

Supplementary Materials

of “Privacy-Protected Sleep Staging using Blurred Videos”

<https://github.com/contactless-healthcare/Privacy-Protected-Sleep-Staging-using-Blurred-Videos>

1 Mini-Review: Comparison of Three Sleep Staging Modalities

To balance at-home deployability and privacy, sleep staging has increasingly shifted from traditional multi-channel PSG toward contactless and semi-contact sensing. The Fig. 1 reviews three representative sleep staging modalities including Radio Frequency (RF)-based, pressure-based, and camera-based, each offering varying degrees of privacy protection:

- (i) RF-based approaches leverage wireless signals for non-contact vital signs monitoring, ensuring strong privacy protection but can only measure motion signals (e.g. respiratory motion, ballistocardiographic motion).
- (ii) Pressure-based methods, which are semi-contact, avoid direct image or audio recording and can be seamlessly integrated into bedding, offering excellent privacy protection and a plug-and-play solution, but still limited to vibration signals.
- (iii) Camera-based methods, while rich in measuring multiple physiological information, pose inherent privacy concerns, making privacy-preserving techniques essential for their applications.

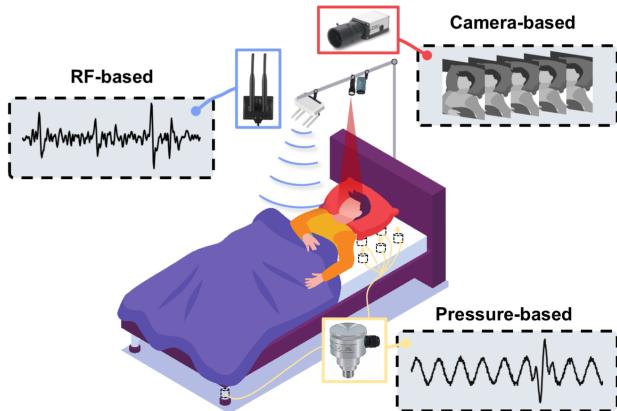


Figure 1: Comparison between the three sleep staging modalities: RF-based, pressure-based, and camera-based.

2 Clinical Setting and Sleep Information Summary

We give a detailed description of the experimental hardware, a summary of participant sleep reports, and an overview of the gold-standard PSG channels:

- (i) This study was conducted in a sleep unit at the Respiratory Disease Institute of Shenzhen People’s Hospital, China, approved by the Institutional Review Board (IRB) with the IRB No.LL-KY-202237401. The experiment used an infrared camera (IDS-UI3860M, 968×548 pixels, 20 fps), equipped with a MidOpt TB850 bandpass filter (850 ± 22 nm) and a 1.5 W infrared supplementary light to ensure high-quality video capture in low-light environments. The Near-Infrared (NIR) video was recorded continuously throughout the night, simulating a real clinical environment.

Table 1: The sleep-related information of 20 participants.

Num	Age	Gender	BMI (kg / m ²)	AHI (events / h)	Efficiency (%)
1	23	male	21.0	4.0	92.0
2	21	male	22.3	4.9	91.3
3	25	male	23.4	22.5	88.3
4	24	male	20.1	5.1	66.5
5	22	male	21.3	3.6	70.4
6	23	male	18.4	1.2	83.5
7	23	male	19.7	1.8	95.4
8	22	female	18.7	0.3	36.0
9	21	female	20.8	1.2	60.8
10	22	female	19.0	1.3	79.2
11	20	female	20.6	1.0	92.0
12	33	male	22.5	12.1	85.3
13	29	female	27.3	0.7	85.2
14	39	male	32.2	81.6	89.8
15	23	female	20.8	5.2	90.8
16	31	male	26.7	16.5	94.4
17	29	female	17.6	4.5	78.5
18	35	female	21.1	8.4	64.9
19	21	male	20.6	2.7	79.2
20	21	female	18.3	7.0	90.2

Participants naturally fell asleep during the monitoring, and including normal behaviors such as bed exiting, nurse check-ups, different sleep postures, and with or without bed coverings.

(ii) A total of 20 adults were recruited for the experiment, consisting of 11 males and 9 females, with an average age of 29 ± 3.6 years and an average body mass index (BMI) of 23 ± 2.3 kg / m². Detailed sleep-related information for the 20 participants, including age, gender, BMI, apnea–hypopnea index (AHI), and sleep efficiency is presented in Table 1. Each participant slept approximately from 10:00 PM to 7:00 AM the next day, with a total duration of 176.8 hours of sleep data collected, averaging 8.5 hours per person. All participants completed a self-assessment of their sleep quality and signed informed consent forms.

(iii) PSG recordings were performed using the Philips Respironics ALICE 6 LDxN system, which provides 76 channels, that collect various physiological signals: Electroencephalogram (EEG, F3/M2, F4/M1, C3/M2, C4/M1, O1/M2, O2/M1), Electromyogram (EMG, chin and leg EMG), nasal airflow (thermosensitive and pressure airflow), chest and abdominal movements (THO and ABD), Electrocardiogram (ECG1, ECG2), Photoplethysmogram (PPG), and SpO₂, among others.

3 Feature Summary and Visualization

To clearly present the video-derived physiological and motion features (four feature categories, 15 parameter types) and the details of the corresponding feature engineering (e.g., multi-scale sliding windows), as well as to intuitively illustrate how the extracted features vary across sleep stages, we provide a summary table and an overnight raw feature trace for visualization.

Feature summary. Table 2 provides an organized overview of the video-derived physiological and motion features used in our analysis, grouped into cardiac/respiratory, heart rate variability (HRV),

and motion domains. The accompanying clarifications further elucidate how these features span multiple temporal scales and body regions.

Table 2: Summary of video-derived physiological and motion features.

