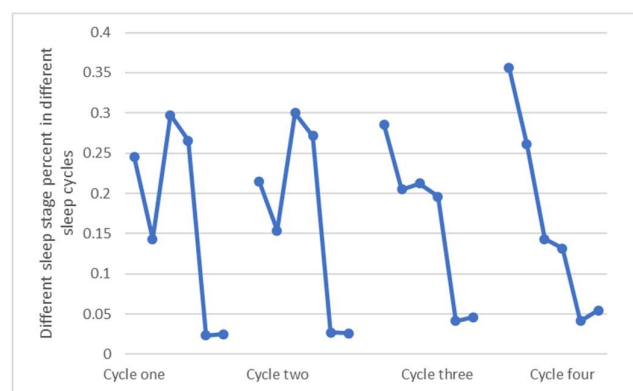


Figure.8 Sleep stage transition probability in different sleep cycles

### 2.2.3 The pre-smooth of sleep stage (LCA algorithm)

Sleep is a continuous process that does not mutate, but some weird stages may exist caused by some external factors, such as diseases and devices. Smoothing rule shows the stable sleep duration and the property of sleep continuity; in other words, in a stable stage of sleep, it is unlikely that the stage suddenly changed [13]. Consequently, we smoothed the data. The sleep stage smoothing rules are based on the local cluster Algorithm (LCA).

The stage transition was ambiguous and the time of stage transition was not very clear. Consequently, we defined the stage transition time domain as the  $T$  minutes which mean that the former  $T/2$  is one sleep stage, and the later  $T/2$  is another sleep stage. For example, if  $T$  is 5, the former 2.5 minutes is in LS, the latter 2.5 minutes in another maybe in WS, DS or REM. The next will explain how to pre-smooth the sleep stage based on the LCA algorithm and  $T$  is set 5. Firstly, the LCA obtains every stage and corresponding count. Secondly, LCA sets domain 2, if the corresponding count less than the domain, the stage will be changed to the adjacent stage whose count is larger. After that, the LCA sets domain 3, the similar processing as the above until the domain equal to  $T$ . Finally, the minimum stage's corresponding count is  $T$ .

As seen in Fig.3, the initial sleep stage is 00000011212222332220022211112, which corresponding count all are 1. After merging the same value, the stage is 01212320212, and the corresponding count is 62114232341 which domain is 1. Next, the domain is set 2, which means that only the count less than domain, the corresponding stage will be smooth. The smoothing rules are as follows: if the former stage's count larger than the latter, the current stage will be

changed to the former; If the latter stage’s count larger than the former, the current stage will be changed to the latter. The same operations of the domain in 3,4,5.

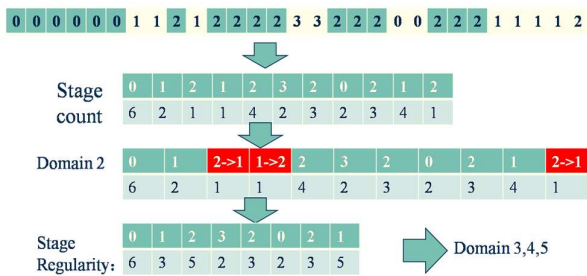


Figure.9 The LCA method for sleep stage smooth.

The LCA pseudocode is as the Algorithm 1. After LCA algorithm, the results are as TABLE IV.

TABLE IV. LCA ALGORITHM RESULTS

Stage	0	2	1
Count	6	18	5

Algorithm 1. Stage Local Cluster Algorithm

Input: sleep Stage, domain

Output: slice Stage and count

FOR s=1: length (sleep Stage)

Count(s)=1;

END FOR

INIT: new Stage= [];

FOR s= 2: length (sleep Stage)-1

IF count(s)>= domain THEN

Continue;

ELSE IF count(s) <domain THEN

IF count (i-1)>=count(i+1) THEN

New stage(i)=new stage(i-1);

ELSE new stage(i)=new stage(i+1);

END IF

END IF

END FOR

IF count (1) <domain THEN

New stage (1) =new stage (2);

END IF

IF count (length (new stage)) < domain THEN

New stage (length (new stage))=new stage(length(new stage)-1);

END IF

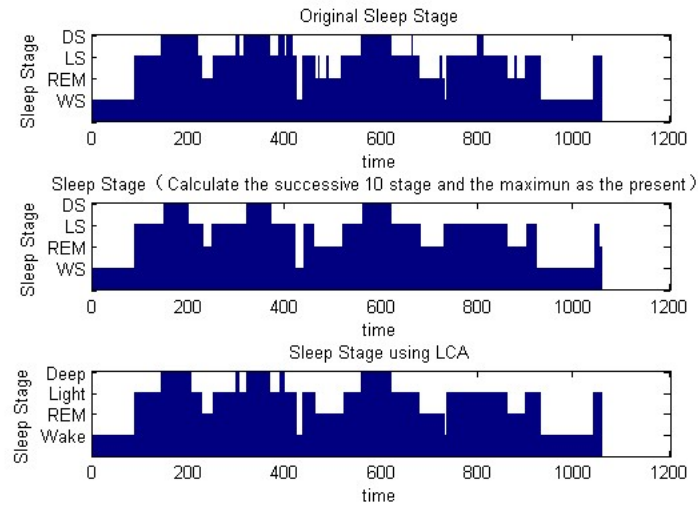


Figure.10 LCA results compared to a Traditional smooth method

### 2.3 Feature extraction and selection

#### 2.3.1 TCSM model

Complete sleep cycle is experience three to five cycles one night, with three or four small loops in every cycle. From the Fig.11, we can obviously find the whole night 4 sleep stage cycles. And each cycle, in the inner cycle, there are 3 sleep stages. Every cycle ends with a REM sleep stage. Every outer cycle may include LS, DS and REM stage. We analyzed the sleep stage structure, and proposed a TCSM concept and on this foundation a new framework of sleep stage classification and prediction are presented based on machine learning and pattern recognition technique.

