

Real-Time Snoring Detection Using Deep Learning: A Home-Based Smartphone Approach for Sleep Monitoring

Joonki Hong^{1,*}, Seung Koo Yang^{2,*}, Seunghun Kim^{1,*}, Sung-Woo Cho^{2,3}, Jayoung Oh², Eun Sung Cho¹, In-Young Yoon⁴, Dongheon Lee¹, Jeong-Whun Kim^{2,3}

¹Asleep Research Institute, Seoul, Republic of Korea; ²Department of Otorhinolaryngology-Head and Neck Surgery, Seoul National University Bundang Hospital, Seoul National University College of Medicine, Seongnam, Republic of Korea; ³Sensory Organ Research Institute, Seoul National University Medical Research Center, Seoul, Republic of Korea; ⁴Department of Psychiatry, Seoul National University Bundang Hospital, Seoul National University College of Medicine, Seongnam, Republic of Korea

*These authors contributed equally to this work

Correspondence: Jeong-Whun Kim, Department of Otorhinolaryngology-Head and Neck Surgery, Seoul National University Bundang Hospital, Seoul National University College of Medicine, 82 Goomiro 173-Boengil, Bundang-gu, Seongnam, 13620, Republic of Korea, Email kimemails7@gmail.com

Background: Despite the prevalence of sleep-related disorders, few studies have developed deep learning models to predict snoring using home-recorded smartphone audio. This study proposes a real-time snoring detection method utilizing a Vision Transformer-based deep learning model and smartphone recordings.

Methods: Participants' sleep-breathing sounds were recorded using smartphones, with concurrent Level I or II polysomnography (PSG) conducted in home or hospital settings. A total of 200 minutes of smartphone audio per participant, corresponding to 400 30-second sleep stage epochs on PSG, were sampled. Each epoch was annotated independently by two trained labelers, with snoring labeled only when both agreed. Model performance was evaluated by epoch-by-epoch prediction accuracy and correlation between observed and predicted snoring ratios.

Results: The study included 214 participants (85,600 epochs). Hospital audio data from 105 participants (42,000 epochs) were used for training, while home audio data from 109 participants were split into 54 participants (21,600 epochs) for training and 55 participants (22,000 epochs) for testing. On the test dataset, the model demonstrated a sensitivity of 89.8% and a specificity of 91.3%. Correlation analysis showed strong agreement between observed and predicted snoring ratios ($r = 0.97$, 95% CI: 0.95–0.99).

Conclusion: This study demonstrates the feasibility of using deep learning for real-time snoring detection from home-recorded smartphone audio. With high accuracy and scalability, the approach offers a practical and accessible tool for monitoring sleep-related disorders, paving the way for home-based sleep health management solutions.

Keywords: snoring, smartphone, artificial intelligence, polysomnography, remote sensing technology

Introduction

Snoring, a common occurrence during sleep, affects a significant portion of the population worldwide. Characterized by the turbulent airflow causing tissues in the upper airway to vibrate, snoring manifests as the often disruptive and audible sound during breathing while asleep. A study suggests that approximately 23% of adults habitually snore, with prevalence rates varying across age groups and genders.¹ Another study of large UK Biobank ($n=408,317$) proposed that 37% of registrants reported others complaining about their snoring.² While snoring is often considered a nuisance or a source of comedic anecdotes, it can also be indicative of underlying health concerns. Persistent and loud snoring may be a symptom of sleep-disordered breathing, including conditions such as obstructive sleep apnea (OSA). Chronic snoring can increase upper airway resistance, leading to airway collapsibility over time. This progression may contribute to intermittent hypoxia and sleep fragmentation, key features of OSA. Early monitoring and intervention could help mitigate the risk of developing OSA and related health complications. Studies have also associated habitual snoring with disruptions in sleep architecture, leading to sleep fragmentation and decreased sleep quality.^{3,4}

Graphical Abstract



When snoring presents by itself, without documented apneas, hypopneas, or hypoventilation, it is often referred to as primary snoring, intermittent snoring, or habitual snoring; and it is classified in the mild end of the spectrum of the sleep-related breathing disorders.⁵ Even in the absence of OSA, some studies showed that snoring is associated cardio-metabolic risk factors such as atherosclerosis and endothelial dysfunction.^{6,7} A recent study proposed the proportion of high blood pressure, hyperlipidemia, diabetes, atrial fibrillation, transient ischemic attack, or high-risk group of stroke risk rating was significantly higher in snoring group in Chinese population over 40 years old.⁴

Currently, there is widespread agreement on the significance of OSA and its treatment. However, many sleep medicine professionals have shifted their attention towards this more severe condition, paying less attention to primary snoring. There is a lack of standardized classification for primary snoring in terms of its severity or frequency, and existing diagnostic tools present challenges such as the high cost and impracticality of PSG for long-term monitoring, the limited focus of portable sleep apnea tests on apneas rather than snoring, and the lack of validation in consumer-grade snoring detection apps. Additionally, it remains uncertain whether the frequency, intensity, or specific location of snoring may contribute to potential health risks such as atherosclerosis.⁸

Primary snoring itself can lead to sleep disruptions for the individuals and, potentially, their bed partner. Even without the respiratory events seen in sleep apnea, the loud and repetitive nature of snoring may contribute to sleep fragmentation, reducing overall sleep quality.⁹ Persistent exposure to a snoring partner can contribute to heightened stress and anxiety levels for the non-snorer.¹⁰ Individuals with primary snoring may experience daytime sleepiness, fatigue, and declined cognitive function, especially in the population of children.¹¹ While these symptoms may not be as severe as those associated with sleep apnea, they can still impact daily activities, work performance, and overall quality of life. In addition, a significantly positive association between depression and various sleep-related parameters, including snoring, was reported.¹²

In this study, we aimed to develop technology capable of real-time monitoring of snoring, which holds significant potential for clinical and home-based applications. Real-time snoring detection could facilitate personalized interventions, such as body weight reduction and positional therapy adjustments in positional snorers or titration of oral appliances in patients undergoing mandibular advancement therapy. Such technology is not only essential for calculating the proportion of snoring during sleep but also holds significant potential for applications in developing real-time interventions to mitigate snoring. By quantifying

the percentage of time spent snoring, researchers and healthcare professionals can gauge the severity of snoring episodes, which helps in categorizing individuals into different severity levels, allowing for a more nuanced understanding of the impact of snoring on sleep quality. While primary snoring alone does not indicate sleep apnea, it is important for healthcare professionals to monitor individuals over time. Primary snoring may be a precursor to the development of sleep apnea in some cases.¹³ Also, for individuals undergoing interventions or treatments aimed at reducing snoring, calculating the percentage of time spent snoring provides a measurable outcome. This metric allows for the assessment of treatment effectiveness and helps in adjusting interventions as needed.