Feature Description	Dimension
Cardiac and Respiratory ^{1,3}	3
Mean value of heart rate	
Mean value of respiratory rate	
Standard deviation of respiratory rate	
Heart Rate Variability ¹	7
Mean value of interbeat interval (IBI)	
Standard deviation of NN intervals (SDNN)	
Root mean square of successive differences (RMSSD)	
High frequency (HF) power	
Low frequency (LF) power	
Very low frequency (VLF) power	
Low frequency to high frequency ratio (LF/HF)	
Motion ²	5
Mean value of raw activity	
Mean value of fast activity plus slow activity	
Mean value of fast activity multiply slow activity	
Motion percentage within 30-sec window	
Static duration between two adjacent large motion	

Some detail clarification:

¹ The multi-scale perception is applied with three scales.

² All features are extracted from both leg and whole body.

³ The extraction window here also shifts ± 15 seconds.

Feature visualization. Fig. 2 presents the feature traces alongside corresponding sleep staging. Overall, heart rate (HR) and breathing rate (BR) delineate the rough sleep structure, motion features are particularly informative for distinguishing Wake and Deep sleep, and HRV provides complementary cues. Together, these four feature categories contribute to the final staging. Key feature changes are highlighted with red boxes, such as elevated HR and BR during REM, opposite LF/HF trends during sleep, and motion patterns related to Light and Deep.

4 Cardiopulmonary Coupling (CPC) Calculation using rPPG

In order to clearly present the CPC analysis workflow using camera-based rPPG, We provide mathematical details and an algorithm diagram below.

Methodological Details. The CPC calculation for ECG signals is based on the R-R interval (RR) sequence and the ECG-derived respiratory signal (EDR). Similarly, for camera PPG signals, CPC

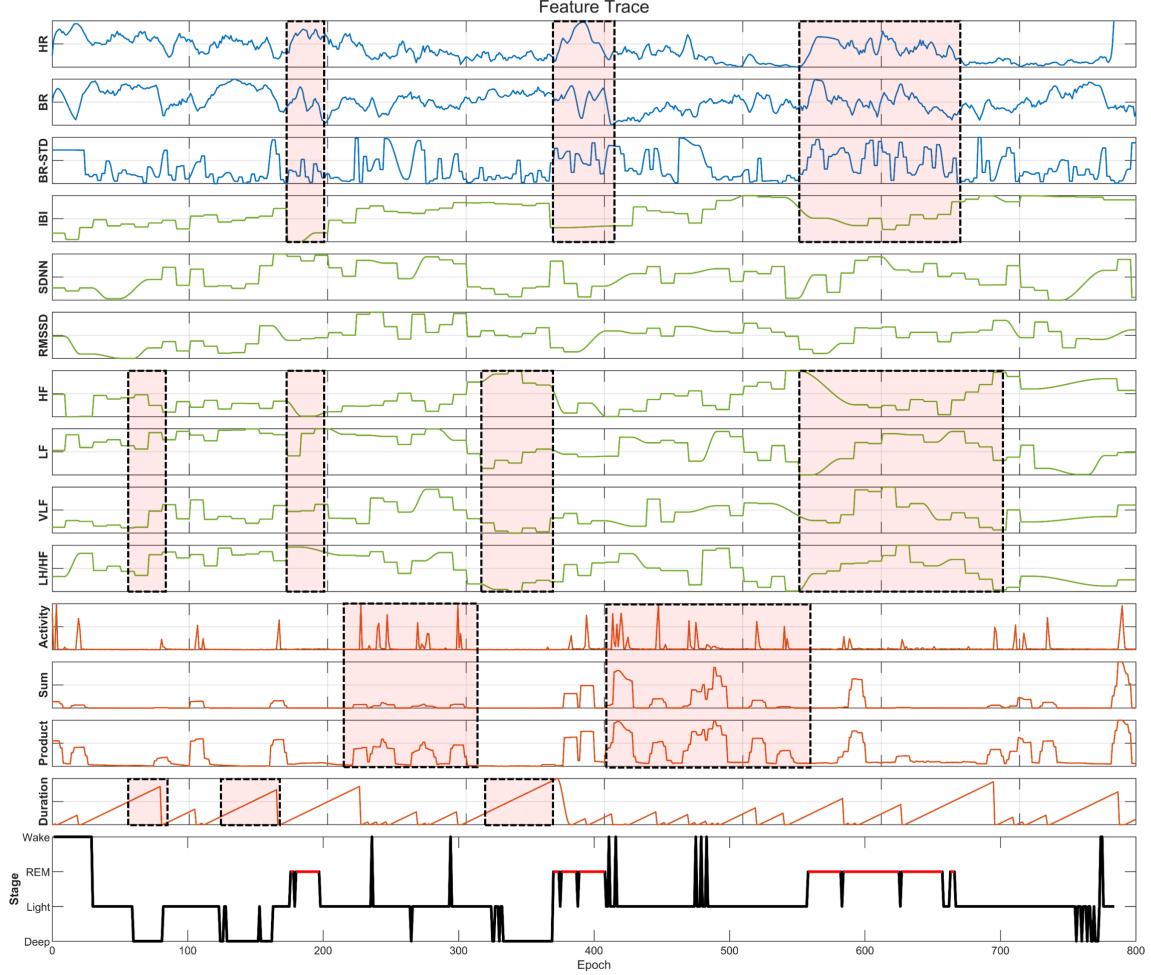


Figure 2: Overview of 15 raw feature traces and corresponding sleep staging. Feature traces are color-coded by category: HR & BR (blue), IBI & HRV (green), and Motion (orange).

combines the PPG-derived IBI (equivalent beat-to-beat information to ECG RR) with the PPG-derived respiratory signal (PDR), as shown in Fig. 3. According to the Nyquist theorem, the sampling rate must be at least twice the highest frequency component of the signal, meaning each respiratory cycle should include at least two RR and EDR to reliably detect the respiratory signal and prevent high-frequency components from aliasing into lower-frequency components. In our experiment, the heart rate of all participants was more than twice the respiratory rate, ensuring this condition was met.

After resampling IBI and PDR sequences at 2 Hz, the cross-spectrum power $\Gamma_m(I, P)$ describes the amplitude and phase relationship between IBI intervals and PDR signals in the frequency domain. It is defined as:

$$\Gamma_m(I, P) = A_{I,m} A_{P,m} e^{j(\Phi_{P,m} - \Phi_{I,m})}, \quad (1)$$

where:

- $A_{I,m}$ and $A_{P,m}$: The amplitudes of IBI and PDR signals at the frequency m ;
- $\Phi_{I,m}$ and $\Phi_{P,m}$: The phases of IBI and PDR signals at the frequency m .