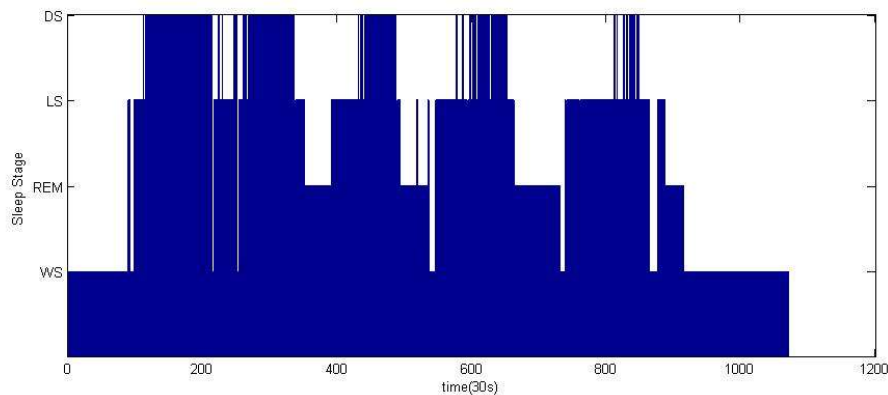


Figure.11 The whole night hypnogram.

#### 2.3.2 Feature extraction and feature reduction

According to the obtained HRV, there are total of 18 features in the time domain, frequency domain and nonlinear domain features are calculated. The features in the time domain are mean, variable, max, min, sdn, sdsd, pnn50, cv; The frequency domain ones are vlf, lf, hf, vhf, tlf, v; Finally, the nonlinear domain features are s-entropy, renyi entropy, Shannon-entropy.

All features used to recognize the pure sleep stage and stage transition are shown as TABLE IV.

TABLE III. FEATURES OF HRV FOR ESTIMATING PURE STAGE AND STAGE TRANSITION

Feature	Description (all feature calculated in a window)
mean	Mean of RRI
var	Variance of all RR intervals
sdnn	Standard deviation of all RR intervals
sdsd	Difference Standard deviation of differences between adjacent RR intervals
rmssd	Square root of the mean of the square of differences between adjacent RR intervals
pnn50	number of adjacent intervals greater than 50ms accounted for the proportion of all RR intervals
cv	Coefficient of variation of differences between adjacent RR intervals
vlf	Total power between 0.015 and 0.04Hz
lf	Total power between 0.04 and 0.15 Hz
hf	Total power between 0.15 and 0.4Hz
vhf	Total power between 0.4 and 0.5 Hz
tlh	Total power between 0.015 and 0.5 Hz
v	Ratio of low to high-frequency power
s_entropy	Sample entropy
renyi_entropy	Renyi entropy
shannon_entropy	Shannon entropy

### 2.3.3 Classification

Firstly, classify the three or four outer cycle and then classify the inner 4 sleep cycle. In the first step, 16HRV features, including temporal, frequency and non-linear features were used to train model1 and model2.

The TCSM model first analyzes the stage transition T and the pure sleep P, which is the first layer of the TCSM model. In the second layer of the model, we have two methods to validate the model proposed in the experiment. The first method is based on course and fine-grained classification of a stage transition. The fine-grained classification is 12 classifications of WL, LW, WD, DW, WR, RW, LD, DL, LR, RL, DR, RD; Coarse sleep stage classification is of WL, LW, LD, DL, LR, RL, DR, RD classification. The second method is based on the classification of pure sleep, including segmentation classification and segmentation classification of pure sleep, and directly classified pure sleep.

The stage transition analysis was performed under different time scales D (Domain), the time scale was 1-10 minutes, and the analysis of sleep stage transition and pure sleep were classified under D. The wake of the sleep stage, the light sleep, the deep sleep and the rapid eye movement period were divided into pure sleep, and the transition of different sleep stages were converted into a stage transition, and the second classification were performed.

1. Use the sliding window for segmentation, the purpose was to build the sample, use the classifier to classified which was pure sleep, which was stage transitions, the time scale was 1-10 minutes respectively; as shown in Fig.12.

2. Further subdivide the detected pure sleep into pieces at different times (1-10 minutes)
3. Directly classified the detected sections, this section did not equal, some were 10 minutes, some were half an hour, some were a few minutes(Fig.12); large did not divided Pure sleep is called a slice, which is called a piece that is cut into fixed lengths(Fig.13);
4. Classified the stage transitions. Based on the first step, the stage transition was detected, and the stage transition was further subdivided into 6 classifications or 12 classifications (Fig.14).