Since snoring has acoustic characteristics and can be recorded using microphones or smartphones, creating a snoring prediction model using smartphone-recorded sounds is both feasible and highly practical. Smartphones present a cost-effective and accessible alternative to traditional sleep monitoring tools, eliminating the need for additional specialized equipment. Their widespread availability enables large-scale home-based sleep monitoring, reducing barriers to participation and allowing for longitudinal tracking of snoring patterns. Several studies have investigated audio-based snoring detection, including smartphone-integrated models, but many face limitations in accuracy, generalizability, and real-time application. Few have been validated against polysomnography in both hospital and home environments. This study addresses these gaps by developing a deep learning-based real-time snoring detection model, optimized for home-based monitoring and validated with level I and II PSG. In a previous study, we developed a sound-based deep learning model for sleep stage classification using home-recorded audio data.¹⁴ Building on this foundation, the current study introduces an advanced model specifically designed and trained to detect snoring events using home-recorded smartphone audio. Unlike previous approaches that primarily relied on PSG-based audio recordings in controlled environments, our model was developed and validated on a large dataset spanning both hospital and home settings, ensuring its robustness across diverse acoustic conditions. Our deep learning approach enhances accuracy by incorporating contextual information from adjacent audio segments. The neural network architecture has been adapted to align with the unique objectives of snoring event prediction. Therefore, the aim of this study is to train and validate a snoring prediction model using sounds recorded using smartphones during sleep at home. The potential benefits of the study may include providing a scalable and accessible alternative to traditional snoring assessment, enabling continuous home-based monitoring without the need for expensive or intrusive equipment. By facilitating early identification of individuals at risk for sleep-disordered breathing and tracking snoring severity over time, this approach enhances both clinical decision-making and sleep health management.

Methods

Study Population and Data Collection

We included participants who underwent polysomnography (PSG) and simultaneously recorded snoring sounds using a smartphone during their sleep between January 2019 and February 2023. The dataset comprised both retrospectively collected data from participants who had previously provided consent for secondary data use under an Institutional Review Board (IRB)-approved protocol and prospectively collected data from newly enrolled participants following additional IRB approval (IRB No. B-2205-755-308, April 2022). Participants were selected based on the inclusion criterion of having at least 200 minutes of PSG-derived sleep stage epochs (excluding wake epochs) to ensure that a sufficient number of 30-second epochs could be labeled for snoring presence. A power analysis was not conducted beforehand, as the study aimed to utilize all available participants meeting this criterion rather than selecting a predetermined sample size. The PSG studies were conducted either at a tertiary care hospital or in the participants' homes. At the tertiary care hospital, Level 1 attended full-night polysomnography (PSG, Embla N 7000, Reykjavik, Iceland) was performed. A smartphone (LG G3, LG Electronics, Inc., Seoul, Republic of Korea) was positioned 1 meter away from the participants to record snoring sounds throughout the night. In the home environment, Level 2 unattended full-night PSG (Embletta MPR/ST+ Proxy, Natus Medical Inc) was performed at the participants' residences. During the sleep period, a smartphone (iPhone 11, Apple Inc., Cupertino, CA, USA) was placed 1 meter away from the participants to capture snoring sounds. Apart from this placement requirement, no additional constraints or guidelines regarding the sleeping environment (eg, room acoustics, bed positioning, or noise control) were imposed. This approach was chosen to allow participants to maintain their habitual sleep conditions, ensuring that the recorded snoring events reflected real-

world variability in home environments. Both Level 1 and Level 2 PSG included the following measurements: electroencephalography, electrooculography, chin and limb electromyography, electrocardiography, nasal pressure transducer, chest and abdomen respiratory inductance plethysmography, and pulse oximetry. This study was conducted following the approval of the institutional review board at the tertiary hospital, with all participants providing written informed consents. The report followed the Standards for Reporting of Diagnostic Accuracy (STARD) 2015 guidelines.¹⁵

Dataset Composition and Preparation for Snoring Prediction

The snoring sound dataset was divided into a training dataset and a test dataset without oversampling or under sampling. Instead, we apply class weighting during training to address this issue. Since snoring epochs were comparatively fewer than non-snoring epochs, we applied approximately 2.5X class weights when training on snoring epochs. To achieve a balanced and representative dataset, home PSG participants were allocated based on snoring percentage distribution, ensuring that the test dataset was not disproportionately different from the training dataset. This method helped maintain diversity in snoring severity and clinical characteristics while preventing data leakage. The training dataset consisted of snoring sound data collected from participants who underwent either Level 1 or Level 2 PSG in hospital or home settings to ensure exposure to diverse acoustic environments. The test dataset, however, included only snoring sound data from participants who underwent Level 2 PSG in home settings, ensuring that the model's performance was evaluated in real-world, home-based conditions. Hospital PSG data, recorded in a controlled sleep laboratory environment with minimal background noise, provided high-quality training samples, whereas home PSG data introduced natural variability in noise levels, room acoustics, and smartphone placement. This dataset design ensured that the model was trained on a broad range of conditions while being tested under realistic home monitoring settings, reinforcing its generalizability and applicability for real-world use. In order to label snoring, 200-minute recorded audio data on a smartphone matched to four hundred 30-second epochs of sleep stages on PSG were sampled from each participant.

