Seoul National University
Graduate School of Data Science

SNUH SEOUL NATIONAL UNIVERSITY HOSPITAL

OuarLab

SlAction: Non-intrusive, Lightweight Obstructive Sleep Apnea Detection using Infrared Video

You Rim Choi[†1] Gyeongseon Eo[†1] Wonhyuck Yoon[3] Hyojin Lee[1] Haemin Jang[1] Dongyoon Kim[4] Hyunwoo Shin[*2,3] Hyung-Sin Kim[*1]

Seoul National University[1] Seoul National University College of Medicine[2] OuarLab[3] Columbia University[3]

INTRODUCTION

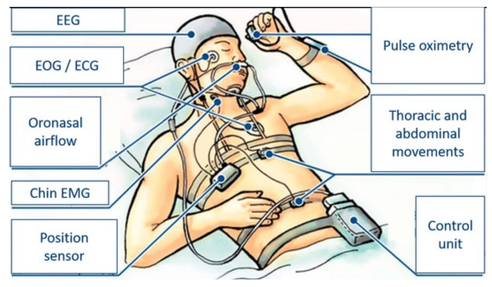
- **Obstructive sleep apnea (OSA)** is one of the most common sleep disorders

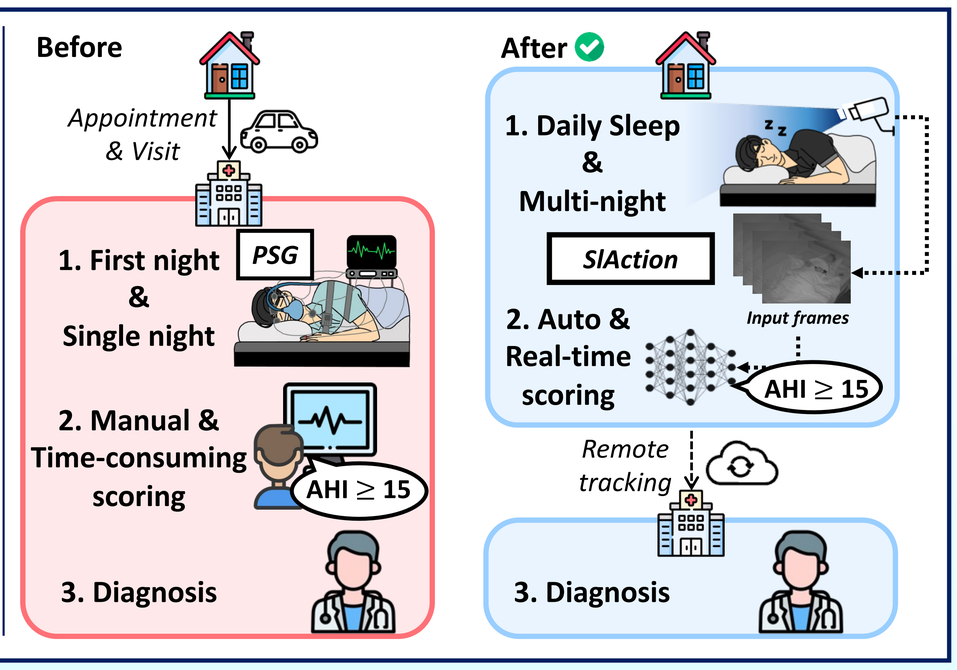
Apnea Hypopnea Index (AHI)

$$= \frac{\text{Apneas + Hypopneas}}{\text{Total sleep time (hours)}}$$

AHI	Rating
< 5	Normal
5 – 15	Mild OSA
15 – 30	Moderate OSA
> 30	Severe OSA

- **Polysomnography (PSG):** Gold standard for sleep evaluation
  - ✓ Expensive / Time-consuming / Interscorer discrepancy
  - ✓ PSG-based OSA diagnoses can be inaccurate due to
    (1) First night effect (Attaching sensors in an unfamiliar environment)
    (2) Single-night stay (No consideration of night-to-night variability)





PRELIMINARY STUDY

- **Dataset**

- Infrared videos from 729 patients, during PSG
- A diverse range of genders and ages spanning from teenagers to individuals over 80 years old
- Various camera angles and room environments across hospitals
- 640 x 480, 5 fps