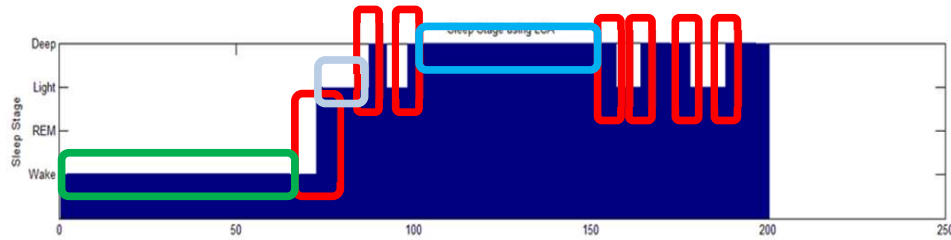


Figure.12 Pure sleep stage classification in slices

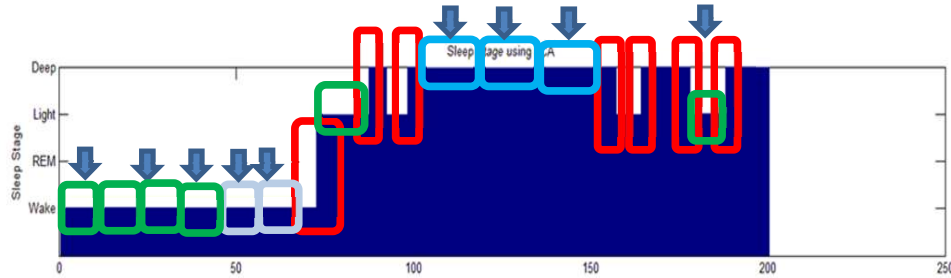


Figure.13 Pure sleep stage classification in pieces

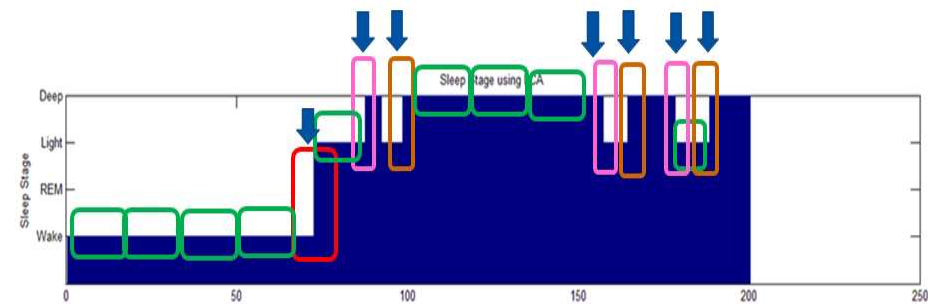


Figure.14 Coarse and fine grained stage transition classification

### 3 Experimental Results and Performance Evaluation

#### 3.1 Performance evaluation Measures

In this section, we presented the evaluation results of the proposed sleep stage classification framework. Four criteria (measures) were used to evaluate the performance of the proposed sleep stage classification system. These criteria were the classification accuracy, sensitivity, specificity, and confusion matrix. The classification performance was measured which are expressed as follows:

$$Accuracy(Ac) = \frac{TP + TN}{TP + FN + TN + FP} (\%)$$

$$Sensitivity(Sn) = \frac{TP}{TP + FN} (\%)$$

$$Specificity(Sp) = \frac{TN}{TN + FP}$$

$$Precision(Pr) = \frac{TP}{TP + FP} (\%)$$

where TP, TN, FP, and FN denote true positives, true negatives, false positives, and false negatives, respectively.

To set the system parameters and to evaluate the system performance, we used a k-fold cross-validation process. To do this, the dataset divided randomly into k equal-sized subsets. At each fold, the (k-1) subsets were used as the training and validating data and 1 subset was used for testing. This process was repeated k times (the number of folds), with each subset used exactly once for testing. We also used a k-fold subject cross-validation process, where k was the number of subjects in the dataset. At each fold, the data of one subject were used for testing, while the data of other subjects were used as the training and validating data. This process was repeated k times. The results of the k folds were averaged and reported as the system performance.

### 3.2 Results

In the experiment, we firstly classified the stage transition and pure stage and then classified the four pure sleep stages. We choose the Random Forest (RF) classifier to classify the different classes. RF classifier, proposed by Breiman [43], was a classification and prediction models, belongs to a kind of combination classifier. It overcomes the problem of easy to converge to the local optimal solution and the problem of excessive fitting which may occur in the single decision tree model and can be a very good tolerance data set of noise and outliers, suitable for processing the unbalanced data sets and high dimension data. In the first step, according to the stage transition definition, we constructed the train set. If the restored stage, we choose the last 2/T stages and the first 2/T stages as the stage transition, and calculated the feature of the corresponding HRV.

Due to the ambiguity of stage transition during sleep in different people, it is difficult to directly set the stage transition to a fixed length [43]. Therefore, in the experiment, this paper classifies the stage transition and pure sleep at different time scales, and uses the detection rate of the stage transition (that is, the degree of effect, which evaluates the proportion of all negative classes detected) is appropriate in the case of evaluating how long the time scale is set. As shown in Fig.15, in 1-10 minutes, as the time scale increases, the detection rate of the stage transition increases; from 5-10 minutes, the detection rate decreases. Too long or too short is a difficult classification during sleep. It can be seen from the experimental results that the stage transition is set to 4 minutes.