The consistency of the phase difference between IBI and PDR signals is evaluated through the magnitude squared coherence. The coherence is defined as the squared average cross-spectrum

divided by the product of the average spectral power of individual signals:

$$\Delta_m(I, P) = \frac{\Gamma_m(I, P)^2}{(A_{I,m} e^{j\Phi_{I,m}})^2 (A_{P,m} e^{j\Phi_{P,m}})^2}. \quad (2)$$

The coherence effectively quantifies the degree of synchronization and phase consistency between the two signals, reflecting their strong correlation in the frequency domain.

To minimize the impact of limited sampling windows on the results, Welch's averaged periodogram method was employed to estimate the cross-correlation matrices. The CPC intensity is quantified as follows:

$$\beta(f_m) = \Gamma_m(I, P)^2 \cdot \Delta_m(I, P)^2, \quad (3)$$

where:

- $\beta(f_m)$: The CPC intensity at frequency f_m ;
- $\Gamma_m(I, P)$: The cross-spectrum power describes the frequency-domain correlation between the signals;
- $\Delta_m(I, P)$: The coherence validates the phase consistency between the signals.

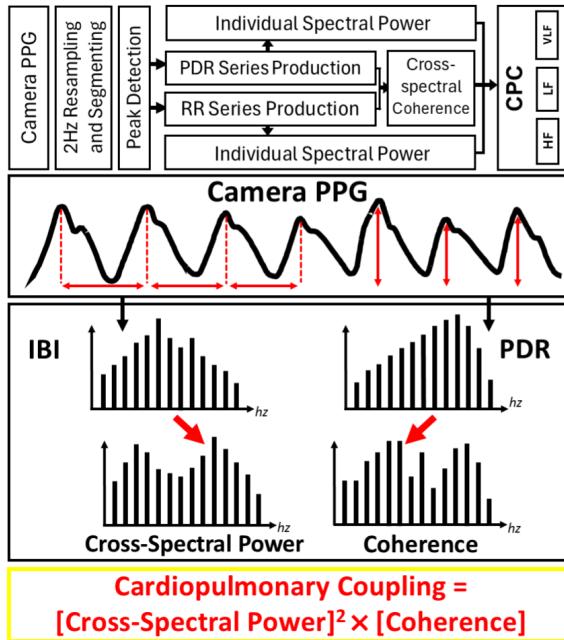


Figure 3: Algorithm diagram of the camera-based CPC analysis.

5 Privacy Protection Test

To provide an intuitive, visual evaluation of the blur experiment and demonstrate how increasing blur suppresses identifiable facial details, we provide snapshots of all 20 participants across different blur conditions in Fig. 4. Under the four levels of blurring, the tasks of face detection, facial landmark detection, and face recognition all failed, which means that the subject identity was protected. Furthermore, image restoration experiments show that subject faces cannot be restored from blurred images using the state-of-the-art algorithm [1], further confirming that our blurring approach is irreversible.

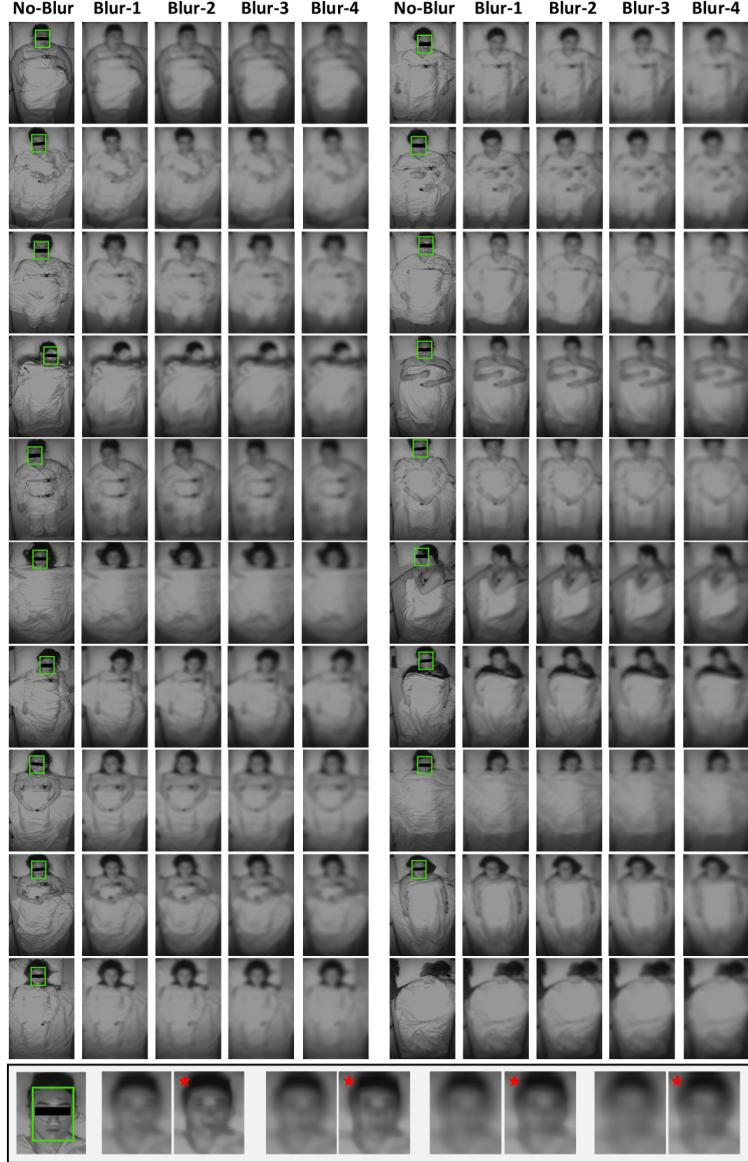


Figure 4: Examples of blurred frames with different degrees of blurring for 20 participants. Green boxes mark detected faces in the original image. The bottom row with a star shows blurred face after image restoration.