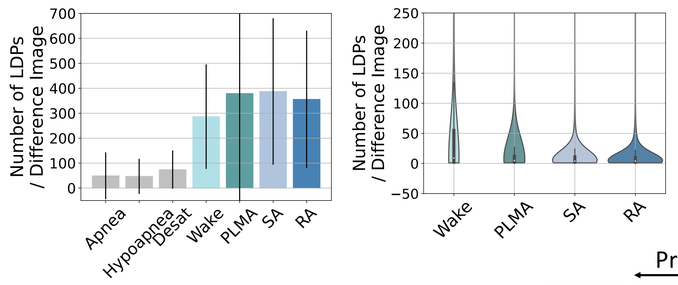
Advantages

- ✓ Subjects can sleep without sensors; useful for daily sleep monitoring
- ✓ Available anywhere and familiar with users
- ✓ Measuring equipment is not expensive

Challenges

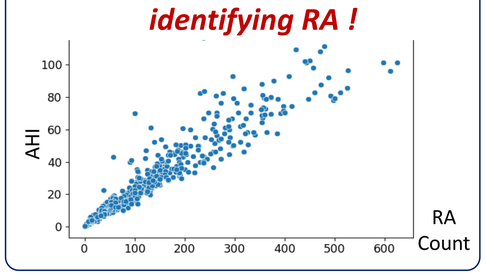
- ✓ Low quality and noisy video
- ✓ Static overall view
- ✓ Not much movement discernible to the naked eyes
- ✓ Blurred information (*e.g.* subject's face)

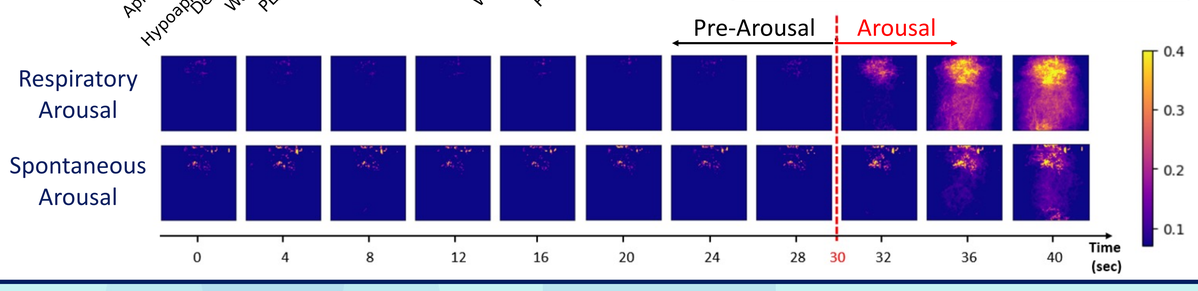
- **Key Idea using vision AI for OSA detection**
  - Difficult to visually distinguish between OSA and normal breathing
  - However, in the case of respiratory arousal (RA) accompanied by OSA, the **movement is greater**



Linear relationship between RA frequency and AHI

Indirectly detect OSA by identifying RA !

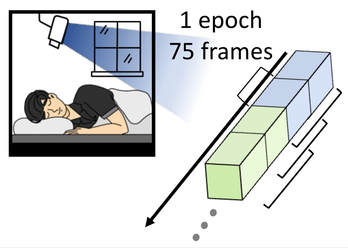




METHOD

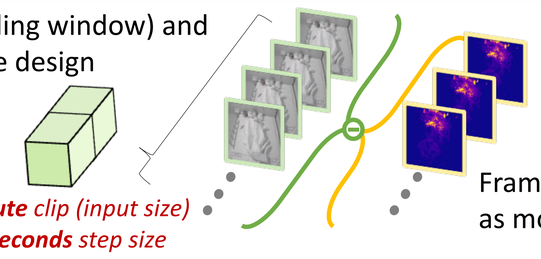
- ***SlAction*** estimates AHI by counting the occurrence of RA during the total sleep time.

Infrared Video Capturing



(1) Input Data Preprocessor

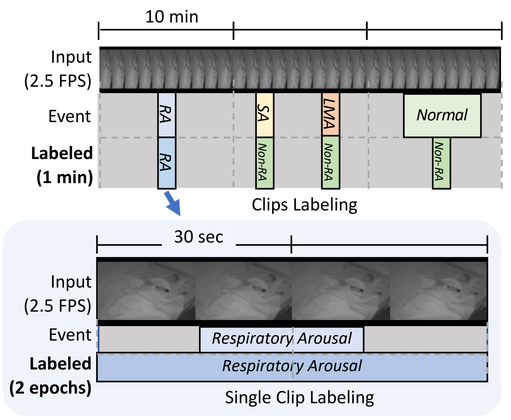
Clip(Sliding window) and step size design



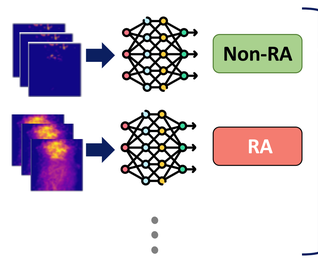
***1 minute** clip (input size)*
***30 seconds** step size*