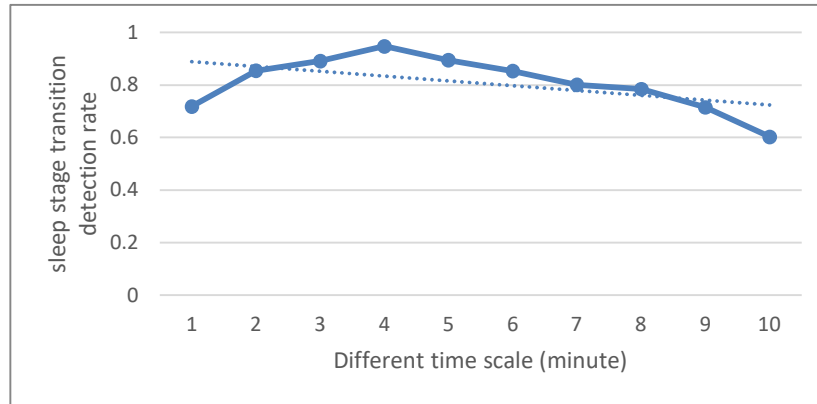


Figure.15 Detection rate of different time scale for pure sleep stage and stage transition

In this experiment, the AdaBoost classifier was used, and the NB and RB classifiers were used for classification. The experimental results are shown in Figure 5-7. As can be seen from the figure, using different classifiers, the experimental results were best when the time scale is set to 4 minutes.

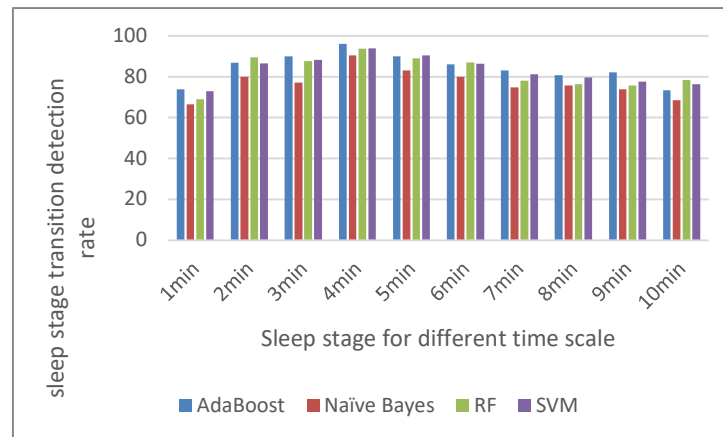


Figure.16 Sleep stage transition using different classifier in a different time scale

Based on the first step, we divide the stage transition and pure sleep. In this step, we classified the classified pure sleep. For the classification of pure sleep, we proposed two methods for classification. First, based on pure sleep, further segmentation is divided into equal-length slices, called sleep segments, and then classified into equal-length segments. In the experiment, by analyzing and comparing the classification results of the fragments at different time scales, it was determined at which time scale the segmentation was performed, and then the classification effect was better, and the classification efficiency was high. Pure sleep was further divided into time slices of equal duration and then classified. Fragmentation was performed from pure sleep at different time scales. The second was to directly classified the classified pure sleep. The length of this pure sleep was different, and it was the unequal length HRV feature. Therefore, when extracting features, we should classify the features that are not related to length. The equal length division of pure sleep, followed by the 4 classifications of pure sleep, as shown in Figure 5-8, is the classification of pure sleep using NB, RF and HMM. Pure sleep was divided into different time scales of 1 to 10 minutes and then

classified. It can be seen from the figure that pure sleep was sliced for 5 minutes, then classified, using three classifiers. The highest accuracy rate is 86.618% with HMM.

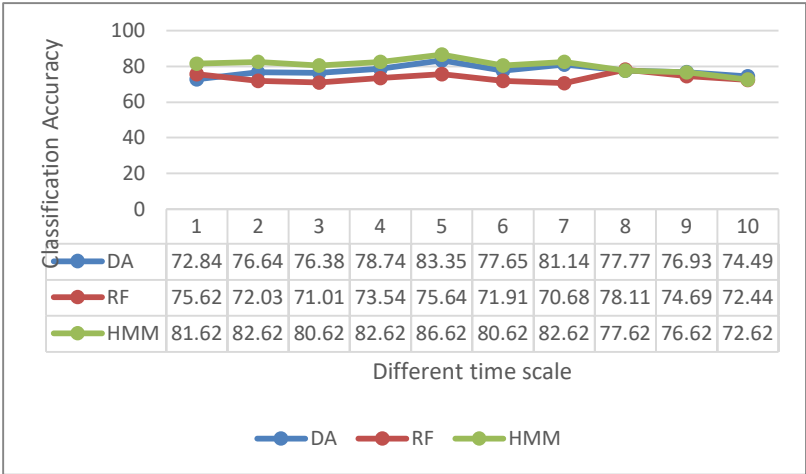


Figure.17 Pure sleep stage classification in pieces using different classifier in a different time scale

After the stage transition and the pure sleep classification, further, the classified pure sleep is directly classified, and the classification efficiency can be improved without fragmentation. Considering the difference in the length of the sections, this paper proposed new features for classification. In the classification of the subsection, considering the non-equal length of each pure sleep, we took the sleep stage of the sleep and the sleep cycle of the sleep stage as characteristics and added it to the traditional classification model to carry out the model training. In practical applications, by recognizing the stage transition, the sleep was divided into pure sleep of unequal length, and one sample loop is used for recognition. In the HMM model, the training of the model was performed using the sleep transition probabilities of different sleep stages. In this paper, three classifiers are used for the classification of pure sleep. From the classification results, the classification effect using HMM classifier was the best, 81.6%.