6 Sleep Parameter Quality

To assess the quality of vital signs measured from blurred videos, the mean absolute error (MAE) between the camera-extracted measurements and synchronized PSG references is calculated over 30-second epochs, consistent with the standard epoch length used in clinical analysis. The specific references are: HR by ECG, BR by nasal airflow, and motion by EMG. Pearson correlation coefficient (R-value) is used to evaluate the linear correlation between parameters measured by the camera and PSG. One point of clarification, the motion detected by the camera is pixel displacement, while PSG records voltage changes through EMG. To facilitate the comparison, the motion signals are binarized (1 for motion, 0 for still), and the detection accuracy is calculated by motion percentage (proportion of motion duration over each epoch).

The quality of sleep parameters is comprehensively evaluated under four blurring conditions

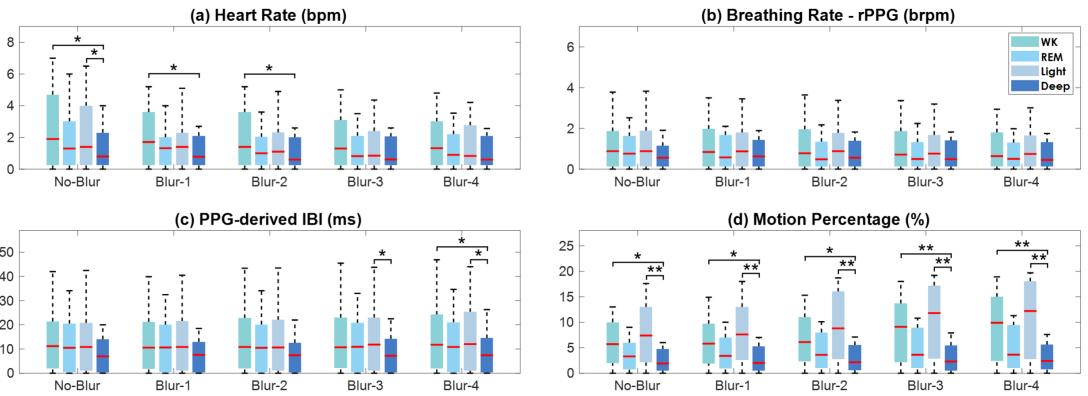


Figure 5: The MAE between camera measurements, including HR, rPPG-derived BR, PPG-derived IBI and motion percentage, and PSG reference under four blurring simulations. Boxplots are used to visualise the distribution of MAEs across four sleep stages (Wake, REM, Light, Deep).

(Blur 1-4) across four sleep stages. First, for different levels of blurring, Fig. 5 (a) - (b) shows that the MAE, upper quartile, and maximum values of HR and BR derived from the rPPG decreased as the blurring level increased, while the between-class statistical significance progressively diminished, indicating that blurring improved the SNR of rPPG signals through pixel fusion, which aligns with the previous studies [2]. For the raw features used for HRV calculation (i.e. IBI), the overall performance remain stable, with no significant change in the median MAE as the blurring level increased. However, with more severe blurring, the upper quartile and maximum values increased, and the between-class statistical significance between deep sleep and other stages became more evident, indicating larger MAE dispersion shown in Fig. 5 (c). Blurred videos lose texture details, making motion analysis more difficult. Fig. 5 (d) shows that the MAE, upper quartile, and maximum values of motion percentage increased significantly with higher blurring, indicating that blurring cause strong disturbances to motion analysis.

Further analysis across four sleep stages reveals that for IBI and motion, IBI experiences more interference in the Light stage compared to the others, while motion parameters are notably affected in all stages except the Deep stage. Overall, blurring pollutes the HRV and motion analysis, but the improved PPG signal quality ensures that HR and PPG-derived BR are still reasonably robust.

7 Individual Feature Trend Analysis

To provide a more comprehensive and statistically meaningful characterization of feature differences across sleep stages, in addition to the raw feature traces shown in Fig. 2 of *Supplementary Materials 3*, we also evaluated the feature trends across the four sleep stages, before applying the extracted features to train the sleep-staging classifier. This analysis helps assessing whether, after blurring, the features can still reflect the expected trends in different sleep stages, and provides physiological interpretability. To minimize the influence of outliers, we used a 5% trimmed mean method [3], which offers a more accurate representation of the overall trends. Statistical analyses included ANOVA and Tukey’s post-hoc tests for multiple comparisons, and the Shapiro-Wilks test was used to check for normality, with a P value < 0.05 is considered to be significant.

HR and BR Features. Across the four sleep stages, the temporal feature trends of heart rate and respiratory rate remain highly consistent across different blurring levels, as shown in Fig. 6 (1) - (3). As sleep deepens, feature values gradually decrease, transitioning from Wake to REM, Light, and Deep. The distinction between REM and non-REM stages is particularly noticeable. These trends reflect the reduction in energy demands and a slowdown in metabolic activity as sleep progresses. Notable physiological phenomena, such as sleep “bradycardia”, reduced respiratory intensity and