Labeling of Snoring

Snoring labeling was conducted in a 30-second epoch-by-epoch manner for the audio recordings. Two independent labelers, selected through a competitive accuracy-based screening process, annotated each epoch for the presence of snoring. To ensure high reliability, a strict consensus-based approach was employed, where an epoch was labeled as containing snoring only if both labelers independently identified snoring within the epoch. If there was any disagreement, the epoch was classified as non-snoring. This approach eliminates ambiguity and ensures that only confidently labeled snoring events were included in the dataset. The final classification was binary, with each 30-second epoch designated as positive for snoring if a snoring event was captured within the epoch window and negative otherwise (Figure 1).

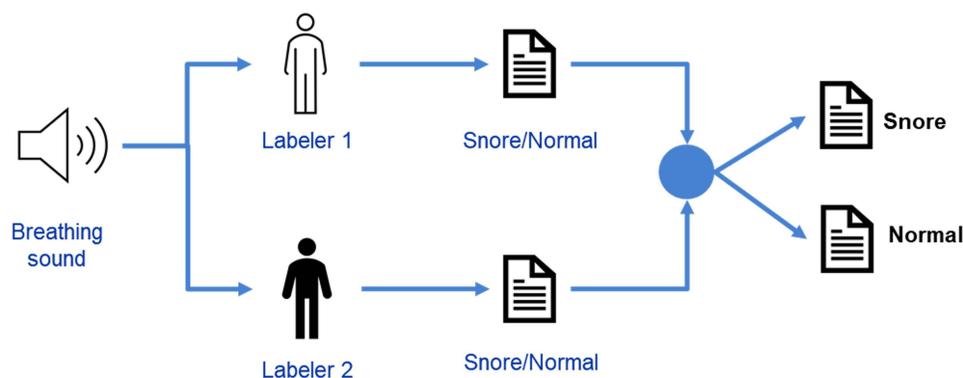


Figure 1 Snoring identification process involving two independent labelers. An epoch was conclusively labeled as snoring only if both labelers independently agreed on its presence.

Snoring Prediction Model

We developed a snoring prediction model based on sound which was transformed into Mel-spectrograms. Before generating Mel spectrograms, the raw audio recordings underwent adaptive noise reduction to minimize environmental noise while preserving the original spectral characteristics of snoring sounds. This method dynamically adjusted to background noise variations without altering the core acoustic properties of snoring events. No additional filtering, amplitude normalization, or artificial data augmentation techniques were applied, as we aimed to maintain the natural variability in snoring characteristics across different participants and recording environments. Instead, the model robustness was enhanced through semi-supervised learning techniques, such as consistency loss and contrastive learning. These approaches helped improve generalization across diverse snoring patterns and recording conditions.

Each Mel spectrogram was labeled as “snoring” or “normal” according to the presence of snoring during each 30-second epoch. We utilized a Vision Transformer-based deep learning model similar to the one employed in our previous research.¹⁶ In brief, we designed an model for detecting snoring events on an epoch-by-epoch basis. It takes 14 input epochs and generates predictions for 10 output epochs, categorizing each as either “snoring” or “normal”. The model comprises two main components: a feature extractor and a multi-epoch detector. The feature extractor processes Mel spectrograms to capture unique acoustic characteristics of snoring. Building on this, the multi-epoch detector identifies snoring events and determines their classification by leveraging contextual information from adjacent epochs.

Evaluation of Model Performance

We validated the performance of the snoring prediction model using the test dataset. Epoch-by-epoch agreement between the snoring prediction model and human-annotated ground-truth snoring. The performance is shown as accuracy, macro F1 score, Cohen kappa value, sensitivity and specificity. We also performed subgroup analyses to evaluate the performance according to the snoring percentage of each participant.

We conducted a thorough analysis comparing the percentage of snoring in a session as predicted by our model with the corresponding values derived from human-annotated ground-truth data. Employing Pearson correlation analysis, we assessed the concordance between the smartphone predictions and the PSG-based percentages of snoring events. The performance evaluation matrix, in this context, includes regression metrics such as the coefficient of determination, mean absolute error, and root mean square error. Statistical analyses were carried out using IBM SPSS Statistics, version 26 (IBM Corp), and continuous parametric variables were presented as means (standard deviation [SD]). A subgroup analysis of subjects with apnea hypopnea index (AHI) less than 15 per hour (hr) and AHI greater than or equal to 15 per hour were performed to evaluate the effect of AHI on the accuracy of snoring percentage prediction.

Results

General Characteristics of Study Participants

The audio dataset in the hospital environment was collected from 105 participants and the hospital dataset was used for training. The audio dataset in the home environments was collected from 109 participants. Among them, 54 was used for training and 55 for testing. Because 400 30-second epochs were sampled from each participant, the total number of epochs used in this study was 85,600 epochs from 214 participants.

Table 1 summarizes the general and PSG characteristics of hospital training dataset (n=105; 42,000 epochs; 69 males), home training dataset (n=54; 21,600 epochs; 26 males), and home test dataset (n=55; 22,000 epochs; 28 males). The mean (SD) age for each subgroup was 52.9 (12.9), 50.3 (14.7), and 46.0 (15.3) years, respectively ($p = 0.757$). The percentage of males was 65.7%, 48.1%, and 50.9%, respectively ($p = 0.055$). The mean (SD) body mass index (BMI) was 26.2 (4.2), 24.4 (5.3), and 25.1 (4.2) Kg/m², respectively ($p = 0.185$). The mean (SD) AHI was 23.3 (23.8), 10.8 (15.2) and 13.2 (16.0) per hour, respectively.

Epoch-by-Epoch Performance of Snoring Event Prediction

The epoch-by-epoch performance of our snoring prediction model was tested using 22,000 epochs of the test dataset. The model showed a good epoch-by-epoch prediction performance for snoring with a sensitivity of 89.8% and an specificity of 91.3%.