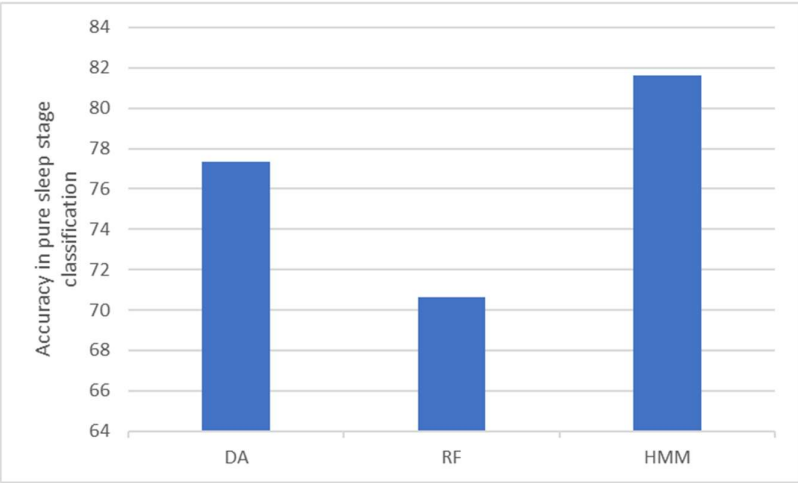


Figure.18 Pure sleep stage classification in slices using different classifier in a different time scale

In this paper, in the fine-grained classification of stage transition, and then the 10-fold cross-validation was performed, and the experimental results are shown in Fig19. It can be seen from the experimental results that for the 12 classifications, since the difference of the 12 stage transitions was relatively small, in the classification, the effect was not very satisfactory. Therefore, this paper proposed a coarse-grained classification of stage transitions, that is, 6 classifications of stage transitions. For this experiment, this paper chooses to use DA and RF classifiers to train the model. The experimental results are shown in Fig20. As can be seen from the figure, the classification accuracy of the DA classifier is about 60, and the classification effect of the RF classifier is not as good as DA. From the time scale analysis, it can be seen that when the time scale was set to 6 minutes, the classification effect was better, and consistent conclusions were drawn from the two classifiers. At the same time, the time scale was 6 minutes, and the time difference between the traditional sleep stage and 5 minutes was not much different. Therefore, in the experiment, this paper used the 6-minute

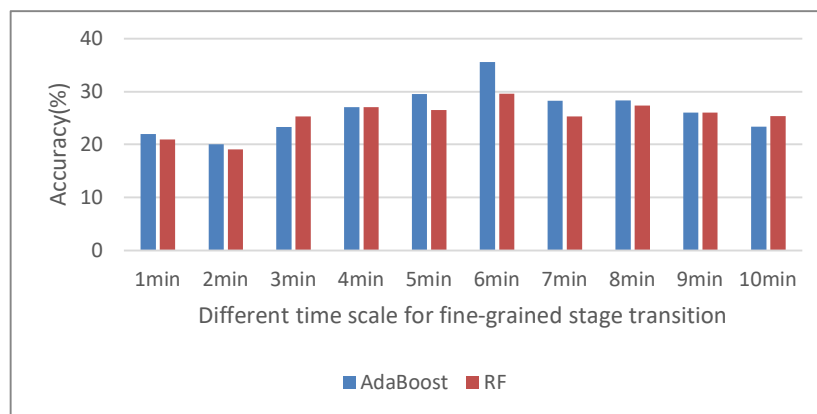


Figure.19 Fine-grained classification accuracy of stage transition at different time scales

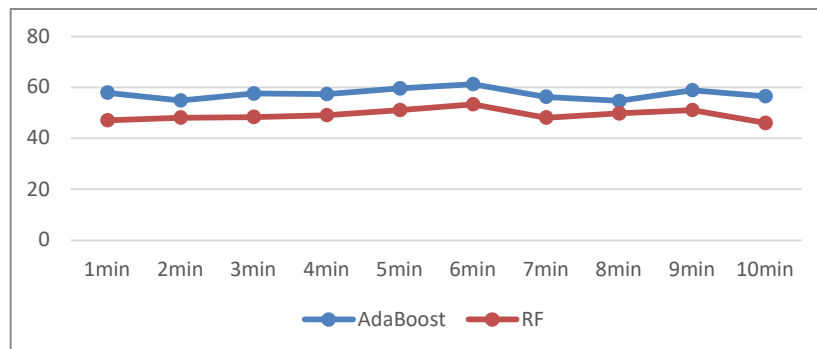


Figure.20 Coarse-grained classification accuracy of stage transition at different time scales

To verify the correctness and effectiveness of the proposed TCSM model, this paper used the traditional sleep staging method to identify the sleep stage. The traditional sleep stage identification method was divided the HRV for 5 minutes, and then count which one of the 10 sleep stages was more frequent in 5 minutes, and then select the sleep stage corresponding to the 5-minute HRV. This article uses data from 400 people for verification. The classification results of traditional sleep stages are shown in Fig.21. As can be seen from the figure, the traditional method was used for sleep staging, and the classification accuracy was about 75%. At the same time, by setting different time scales, the paper analyzed the HRV segmentation, the scale was set to 6 minutes, and the classification effect was better.