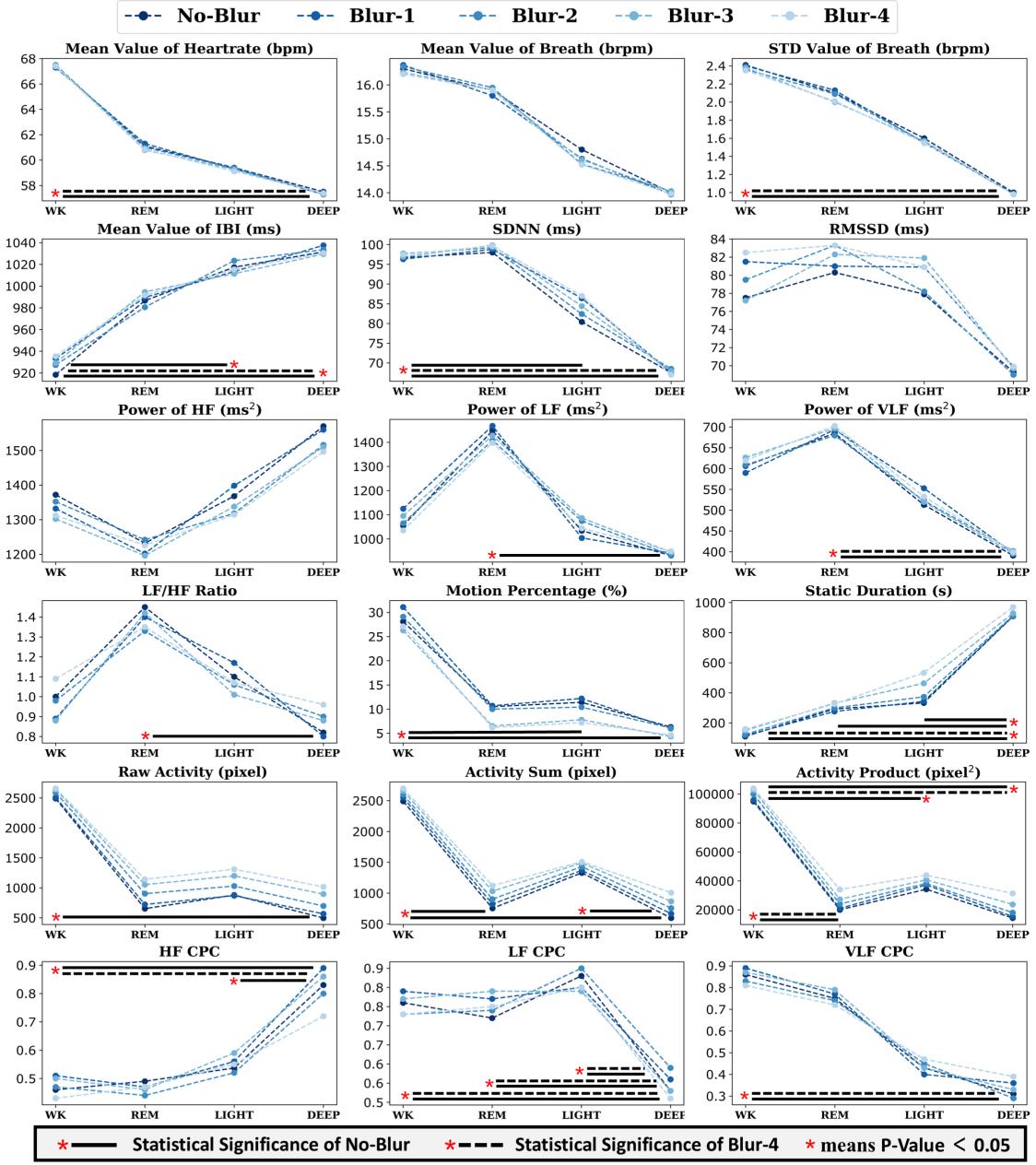


Figure 6: Overview of the trend analysis of 15 prior features and 3 CPC features at the bottom across *Wake*, *REM*, *Light*, and *Deep* sleep stages and blur gradient *Blur-1*, *Blur-2*, *Blur-3*, and *Blur-4*. All metrics are performed on all 20 volunteers, and a Trimmed Mean is used to focus on 95% of the data. Solid lines show significant differences (p -value < 0.05) across sleep stages of the No-Blur condition, while dashed lines indicate the most blurred condition, no markers are shown for non-significant differences. All these metrics can be categorized as: *HR* & *BR* (1-3), *IBI* & *HRV* (4-10), *Motion* (11-15), *CPC* (16-18).

variability, are stabilized under vagal nerve regulation [4, 5, 6]. Statistical analysis reveals significant differences (p -value < 0.05) in mean heart rate and respiratory rate variability between Wake and Deep, consistently observed under both No-Blur and Blur-4 conditions. Overall, pulsatile and respiratory features remain unaffected by blurring and consistently reflect physiological changes during sleep stage transitions, further demonstrating the robustness of PPG signals under blurred conditions.

HRV Feature. Under varying blurring conditions, time-domain parameters of HRV still remain

effective for indicating autonomic nervous system regulation. As shown in Fig. 6 (4) - (6), the mean IBI exhibits a steady upward trend across Wake, REM, Light, and Deep, while SDNN generally shows a downward trend, which are consistent with studies by Bušek and Herzig [7, 8]. However, SDNN values during the REM stage are slightly higher than in Wake, possibly due to limited data collection and the fact that most wakefulness data reflect the relaxed pre-sleep state in an ideal, undisturbed environment. The mean RMSSD follows a trend similar to SDNN, but at higher blurring levels (Blur-3 and Blur-4), variations in the Wake and REM stages show greater fluctuations. IBI statistically distinguishes Wake from both Light and Deep. However, with Blur-3 and Blur-4, the distinction between Wake and Light diminishes to non-significance. SDNN maintains significant differences between Wake and Deep. Overall, the time-domain parameters of HRV show consistent trends across four blurring conditions, although higher blurring levels reduce their statistical significance.

The frequency-domain features of HRV are more significantly affected by blur interference compared to time-domain. By analyzing the distribution and dynamic changes of cardiac activity parameters across different frequency ranges, frequency-domain features can more sensitively reflect autonomic nervous system regulation across sleep stages, but they also demand higher signal quality. As shown in Fig. 6 (7) - (9), high-frequency (HF) power increases with deeper sleep stages, primarily reflecting parasympathetic nervous system activity. Low-frequency (LF) power, which indicates sympathetic nervous system activity, exhibits a continuous decline, consistent with the physiological pattern of parasympathetic dominance during deeper sleep and sympathetic dominance during Wake [6]. While very low-frequency (VLF) power, focusing on a narrower low-frequency range, shows a more pronounced and stable downward trend.

The overall trends of the frequency components remain largely consistent with the reference (No-Blur). LF and VLF components show greater stability across different blurring levels, while HF components gradually weaken as blurring increases, despite only minor changes in absolute power. The LF/HF ratio serves as a comprehensive indicator of the relative balance between sympathetic and parasympathetic nervous system activity. As shown in Fig. 6 (10), the ratio rises during Wake and gradually declines across REM, Light, and Deep. The peak ratio of the REM stage is associated with increased sympathetic activities and decreased parasympathetic activities [9]. However, as the blurring level increases, this trend gradually flattens. After Blur-2, the statistical significance of differences between REM and Deep disappears. These findings indicate that while the overall trends of frequency components are preserved under blurring, their proportions fluctuate considerably. HF components are particularly affected by blurring, disrupting the overall stability of HRV frequency-domain metrics and further impacting the accuracy of sleep stage predictions.