Table 1 Demographic and Polysomnographic Characteristics of Study Subjects

	Training Dataset		Test Dataset
	Hospital Audio (n=105, 42,000 epochs)	Home Audio (n=54, 21,600 epochs)	Home Audio (n=55, 22,000 epochs)
Age (years), mean (SD)	52.8 (12.7)	50.3 (14.6)	46.0 (15.2)
Males, n (%)	69 (65.7%)	26 (48.1%)	28 (50.9%)
BMI (Kg/m ²), mean (SD)	26.1 (4.2)	24.3 (5.3)	25.0 (4.1)
BMI<25 kg/m ² , n (%)	46 (43.8%)	39 (72.2%)	28 (50.9%)
BMI≥25 kg/m ² , n (%)	59 (56.1%)	15 (27.7%)	27 (49.0%)
AHI (/hr), mean (SD)	23.2 (23.8)	10.8 (15.2)	13.2 (16.0)
AHI<15/hr, n (%)	52 (49.5%)	42 (77.7%)	35 (63.6%)
AHI≥15/hr, n (%)	53 (50.4%)	12 (22.2%)	20 (36.3%)

Abbreviations: SD, standard deviation; BMI, body mass index; AHI, apnea hypopnea index.

Subgroup analyses were conducted in 4 subgroups divided according to the snoring percentage of each participant in the test dataset such as subgroups 0 to 24.9% (n=32, 12,800 epochs), 25% to 49.9% (n=13, 5200 epochs), 50 to 74.9% (n=5, 2000 epochs) and 75 to 100% (n=5, 2000 epochs). The observed ground truth snoring percentage of each subgroup was 4.8%, 38.7%, 59.8% and 78.1%, respectively and the predicted snoring of each subgroup was 8.1%, 44.5%, 62.1% and 84.5%, respectively. The AHI of each subgroup was 4.6, 16.8, 33.6 and 34.1/hr, respectively (Table 2).

The sensitivity of the snoring prediction model was 0.838, 0.859, 0.940 and 0.940 for the 4 different subgroups, respectively. The specificity of the model was 0.952, 0.816, 0.717 and 0.777 for the 4 subgroups, respectively (Figure 2). A representative subject showing epoch-by-epoch comparison between observed and predicted snoring is demonstrated in Figure 3.

Correlation Between Observed and Predicted Snoring Percentage

The mean percentage of ground truth observed and predicted snoring of 55 participants in the training dataset was 24.5 ±26.6% and 28.6±28.3%, respectively. The correlation coefficient between observed and predicted snoring ratio was 0.97 (95% CI, 0.95–0.99). The mean absolute error and root mean squared error were 5.09% (95% CI, 3.42–6.71) and 7.96% (95% CI, 5.82–9.93), respectively.

Table 2 Epoch-by-Epoch Prediction Performance and Characteristics According to the Snoring Percentage of Participants in the Test Dataset

	Snoring Subgroup			
	0–24.9% (n=32, 12,800 epochs)	25–49.9% (n=13, 5200 epochs)	50–74.9% (n=5, 2000 epochs)	75–100% (n=5, 2000 epochs)
Epoch-by-epoch prediction				
Sensitivity	0.838	0.859	0.940	0.940
Specificity	0.942	0.816	0.717	0.777
Subgroup characteristics				
Observed snoring (%), mean (SD)	4.8 (7.0)	38.7 (7.0)	59.8 (7.3)	78.1 (9.8)
Predicted snoring (%), mean (SD)	8.1 (10.1)	44.5 (10.6)	62.1 (10.4)	84.5 (8.8)
Age (years), mean (SD)	44.4 (16.4)	49.0 (16.1)	45.6 (13.3)	49.0 (5.1)
Males, n (%)	11 (34.3%)	8 (61.5%)	4 (80%)	5 (100%)
BMI (Kg/m ²), mean (SD)	24.3 (4.0)	25.3 (5.0)	26.1 (2.1)	28.3 (2.7)
AHI (/hr), mean (SD)	4.6 (8.6)	16.8 (9.8)	33.6 (24.3)	34.1 (10.6)

Abbreviations: SD, standard deviation; BMI, body mass index; AHI, apnea hypopnea index.

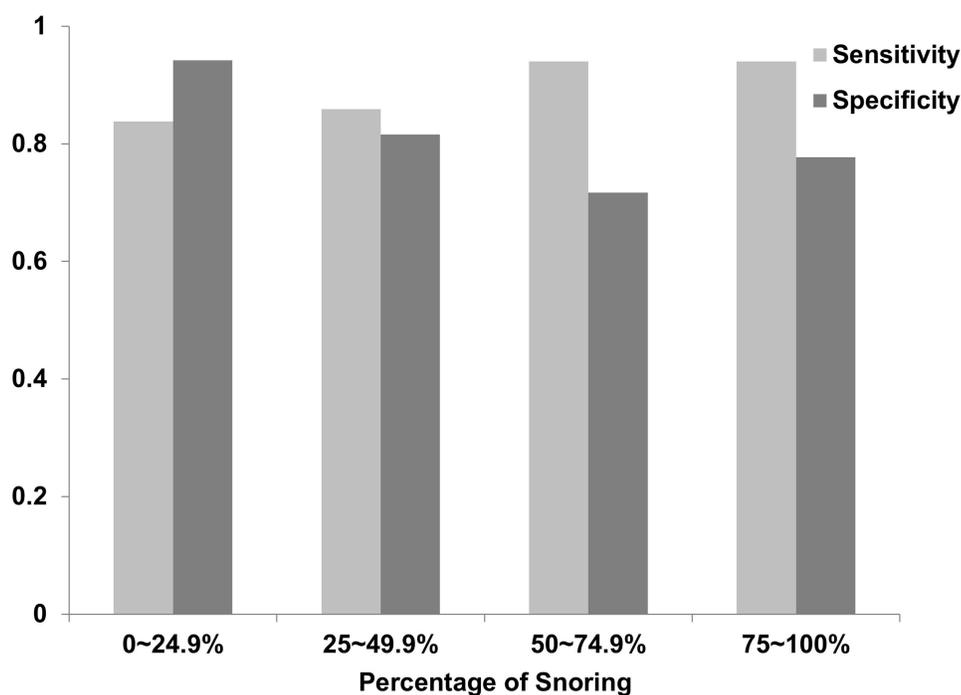


Figure 2 Epoch-by-epoch performance of snoring prediction in 4 subgroups divided according to the snoring percentage of each participant of the test dataset.