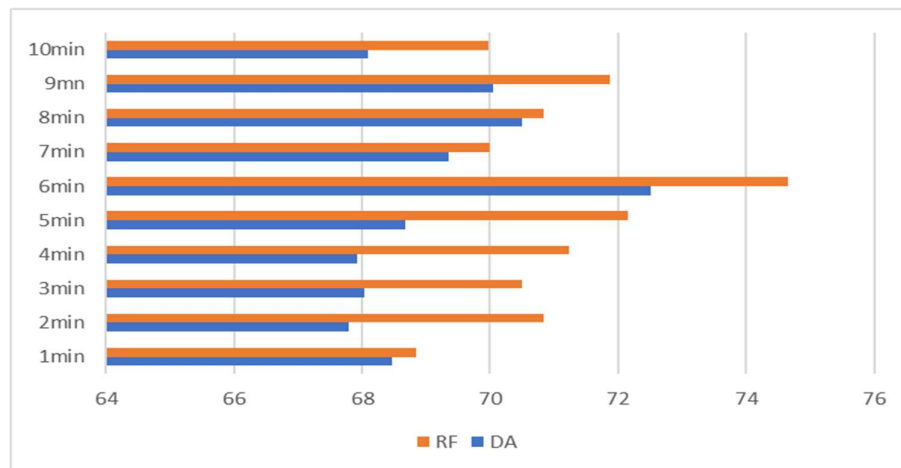


Figure.21 Traditional method classification accuracy at different time scales

This article compared the traditional method with TCSM and used 34 people's data to illustrate. From the data, it can be seen that there are 30 people whose classification effect is higher than the traditional classification method, more than 7%. The model presented in this paper is valid and feasible, as shown in Fig.22.

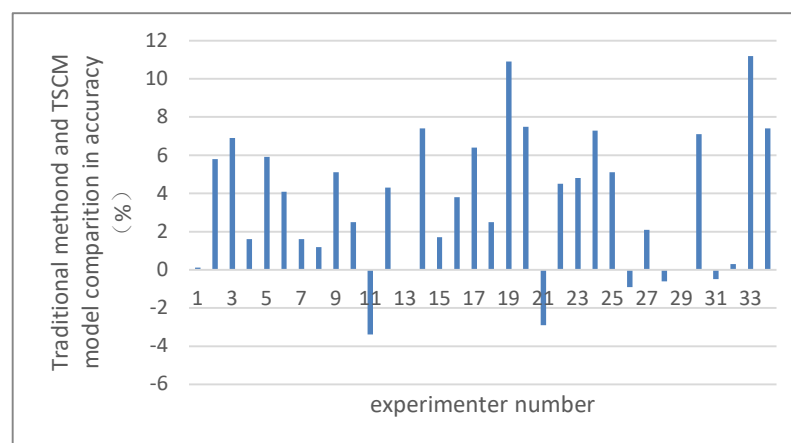


Figure.22 TCSM and Traditional method comparison in 35 subjects

TCSM is better than traditional ones in terms of classification accuracy and accuracy, as well as classification efficiency. During the experiment, the accuracy of the classification of pure sleep segmentation was 79.7%, and the recall rate is 79.6%. The experimental results shown that the research method is different from the traditional long-segmentation of sleep and then the staging method. The sleep staging method based on TCSM is based on the two-layer sleep structure model proposed by the sleep structure for the first time. The classification of non-equal length pure sleep was directly classified, and the classification efficiency was high.

#### 4 Discussion and Conclusion

The paper presents a new framework of sleep stage classification based on the proposed concept TCSM. In this study, we demonstrated the performance of sleep stage classification by the indices obtained from ECG signals. We studied 4340 epochs (length 5 min) of PSG data in 800 subjects. Our results indicated that TCSM model and two steps sleep stage classification was useful for discriminating stage transitions from pure sleep stage and for discriminating 4 pure sleep stage(Wake, Light sleep, Deep sleep, REM); Although a plenty of earlier studies

have reported various algorithms and methods for sleep stage classification by ECG data, this study is unique from them in (1) providing TCSM model, (2) proposing two-step approaches separately discriminating stage transitions from pure sleep stage and for discriminating 4 pure sleep stage (WS, LS, DS, REM), and (3) examining models with a large sample size. Our observations seem to reflect the performance of this approach when applied to real-world data from a wide range of people.

The study has some limitations:

- 1) Only signals from a single ECG channel were used in classification,
- 2) The pure sleep stage and stage transition classification were not satisfactory,
- 3) We have not considered the effects of diseases, such as narcolepsy, insomnia and sleep disorder, on their influences on classification.