Motion Feature. Signal-level evaluation indicates that motion features are significantly affected by blurring, and feature analysis further confirms such interference. As shown in Fig. 6 (13) - (15) motion features raw activity, activity sum and activity product (activity sum and activity product combined short-term and long-term motion for capturing slight motion), extracted based on intensity, exhibit a consistent trend: motion is most prominent during Wake, while it decreases sequentially across the Light, REM, and Deep. This pattern aligns with sleep physiology, where reduced motion intensity during the REM and Deep is associated with the self-protective mechanism of “muscle atonia” [10], in which most skeletal muscles lose tone during the REM sleep. As the blurring increases, mean values of all motion features steadily rise, accompanied by an increase in the standard deviation of features across all sleep stages. This indicates that blurring introduces surrounding noise (e.g., pixels without texture information) into the real motion signals, causing a continuous elevation of the baseline. While blurring diminishes texture, it introduces more baseline noise at the same time, reducing the SNR of motion signals.

State-based motion features further validate that threshold will fail in some blur cases. As

shown in Fig. 6 (11) - (12), the motion percentage shows a decreasing trend, indicating fewer movements exceeding the threshold detected per unit of time. Static duration reveals a gradual increase in the intervals between movements, further reflecting the reduction in movements exceeding the threshold. The increasing standard deviation of these features suggests challenges in distinguishing motion signals from the background, leading to a significant decline in the sensitivity for motion monitoring. Statistical results indicate that activity sum lost the significant difference between Light vs. Deep, Wake vs. REM as early as Blur-1, while activity product lost the distinction between Wake and Light after Blur-3. This suggests that the sensitivity of motion features are significantly impacted by blur. From the perspective of predicting sleep stages, blur has minimal impact on distinguishing Wake and Deep but significantly weakens the ability to represent the Light.

CPC Feature. Across the four sleep stages, CPC metrics show consistent patterns under all blurring conditions, as shown in Fig. 6 (16) - (18). HF CPC is significantly higher in the Deep than in other stages, indicating stronger coupling driven by enhanced vagal modulation, with more synchronous interactions between respiratory and cardiac rhythms under stable parasympathetic control [11, 6]. LF coupling is more prominent during Wake or light sleep conditions as shown in Fig. 6 (17) - (18), when sympathetic activity[12] and long-period oscillations [11] are more pronounced. LF CPC reaches the lowest level in Deep, while VLF CPC shows a clearer decline across all four stages. Statistical analysis reveals significant differences ($p\text{-value} < 0.05$) between Deep and other stages for HF CPC and LF CPC, and between Wake and Deep for VLF CPC. These findings are consistent under both No-Blur and Blur-4 conditions. Although both HRV and CPC reflect autonomic nervous system regulation during sleep, CPC features are less affected by blurring and provide a more stable contribution to sleep stage classification.

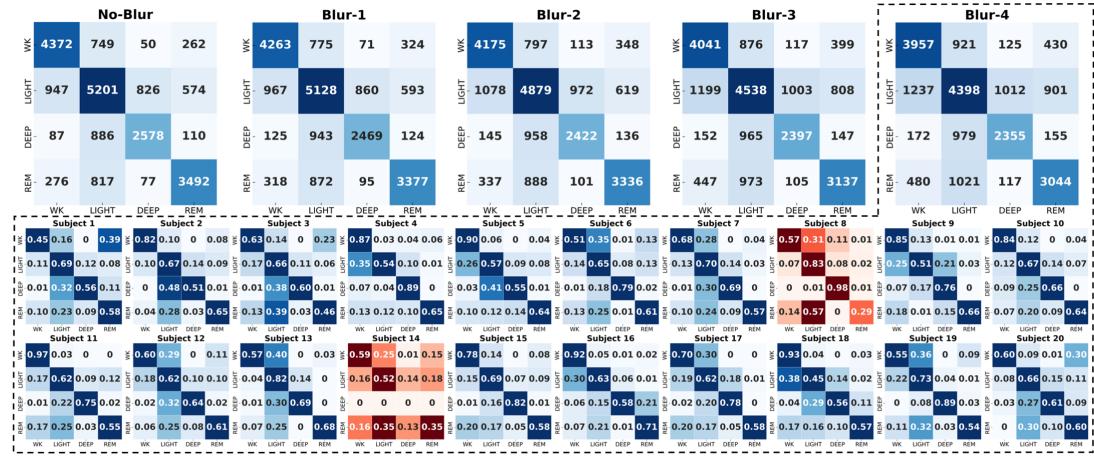


Figure 7: The upper five panels show the overall confusion matrices of Set - 4 using RF classifier across all subjects under different blur conditions (No-Blur to Blur-4), while the lower panels display the confusion matrices of 20 individual subjects under the most-blurred (Blur-4) condition, with red-highlighted matrices indicating subjects with abnormal sleep patterns.

8 Overall and Subject-Level Confusion Matrices

The overall four-class (Wake, REM, Light, Deep) confusion matrices from No-Blur to Blur-4 are provided to assess how blur affects the performance of each individual class. And subject-specific matrices under the most blurred condition are provided to illustrate inter-subject variability.

As shown in the confusion matrices of Fig. 7 and recall values in Table 3, among all sleep stages, Light is the most affected by blurring. As a dominant category accounting for more than 40%, the

Table 3: Recall Value of Set-4 (HR, BR, HRV, Motion and CPC) using RF classifier.