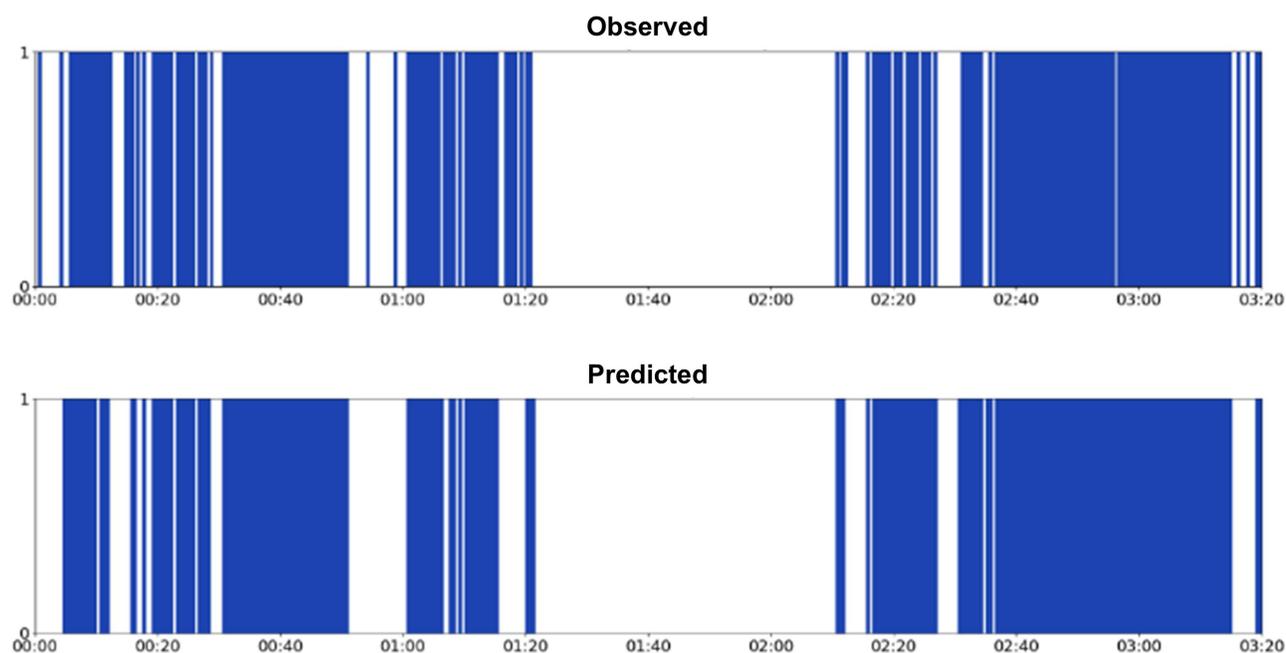


Figure 3 A representative subject showing epoch-by-epoch comparison between observed and predicted snoring. Upper and lower panels show observed and predicted snoring, respectively. The highlighted blue lines demonstrate snoring epochs.

The correlation analyses were performed in two subgroups divided according to AHI of each participant of the test dataset such as subgroups $AHI < 15/hr$ ($n=35$) and $AHI \geq 15/hr$ ($n=20$). The correlation coefficient of the two subgroups was 0.97 (95% CI, 0.95–0.99) and 0.94 (95% CI, 0.86–0.98), respectively (Figure 4). The mean absolute error and root mean squared error were 4.27% (95% CI, 2.56–6.05) and 6.88% (95% CI, 4.45–9.26), respectively for $AHI < 15/hr$, and 6.53% (95% CI, 3.59–9.49) and 9.57% (95% CI, 5.47–12.96), respectively for $AHI \geq 15/hr$.

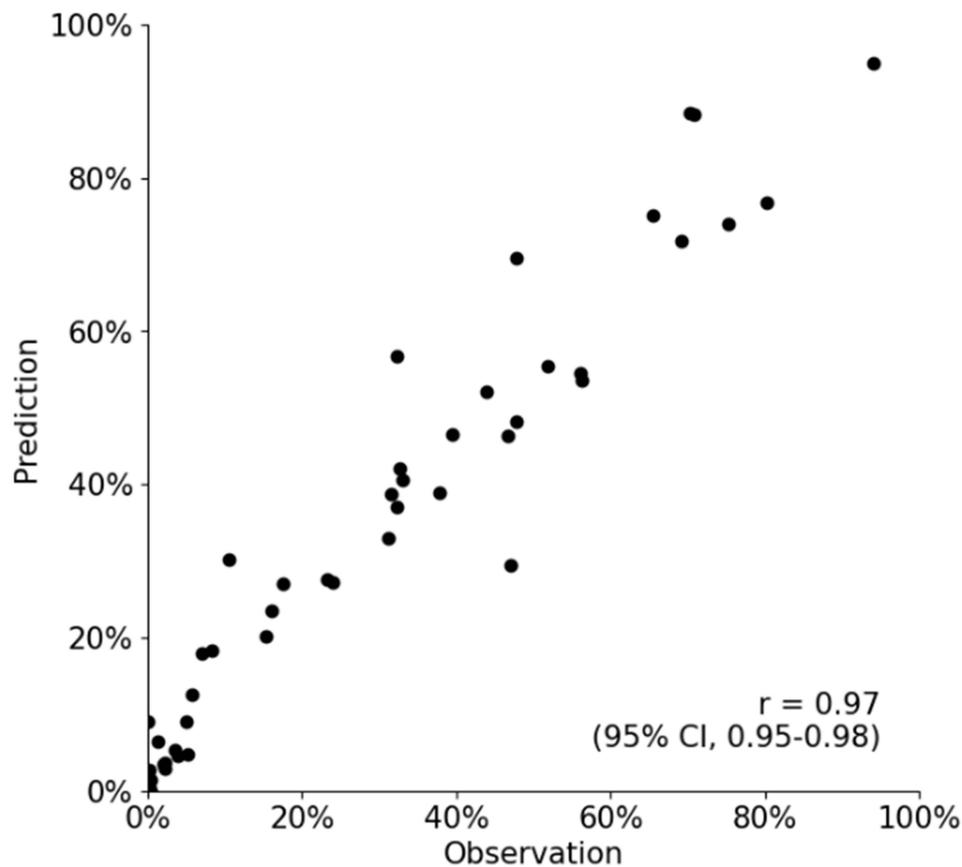


Figure 4 Correlation between observed snoring ratio and predicted snoring ratio in the test dataset.