In future work, we will focus on sleep stage scoring using multiple PSG signals, such as EOG and EMG, to improve the performance of our approach. Moreover, we will design more elegant algorithms for analyzing and classifying sleep data. Additionally, we will apply our approach to more datasets, especially to the datasets of different sleep disease studies, which can be obtained from the NSRR website [44]. In future work, the inner cycle will be classified directly rather than slicing into 5 minutes.

## Acknowledgment

This work is supported by the National Natural Science Foundation of China (No. 61332013).

## References

1. Ellenbogen J M, Payne J D, Stickgold R. The role of sleep in declarative memory consolidation: passive, permissive, active or none? [J]. *Current Opinion in Neurobiology*, 2006, 16(6):716-722.
2. Liu F, Zhou X, Wang Z, et al. Identifying Obstructive Sleep Apnea by Exploiting Fine-Grained BCG Features Based on Event Phase Segmentation[C]// *IEEE, International Conference on Bioinformatics and Bioengineering*. IEEE, 2016:293-300.
3. Kwon H, Yoon H, Jung D, et al. Heart rate variability in patients with major depressive disorder and healthy controls during non-REM sleep and REM sleep[J]. *Conf Proc IEEE Eng Med Biol Soc*, 2017:2312-2315.
4. Sors A, Bonnet S, Mirek S, et al. A convolutional neural network for sleep stage scoring from raw single-channel EEG[J]. *Biomedical Signal Processing & Control*, 2018, 42:107-114.
5. H. G. Jo, J. Y. Park, C. K. Lee, S. K. An, and S. K. Yoo, "Geneticfuzzy classifier for sleep stage identification," *Computers in Biology and Medicine*, vol. 40, no. 7, pp. 629-634, 2010
6. Memar P, Faradji F. A Novel Multi-Class EEG-Based Sleep Stage Classification System[J]. *IEEE Transactions on Neural Systems & Rehabilitation Engineering A Publication of the IEEE Engineering in Medicine & Biology Society*, 2018, 26(1):84.
7. P. Tian et al., "A hierarchical classification method for automatic sleep scoring using multiscale entropy features and proportion information of sleep architecture," *Biocybern. Biomed. Eng.*, vol. 37, no. 2, pp. 263-271, 2017.
8. M. A. Carskadon and W. C. Dement, "Normal human sleep: An overview," in *Principles and Practice of Sleep Medicine*, M. Kryger, T. Roth, and W. C. Dement, Eds., 6th ed. Amsterdam, The Netherlands: Elsevier, 2017, pp. 15-24. [Online]. Available: <https://doi.org/10.1016/B978-0-323-24288-2.00002-7>
9. A. Laffan, B. Caffo, B. J. Swihart, and N. M. Punjabi, "Utility of sleepstage transitions in assessing sleep continuity," *Sleep*, vol. 33, no. 12, p.1681, 2010.
10. J. V. Holland, W. C. Dement, and D. M. Raynal, "Polysomnography: A response to a need for improved communication," presented at the 14th Annu. Meeting Assoc. Psychophysiology. Study Sleep, 1974. [Online]. Available: <https://www.lib.uchicago.edu/e/scrc/findingaids/view.php?eadid=ICU.SPCL.SLEEP#idp90478928>
11. R. A. U., O. Faust, N. Kannathal, T. Chua, and S. Laxminarayan, "Nonlinear analysis of EEG signals at various sleep stages," *Comput. Methods Programs Biomed.*, vol. 80, no. 1, pp. 37-45, 2005.
12. Fonseca P, Den T N, Long X, et al. Cardiorespiratory Sleep Stage Detection Using Conditional Random Fields[J]. *IEEE Journal of Biomedical & Health Informatics*, 2017, 21(4):956-966.