	WK	LIGHT	DEEP	REM
No-Blur	0.76 ± 0.09	0.69 ± 0.09	0.70 ± 0.07	0.72 ± 0.08
Blur-1	0.75 ± 0.11	0.68 ± 0.10	0.67 ± 0.08	0.71 ± 0.09
Blur-2	0.73 ± 0.11	0.65 ± 0.11	0.66 ± 0.09	0.70 ± 0.09
Blur-3	0.72 ± 0.13	0.62 ± 0.12	0.65 ± 0.09	0.66 ± 0.11
Blur-4	0.71 ± 0.13	0.60 ± 0.14	0.64 ± 0.09	0.64 ± 0.12

accuracy of Light dropped from 68.91% to 58.27%, with nearly 3% decrease per blur level. This change is consistent with the feature trend analysis, where motion features lost their significant differences between Light and other stages. For the REM stage, although key HRV features such as LF/HF progressively degraded under blurring, the classification accuracy remained at 65.29%. This may be attributed to the REM indicator [13], which fuses time-domain HR and BR signals derived from PPG, as the raw PPG signal quality was minimally affected by blurring. We further examined the individual confusion matrices and referred to their sleep-health information (AHI and sleep efficiency) in Table 1 to interpret performance differences. Overall, most subjects exhibited balanced classification performance across sleep stages. Notably, Subject 8 had a sleep efficiency of only 36% and entered deep sleep briefly at the beginning of the night, spending the remainder alternating between light sleep and wakefulness, a pattern consistent with the “first-night effect” [14]. Subject 14, with an AHI of 81.6 indicating severe OSA, experienced almost no deep sleep throughout the night. These subject-level observations provide context for the variability in classification outcomes.

9 Additional Analysis of the Blur Impact

We added line graphs showing the trends of accuracy, Cohen’s kappa, and F1-score to more intuitively illustrate the blur impact. Beyond the ablation study, which examines how different features are affected by blur, we also designed additional comparative experiments to investigate its impact on staging performance:

- Frame rate impact analysis focuses on HRV features by resampling the original 20 Hz signal to 200 Hz and 500 Hz, evaluating the improvement in HRV signal quality and staging performance with higher frame rates;
- Block size analysis evaluates different receptive field configurations by varying the size of local blocks: original size ([12 24 48] and [48 72 96 120]), reduced receptive field ([6 12 24] and [48 72 96 120]), and multi-level receptive field ([6 12 24] and [24 36 48 60]), to evaluate the impact of blurring on staging and its adaptability.

The results show that the classification performance of all combinations decrease with the increase of blurring in Fig. 8. The feature Set-4 (HR, BR, HRV, motion and CPC) performed the best and showed a buffering compensatory effect against blurring. The combination of HR and BR alone (green line) showed stable performance under different blurring conditions, consistent with the results of signal-level and feature-level analysis. After adding HRV features (red line), performance decreased significantly with gradient blurring, particularly in Blur-3 and Blur-4. This indicates that blurring interferes with HRV signals, and, based on feature trend analysis, is more

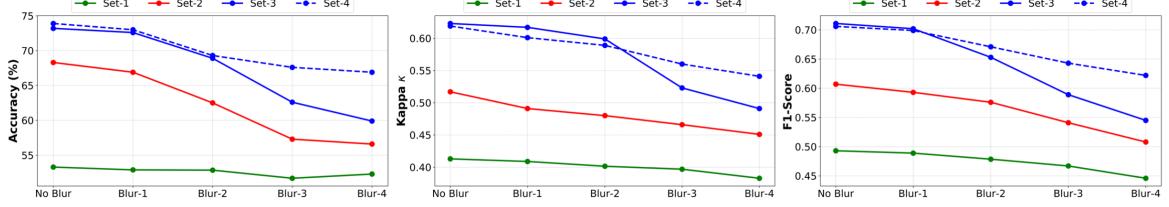


Figure 8: Line graph of Accuracy (%), Cohen’s κ , F1-score using the four features combinations including Set-1 (HR and BR), Set-2 (HR, BR, and HRV), Set-3 (HR, BR, HRV, and motion), and Set-4 (HR, BR, HRV, motion, and CPC).

likely to disturb high-frequency features. Adding motion signals (blue line) still showed a consistent decline, with a more noticeable downward trend in Blur-3 and Blur-4 compared to the red line, further confirming the significant interference of high-intensity blurring on motion analysis. After adding CPC features, the sharp accuracy drop under Blur-4 was mitigated, which improved from 59.8% to 66.9%, indicating that CPC is more robust to blurring than HRV. This is further supported by the ablation study, where CPC alone yielded higher performance than HRV alone under blurred conditions. Furthermore, in the current sleep dataset with relatively healthy subjects, the four-class performance under effective privacy protection at Blur-2 (accuracy = 69.5%, kappa = 0.59, F1-score = 0.67) still lags behind the clinical PSG and semi-contact pressure-based modalities, indicating the need of further optimization toward real-world applications.

Additionally, Table 4 summarizes our staging evaluation under various configurations, aiming to investigate potential factors that influence sleep staging performance under blurring framework, such as HRV sampling rate and multi-scale block size. Due to the camera’s sampling rate (20 Hz), the spatial resolution of the HRV signal is relatively low. Experimental results show that resampling had minimal impact, and performance consistently declined as blur increased. Reducing the size of PPG boxes to shrink the receptive field and suppress environmental noise led to minor improvements in staging performance. However, feature degradation remained evident with increasing blur, and neither method mitigated its impact.

Table 4: Accuracy (%) with Different Configurations using Feature Set-3.

HRV sampling	20 Hz	200 Hz	500 Hz
No-Blur	73.3 ± 8.1	73.6 ± 8.9	73.0 ± 8.7
Blur-1	72.6 ± 8.0	72.8 ± 9.3	72.2 ± 9.3
Blur-2	68.9 ± 8.6	69.5 ± 9.9	68.5 ± 10.1
Blur-3	62.6 ± 9.5	63.0 ± 10.8	62.3 ± 11.1
Blur-4	59.8 ± 10.0	58.6 ± 11.6	58.7 ± 11.7

Block size	Size-1	Size-2	Size-3
No-Blur	73.3 ± 8.1	73.4 ± 8.0	73.2 ± 8.1
Blur-1	72.6 ± 8.0	73.0 ± 8.0	72.7 ± 8.2
Blur-2	68.9 ± 8.6	69.5 ± 8.8	69.2 ± 8.7
Blur-3	62.6 ± 9.5	63.1 ± 9.7	63.4 ± 9.7
Blur-4	59.8 ± 10.0	60.0 ± 10.1	59.9 ± 10.3

Size-1: PPG box [12 24 48] and motion box [48 72 96 120].