Discussion

Snoring not only diminishes the sleep quality of bed partners but can also negatively impact the snorer's own sleep quality. It is, therefore, an important condition in itself and a key symptom of OSA, making its early diagnosis critical. Developing technology that allows for the easy and accurate diagnosis of snoring at home has significant implications. Such advancements can help individuals recognize the severity of their snoring and identify lifestyle habits contributing to it. Many individuals remain unaware of their habitual snoring or its potential consequences, leading to delayed diagnosis and treatment of sleep-disordered breathing conditions. By offering a simple and user-friendly method to track snoring percentage, home-based detection technology could raise awareness and encourage timely medical consultations, reducing the burden of undiagnosed OSA. From a clinical perspective, home-based snoring monitoring could serve as a first-line screening tool to help individuals and healthcare providers recognize patterns indicative of potential OSA risk. This technology could be particularly valuable in remote patient monitoring, allowing individuals to assess snoring burden before undergoing more comprehensive diagnostic evaluations such as polysomnography or home sleep apnea testing.

The development of a real-time snoring prediction model using deep learning algorithms may represent a significant advancement in the field of sleep medicine, offering a promising avenue for non-invasive, accessible monitoring of sleep-related breathing disorders. Our study, incorporating training datasets obtained in a home setting utilizing smartphone recordings of sleep-breathing audios, aimed to validate the accuracy and reliability of this innovative approach. With a high epoch-by-epoch sensitivity and specificity and correlation coefficient between observed and predicted snoring, our model exhibited remarkable prediction performance in quantifying snoring patterns. These findings underscore the feasibility and efficacy of utilizing smartphone recordings for real-time snoring prediction, offering a convenient, cost-effective alternative to traditional diagnostic methods.

While previous studies have explored the prediction of OSA using various home devices, including smartphones, the emphasis has primarily been on diagnosing OSA based on AHI. A cross-sectional study of 423 subjects using audio recordings during sleep using a smartphone demonstrated accuracies of 82.3% for an AHI threshold of 15/hr.¹⁷ Another prospective study of 101 participants using home audio recordings showed accuracies 93.3% for an AHI level of 15/hr.¹⁸

With regard to prediction of snoring in particular, previous studies have utilized sound samples, audio maps, or recordings obtained with microphones. In a recent study of a multi-branch convolutional neural network for classifying snoring and non-snoring events based on publicly available dataset consisting of 1000 one-second sound samples with an accuracy of 99.5% was achieved.¹⁹ The one-second recordings were classified into snoring (500 samples) and non-snoring (500 samples). In contrast, our study analyzed 30-second audio epochs recorded in real-world home and hospital settings, capturing realistic sleep conditions. Each epoch was annotated by two trained human labelers, requiring consensus for snoring classification, ensuring high labeling accuracy. Our dataset is substantially larger, encompassing 85,600 30-second epochs from 214 participants. Our recordings reflect actual sleep environments, including various background noises, enhancing the generalizability of our findings. The use of 30-second epochs allows for a more detailed analysis of breathing patterns over time, unlike the one-second snippets used in the above dataset. Consensus-based human labeling provides robust ground truth data, reducing misclassification errors, and the larger and more diverse dataset improves the reliability and applicability of our deep learning model.

Another study based on the combination of convolutional neural network, deep neural network, and Long and Short memory network, analyzed various descriptors extracted from audio maps, and showed that the Mel-spectrogram can better distinguish the differences between snoring and non-snoring sound segments than other descriptors.²⁰ They utilized 4600 minutes of audio recordings sourced from YouTube, which were segmented into 1-minute intervals and annotated by one of four annotators. Ultimately, 1147 segments containing 18,309 snore events were included in their dataset. In contrast, our study employed 30-second epochs annotated by two trained human labelers, ensuring a high degree of labeling accuracy. Unlike their dataset, which consisted of selected 1-minute segments, our approach captures more detailed and granular data across diverse environments, including both home and hospital settings, thereby enhancing real-world applicability. Furthermore, our dataset is substantially larger, comprising 85,600 epochs compared to their 1147 segments, offering a more robust and comprehensive foundation for training deep learning models.

In a study of a hybrid convolutional neural network model for the automatic snoring prediction using 88 snoring recordings obtained using a high-resolution microphone, with sensitivity of 89.7% and specificity of 88.5%.²¹ The referenced study and our study share the common feature of manually annotating snoring. However, while they used a portable PSG without EEG, our study employed a Level 2 full PSG, enabling the precise association of respiratory sounds with sleep stages. Additionally, we used smartphones in realistic home environments, mirroring real-world conditions while maintaining robust data quality. Furthermore, our dataset is significantly larger (85,600 epochs vs 5441 episodes). Another study analyzed the accuracy of snoring rates based on 201 snoring records of 11 patients on snoring prediction application for smartphones, showed a mean snoring prediction accuracy rate of 95% and a correlation coefficient between predicted and observed snoring rates of 0.91.²² However, their study sample size was much smaller than ours.

Our snoring prediction model demonstrated robust performance across various metrics. When evaluated on 22,000 epochs from the test dataset, the model achieved a sensitivity of 89.8% and a specificity of 91.3% in detecting snoring events. The strong correlation between observed and predicted snoring percentages (correlation coefficient of 0.97) across both AHI subgroups further highlights the model's accuracy. Notably, the mean absolute error was 5.09%, reflecting precise prediction capabilities. Given the significant correlation between snoring and AHI,²³ our results underscore the potential of snoring analysis in the early detection of OSA. These performance metrics suggest that the model is well-suited for clinical applications, particularly in home-based monitoring and early screening for sleep-related breathing disorders. By effectively detecting snoring events while minimizing false positives, the model ensures reliable identification of individuals at risk of sleep-disordered breathing. Given the natural night-to-night variability in snoring, the observed error margins are unlikely to significantly impact clinical decision-making, further reinforcing the model's potential for real-world implementation.