13. Li X, Cui L, Tao S, et al. HyCLASS: A Hybrid Classifier for Automatic Sleep Stage Scoring[J]. IEEE J Biomed Health Inform, 2017, PP (99):1-1.
14. S.-F. Liang, C.-E. Kuo, Y.-H. Hu, and Y.-S. Cheng, "A rule-based automatic sleep staging method," *Journal of neuroscience methods*, vol.205, no. 1, pp. 169–176, 2012.
15. R. B. Berry, R. Brooks, C. E. Gamaldo, S. Harding, C. Marcus, and B. Vaughn, "The aasm manual for the scoring of sleep and associated events," *Rules, Terminology and Technical Specifications*, Darien, Illinois, American Academy of Sleep Medicine, 2012.
16. Memar P, Faradji F. A Novel Multi-Class EEG-Based Sleep Stage Classification System[J]. IEEE Transactions on Neural Systems & Rehabilitation Engineering A Publication of the IEEE Engineering in Medicine & Biology Society, 2018, 26(1):84.
17. Supratak A, Dong H, Wu C, et al. Deep Sleep Net: a Model for Automatic Sleep Stage Scoring based on Raw Single-Channel EEG[J]. IEEE Transactions on Neural Systems & Rehabilitation Engineering A Publication of the IEEE Engineering in Medicine & Biology Society, 2017, PP (99):1-1.
18. Shokrollahi M, Krishnan S. A Review of Sleep Disorder Diagnosis by Electromyogram Signal Analysis[J]. *Critical Reviews in Biomedical Engineering*, 2015, 43(1):1.
19. Zhao M, Yue S, Katabi D, et al. Learning Sleep Stages from Radio Signals: A Conditional Adversarial Architecture[C]// *International Conference on Machine Learning*. 2017.
20. Kim J, Lee J, Shin M. Sleep stage classification based on noise-reduced fractal property of heart rate variability[J]. *Procedia Computer Science*, 2017, 116:435-440.
21. Yoon H, Hwang S H, Choi J W, et al. REM sleep estimation based on autonomic dynamics using R-R intervals[J]. *Physiological Measurement*, 2017, 38(4):631-651.
22. Flexer A, Gruber G, Dorffner G. A reliable probabilistic sleep stager based on a single EEG signal [J]. *Artificial Intelligence in Medicine*, 2005, 33(3):199-207.
23. Bianchi M T, Eiseman N A, Cash S S, et al. Probabilistic sleep architecture models in patients with and without sleep apnea[J]. *Journal of Sleep Research*, 2012, 21(3):330–341.
24. J. M. Kortelainen, M. O. Mendez, A. M. Bianchi, M. Matteucci, and S. Cerutti, "Sleep staging based on signals acquired through bed sensor," *IEEE Trans. Inf. Technol. Biomed.*, vol. 14, no. 3, pp. 776–85, May 2010.
25. D. Novak, Y. Al-Ani, and L. Lhotska, "Electroencephalogram processing 'using hidden markov models,'" in *5th EUROSIM Congress on Modelling and Simulation (EUROSIM'04)*, 2004.
26. L. Doroshenko and V. Konyshov, "Usage of hidden markov models for automatic sleep stages classification," in *Russian-Bavarian Conference on Bio-Medical Engineering*. Citeseer, 2007, p. 19.
27. S.-T. Pan, C.-E. Kuo, J.-H. Zeng, S.-F. Liang et al., "A transition constrained discrete hidden markov model for automatic sleep staging,"
28. Kim J, Lee J, Shin M. Sleep stage classification based on noise-reduced fractal property of heart rate variability[J]. *Procedia Computer Science*, 2017, 116:435-440.
29. Hayano J, Yuda E, Yoshida Y. Sleep stage classification by combination of actigraphic and heart rate signals[C]// *IEEE International Conference on Consumer Electronics - Taiwan*. IEEE, 2017:387-388.
30. Yoon H, Hwang S H, Choi J W, et al. REM sleep estimation based on autonomic dynamics using R-R intervals[J]. *Physiological Measurement*, 2017, 38(4):631-651.
31. Scherz W D, Fritz D, Velicu O R, et al. Heart rate spectrum analysis for sleep quality detection[J]. *Eurasip Journal on Embedded Systems*, 2017, 2017(1):26.
32. S Devot, R Dratwa, E Naujokat, in *Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE*. Sleep/wake detection based on cardiorespiratory signals and actigraphy ,(2010), pp. 5089–5092
33. Yu S, Chen X, Wang B, et al. Automatic sleep stage classification based on ECG and EEG features for day time short nap evaluation[C]// *Intelligent Control and Automation*. IEEE, 2012:4974-4977.
34. A. Rechtschaffen and A. Kales, *A Manual of Standardized Terminology, techniques and Scoring System for Sleep Stages of Human Subjects*, ser. Publication. Brain Information Service/Brain Research Institute, University of California, 1968.
35. H. Schulz, "The AASM manual for the scoring of sleep and associated events," *J. of Clinical Sleep Medicine*, vol. 3, no. 7, p. 752, 2007.
36. Quan, S.F., Howard, B.V., Iber, C., Kiley, J.P., Nieto, F.J., O'Connor, G.T., Rapoport, D.M., Redline, S., Robbins, J., Samet, J.M., & Wahl, P.W. (1997). The sleep heart health study: design, rationale and methods. *Sleep*, 20, 1077-1085
37. Sleep Heart Health Study S.R.C. (2002). Reading Center Manual of Operations, Tech Report VMLA-039-02, Case Western University
38. Sors A, Bonnet S, Mirek S, et al. A convolutional neural network for sleep stage scoring from raw single-channel EEG[J]. *Biomedical Signal Processing & Control*, 2018, 42:107-114.
39. PAN, J., TOMPKINS, W. J., "A Real-Time QRS Detection Algorithm" *IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING*, VOL. BME-32, NO. 3, MARCH 1985.
40. M. Sami, M. Isa, Ito Wasito and Aniti Mur, "Kernel Dimensionality Reduction on Sleep Stage Classification using ECG Signal," *IJCSI International Journal of Computer Science Issues*, Vol. 8, Issue 4, No 2, July 2011.
41. R.D. Berger, S. Akselrod and D. Gordon, "An efficient algorithm for spectral analysis of the heart rate variability," *IEEE Trans Biomed Eng.*, vol.33, pp. 220-222, 1986.
42. Azar, A.T., El-Said, S.A.: 'Performance analysis of the support vector machine classifiers in breast cancer mammography recognition', *Neural Comput. Appl.*, 2014, 24, (5), pp. 1163–1177

43. Li H, Peng C, Ye D. A study of sleep staging based on a sample entropy analysis of electroencephalogram[J]. Bio-Medical Materials and Engineering, 2015, 26(s1): 1149-115
44. "National sleep research resource website, <https://sleepdata.org/>."