Size-2: PPG box [6 12 24] and motion box [48 72 96 120].

Size-3: PPG box [6 12 24] and motion box [24 36 48 60].

10 CPC Compensation Reasoning

Although CPC incorporates IBI, a time-domain HRV feature affected by blurring, the staging results were still improved under blurring after introducing CPC. To determine whether the robustness of CPC under blurring is because that one input signal being less affected and thus dominating the calculation, we analyzed the two input signals (IBI and PDR) separately.

Table 5: Repeated measures correlation coefficients between Camera PPG and PSG ECG.

	No Blur	Blur-1	Blur-2	Blur-3	Blur-4
IBI vs RR	0.860	0.848	0.821	0.802	0.778
PDR vs EDR	0.906	0.898	0.879	0.870	0.853

As shown in Table 5, we analyzed the repeated measures correlation between PPG-IBI and the gold standard ECG-RR, as well as between PDR and EDR under different blurring conditions. Regardless of the blurring condition, the correlation for PDR was always higher than that for IBI, and the degree of PDR attenuation was much smaller than that of IBI. This suggests that PDR is less affected by blurring, reflecting the stability of respiratory amplitude.

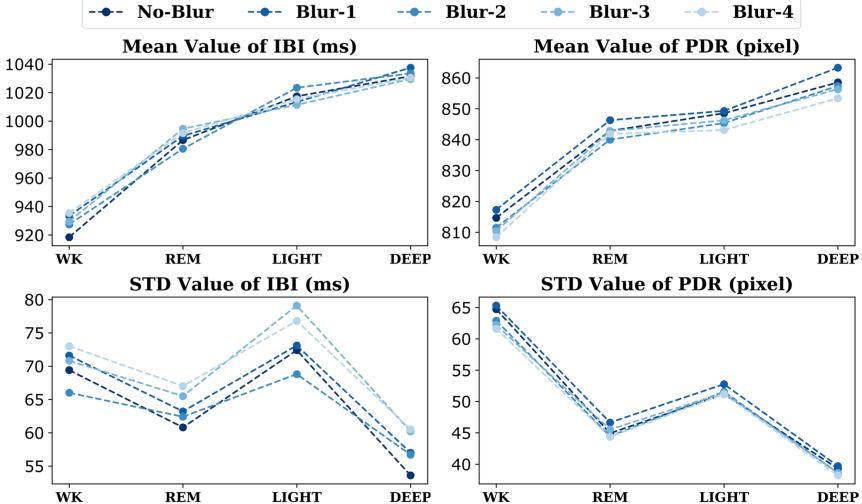


Figure 9: The stability of two input signals (IBI and PDR) of CPC under different blur conditions.

To more intuitively demonstrate the response of the two input signals of CPC to blurring across the four sleep stages, Fig.9 shows the consistency analysis of the mean PDR and IBI (RR values) under different blurring conditions. The STD value of PDR maintained good stability, while IBI was more significantly affected by blurring. Therefore, the resistance of CPC to interference of blurring likely stems primarily from the PDR signal. The PDR signal preserved the original features well, and IBI and PDR are orthogonal (independent) in the signal space. Even if IBI is affected by blurring, PDR still dominates the calculation of the cross-power spectrum, compensating for overall coherence and enhancing CPC's robustness.

References

- [1] Syed Waqas Zamir et al. “Multi-stage progressive image restoration”. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 2021, pp. 14821–14831.
- [2] Yingen Zhu, Hong Hong, and Wenjin Wang. “Privacy-protected Contactless Sleep Parameters Measurement using A Defocused Camera”. In: *IEEE Journal of Biomedical and Health Informatics* (2024).
- [3] Tianming Hu and Sam Yuan Sung. “A trimmed mean approach to finding spatial outliers”. In: *Intelligent Data Analysis* 8.1 (2004), pp. 79–95.
- [4] Walter Baust et al. “The regulation of heart rate during sleep”. In: *Experimental brain research* 7 (1969), pp. 169–180.
- [5] Yasutaka Mukai and Akihiro Yamanaka. “Functional roles of REM sleep”. In: *Neuroscience Research* 189 (2023), pp. 44–53.
- [6] David W Carley et al. “Physiology of sleep”. In: *Diabetes spectrum: a publication of the American Diabetes Association* 29.1 (2016), pp. 5–9.
- [7] P Bušek et al. “Spectral analysis of heart rate variability in sleep”. In: *Physiol res* 54.4 (2005), pp. 369–376.
- [8] David Herzig et al. “Reproducibility of heart rate variability is parameter and sleep stage dependent”. In: *Frontiers in physiology* 8 (2018), p. 1100.
- [9] I Berlad et al. “Power spectrum analysis and heart rate variability in Stage 4 and REM sleep: evidence for state-specific changes in autonomic dominance”. In: *Journal of sleep research* 2.2 (1993), pp. 88–90.
- [10] Yasuo Hishikawa and Tetsuo Shimizu. “Physiology of REM sleep, cataplexy, and sleep paralysis.” In: *Advances in neurology* 67 (1995), pp. 245–271.
- [11] Robert Joseph Thomas et al. “An electrocardiogram-based technique to assess cardiopulmonary coupling during sleep”. In: *Sleep* 28.9 (2005), pp. 1151–1161.
- [12] Virend K Somers et al. “Sympathetic-nerve activity during sleep in normal subjects”. In: *New England Journal of Medicine* 328.5 (1993), pp. 303–307.
- [13] Qiongyan Wang, Hanrong Cheng, and Wenjin Wang. “Video-PSG: An Intelligent Contactless Monitoring System for Sleep Staging”. In: *IEEE Transactions on Biomedical Engineering* (2024).
- [14] Jong-Ho Byun et al. “The first night effect during polysomnography, and patients’ estimates of sleep quality”. In: *Psychiatry research* 274 (2019), pp. 27–29.