Subgroup analyses, categorized by participants' snoring percentages, revealed that as the proportion of snoring increased, sensitivity improved (ranging from 83.8% to 94.0%), while specificity decreased (from 95.2% to 77.7%). This pattern suggests that the model maintains relatively high sensitivity in detecting snoring even among individuals with lower snoring

frequencies, albeit with a trade-off in specificity. The subgroup analysis based on snoring percentage provides important clinical insights into the potential relationship between snoring burden and sleep-disordered breathing severity. While primary snoring is generally considered distinct from OSA, higher snoring percentages have been associated with increased upper airway resistance, more frequent oxygen desaturation, and greater sleep fragmentation, which are key features of OSA. Although our study did not directly evaluate OSA severity, the findings suggest that individuals with higher snoring percentages may warrant further clinical assessment for possible sleep-disordered breathing.

One limitation of our study is the reliance on manual annotation of snoring by human labelers. While this approach ensures high labeling accuracy through a consensus method, it reflects the absence of a universally accepted academic definition of snoring.²⁴ Our use of human scoring aimed to address this gap and provide a reliable standard for model training. A systematic review on snoring detection highlighted significant inconsistencies in measurement methodologies across studies.²⁴ Different studies have employed various sensor types, including microphones, piezoelectric sensors, and nasal transducers, with varying intensity thresholds and annotation methods, leading to conflicting results. Additionally, concerns about ambient noise and microphone placement variability have raised questions about the reliability of smartphone-based snoring detection. These methodological differences must be considered when interpreting the findings of snoring detection models, emphasizing the need for standardized measurement criteria and further validation across diverse populations. Additionally, the dataset, although diverse, may not fully capture variations in snoring patterns across different populations, such as those with distinct ethnic or health profiles.²⁵ Another potential limitation is the relatively small sample size in the higher snoring percentage subgroups, which is an inherent challenge given that snoring distribution varies among individuals, making it difficult to recruit participants with very high snoring percentages. However, despite this limitation, the model demonstrated stable performance across all subgroups, suggesting that the small sample size did not significantly affect its reliability. Furthermore, while the number of participants in these subgroups was low, the total number of epochs analyzed remained substantial, helping to mitigate concerns related to statistical instability. Nevertheless, we acknowledge this limitation and plan to further validate our findings by expanding our dataset with additional home-based recordings in future studies. The age and BMI distributions, which are somewhat concentrated within middle-aged and overweight individuals, present a potential limitation. The mean age of participants suggests that younger individuals are underrepresented, and the BMI range is skewed toward overweight participants, potentially limiting the generalizability of the model to populations with different demographic profiles. However, the relatively balanced gender distribution helps mitigate concerns about sex-based bias. To enhance the model's applicability across broader populations, future studies will focus on expanding the dataset to include a wider range of ages and BMI values. By incorporating more participants from diverse demographic and clinical backgrounds, we aim to ensure that the model remains robust across varying physiological and acoustic characteristics. The smartphone-specific recording setup may also limit reproducibility across devices or varying home environments.

While this study demonstrates the feasibility of deep learning-based snoring detection using smartphone audio, several avenues for future research remain. One key area for advancement is the integration of additional physiological signals, such as oxygen saturation, heart rate variability, and EEG-derived sleep parameters. By incorporating multimodal data, future models could enhance differentiation between primary snoring and sleep-disordered breathing conditions, including OSA. Developing a mobile application that allows individuals to monitor their snoring patterns, receive feedback, and securely share data with healthcare providers could facilitate early intervention and personalized sleep health management. Future research should explore how snoring detection can be incorporated into telemedicine platforms, digital health records, or integrated with wearable sleep monitors to support remote patient monitoring.

In conclusion, this study highlights the potential of deep learning combined with smartphone recordings as a practical and accessible tool for snoring prediction and analysis. By leveraging real-world data from diverse environments and employing rigorous annotation methods, our model achieves high accuracy in detecting snoring and estimating its prevalence. This not only provides an objective metric for assessing snoring severity but also lays the groundwork for integrating snoring analysis into broader sleep health management frameworks. The real-time, epoch-by-epoch prediction capability of the model further enhances its applicability, offering possibilities for immediate feedback and personalized interventions. As smartphones continue to proliferate globally, our approach bridges the gap between advanced sleep disorder diagnostics and everyday accessibility, paving the way for its adoption in clinical practice and home-based monitoring to improve sleep health outcomes.

Highlights

1. A novel application of real-time, epoch-by-epoch snoring prediction with high sensitivity and specificity, offering potential for immediate feedback and personalized intervention.
2. A dataset collected in diverse sleep environments, including home settings, enhancing the generalizability and practical relevance of our findings.
3. The demonstration of a strong correlation between observed and predicted snoring percentages, underscoring the reliability of our approach.

Data Sharing Statement

Due to privacy and ethical concerns, the data supporting this study cannot be made publicly available. Mel-spectrogram data could potentially be reverse-engineered to reconstruct participants' voices, posing privacy risks.

Ethics Statement

This study was conducted following ethical approval from the Seoul National University Bundang Hospital Institutional Review Board (IRB No. B-2205-755-308). Informed consent was obtained from all individual participants included in the study. The research was performed in accordance with the ethical standards as laid down in the 1964 Declaration of Helsinki and its later amendments.

Acknowledgment

This work was partly supported by SNUBH grant #14-2024-0025.