Frame difference as motion input

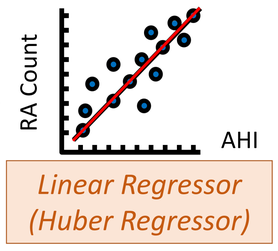
(2) RA Detector



- Deep neural network that classifies RA or not
- Training data is strategically curated



(3) AHI Estimator



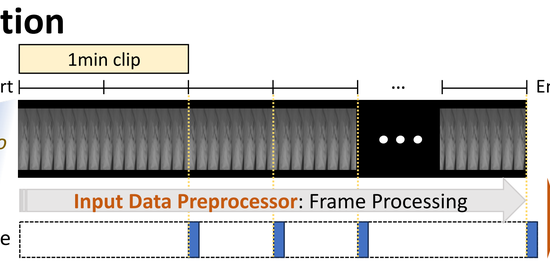
Linear Regressor (Huber Regressor)

Final Diagnosis

If AHI ≥ 15: OSA
Else: No OSA

- **On-device inference operation**

NVIDIA Jetson Nano
- quard-core ARM A57 CPU
- 4GB 64-bit LPDDR4 memory



EVALUATION

- **Evaluation Dataset**

Train A Hospital, 449 patients

Valid

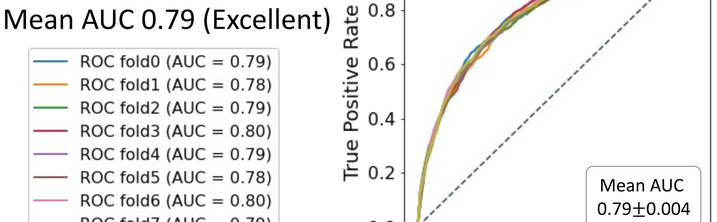
A	B	C	Total
50	0	35	85

✓ AHI Estimator fitting

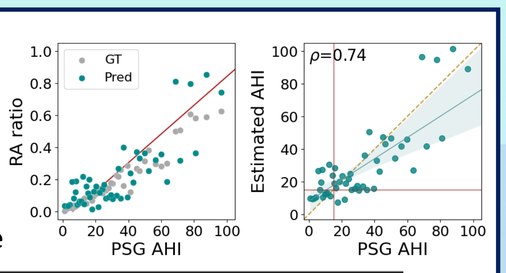
Test

A	B	C	Total
50	115	80	245

- **RA Detector**
  - 9-fold cross validation
  - Mean AUC 0.79 (Excellent)



- **AHI Estimator & OSA Prediction**
  - High Spearman Correlation coefficient values between the estimated AHI and PSG AHI
  - Good OSA prediction performance on hospital A (F1 score 0.876) and B (F1 score 0.891)
  - However, the performance for hospital C was relatively lower
    → Refine the AHI estimator using A valid set and C valid set → Shows improved performance



Dataset	Estimator fitting dataset	AHI Estimation Spearman correlation coefficient (ρ)	P-value	OSA Prediction Accuracy (%)	Precision	Recall	F1 Score
A valid (50 patients)	A valid	0.827	1.37e-13	84.0	0.886	0.886	0.886
A test (50 patients)	A valid	0.744	5.89e-10	82.0	0.842	0.914	0.876
B test (115 patients)	A valid	0.756	8.60e-23	83.4	0.867	0.918	0.891
C test (80 patients)	A valid	0.834	8.16e-22	65.0	1.000	0.594	0.745
C test (80 patients)	A & C valid	0.834	8.16e-22	83.7	0.924	0.884	0.903

- **On-device Performance**

Model size (FP16)	Model load Frame capture ready	Frame processing (1 min. clip)	Inference	Total Operation	Peak Memory (RSS)	Peak Memory (Runtime)
5.1 (MB)	1.035±0.007 (s)	0.224±0.042 (s)	3.040±0.046 (s)	3.264±0.088 (s)	839±15.5 (MB)	2.67±0.016 (GB)

CONCLUSION

- We propose a **practical approach** to address PSG limitations for OSA diagnosis, utilizing ***human motion*** during sleep.
- Our research **pioneers the use of sleep video** to detect OSA events, identifying unique features **with clinical expertise and empirical analysis.**
- We introduce effective **input data curation** for identifying slow and long-term motions in a lightweight manner. *SlAction* is **implemented on the resource constrained device for fast AHI estimation.**
- Through extensive experiments, we demonstrate that *SlAction* maintains **robust accuracy regardless of various environmental changes** such as camera location, sleeping position, blanket occlusion. These results **highlight the broad applicability of *SlAction* in diverse real-world settings.**