Author Contributions

Joonki Hong: Conceptualization, Formal Analysis, Writing, Methodology

Seung Koo Yang: Conceptualization, Investigation, Visualization, Writing – Original Draft

Seunghun Kim: Writing, Software

Sung-Woo Cho: Conceptualization, Validation

Jayoung Oh: Data Curation, Investigation

Eun Sung Cho: Data Curation, Resources, Software

In-Young Yoon: Supervision, Validation

Dongheon Lee: Conceptualization, Project Administration

Jeong-Whun Kim (corresponding author): Methodology, Supervision, Validation, Funding Acquisition, Writing – Review & Editing

All authors have drafted or substantially revised the manuscript, reviewed and agreed on all versions of the article before submission, during revision, and in the final stage before publication, and approved the journal to which the article was submitted. Additionally, all authors agree to be accountable for the content of the manuscript and any significant changes introduced at the proofing stage.

Disclosure

There is no conflict of interest regarding the publication of this article.

References

1. Enright PL, Newman AB, Wahl PW, Manolio TA, Haponik EF, Boyle PJ. Prevalence and correlates of snoring and observed apneas in 5201 older adults. *Sleep*. 1996;19(7):531–538. doi:10.1093/sleep/19.7.531
2. Campos AI, Garcia-Marin LM, Byrne EM, Martin NG, Cuellar-Partida G, Renteria ME. Insights into the aetiology of snoring from observational and genetic investigations in the UK biobank. *Nat Commun*. 2020;11(1):817. doi:10.1038/s41467-020-14625-1
3. Lofaso F, Coste A, Gilain L, Harf A, Guilleminault C, Goldenberg F. Sleep fragmentation as a risk factor for hypertension in middle-aged nonapneic snorers. *Chest*. 1996;109(4):896–900. doi:10.1378/chest.109.4.896
4. Zhang Y, Zhang T, Xia X, et al. The relationship between sleep quality, snoring symptoms, night shift and risk of stroke in Chinese over 40 years old. *Front Aging Neurosci*. 2023;15:1134187. doi:10.3389/fnagi.2023.1134187

5. Sateia MJ. International classification of sleep disorders-third edition: highlights and modifications. *Chest*. 2014;146(5):1387–1394. doi:10.1378/chest.14-0970
6. Deeb R, Judge P, Peterson E, Lin JC, Yaremchuk K. Snoring and carotid artery intima-media thickness. *Laryngoscope*. 2014;124(6):1486–1491. doi:10.1002/lary.24527
7. Lee SA, Amis TC, Byth K, et al. Heavy snoring as a cause of carotid artery atherosclerosis. *Sleep*. 2008;31(9):1207–1213. doi:10.1093/sleep/31.9.1207
8. Meira ECM, Soca R, Kryger M. How much is too much after all? Primary snoring as a remaining unsolved issue. *J Clin Sleep Med*. 2020;16(6):991. doi:10.5664/jcs.m.8442
9. Blumen M, Quera Salva MA, d'Ortho MP, et al. Effect of sleeping alone on sleep quality in female bed partners of snorers. *Eur Respir J*. 2009;34(5):1127–1131. doi:10.1183/09031936.00012209
10. Leung AK, Robson WL. The ABZzzzs of snoring. *Postgrad Med*. 1992;92(3):217–222. doi:10.1080/00325481.1992.11701451
11. Smith DL, Gozal D, Hunter SJ, Kheirandish-Gozal L. Frequency of snoring, rather than apnea-hypopnea index, predicts both cognitive and behavioral problems in young children. *Sleep Med*. 2017;34:170–178. doi:10.1016/j.sleep.2017.02.028
12. Liu Y, Peng T, Zhang S, Tang K. The relationship between depression, daytime napping, daytime dysfunction, and snoring in 0.5 million Chinese populations: exploring the effects of socio-economic status and age. *BMC Public Health*. 2018;18(1):759. doi:10.1186/s12889-018-5629-9
13. Yu S, Guo X, Li G, Yang H, Sun Y. Influence of snoring on the incidence of metabolic syndrome: a community-based prospective cohort study in rural northeast China. *J Clin Med*. 2023;12(2):217. doi:10.3390/jcm12020447
14. Tran HH, Hong JK, Jang H, et al. Prediction of sleep stages via deep learning using smartphone audio recordings in home environments: model development and validation. *J Med Internet Res*. 2023;25:e46216. doi:10.2196/46216
15. Cohen JF, Korevaar DA, Altman DG, et al. STARD 2015 guidelines for reporting diagnostic accuracy studies: explanation and elaboration. *BMJ Open*. 2016;6(11):e012799. doi:10.1136/bmjopen-2016-012799
16. Le VL, Kim D, Cho E, et al. Real-time detection of sleep apnea based on breathing sounds and detection reinforcement using home noises: algorithm development and validation. *J Med Internet Res*. 2023;25:e44818. doi:10.2196/44818
17. Cho SW, Jung SJ, Shin JH, Won TB, Rhee CS, Kim JW. Evaluating prediction models of sleep apnea from smartphone-recorded sleep breathing sounds. *JAMA Otolaryngol Head Neck Surg*. 2022;148(6):515–521. doi:10.1001/jamaoto.2022.0244
18. Han SC, Kim D, Rhee CS, et al. In-home smartphone-based prediction of obstructive sleep apnea in conjunction with level 2 home polysomnography. *JAMA Otolaryngol Head Neck Surg*. 2024;150(1):22–29. doi:10.1001/jamaoto.2023.3490
19. Dong H, Wu H, Yang G, Zhang J, Wan K. A multi-branch convolutional neural network for snoring detection based on audio. *Comput Methods Biomech Biomed Engin*. 2024;27(2):1–12. doi:10.1080/10255842.2024.2317438
20. Jiang Y, Peng J, Zhang X. Automatic snoring sounds detection from sleep sounds based on deep learning. *Phys Eng Sci Med*. 2020;43(2):679–689. doi:10.1007/s13246-020-00876-1
21. Li R, Li W, Yue K, Zhang R, Li Y. Automatic snoring detection using a hybrid 1D-2D convolutional neural network. *Sci Rep*. 2023;13(1):14009. doi:10.1038/s41598-023-41170-w
22. Chiang JK, Lin YC, Lin CW, Ting CS, Chiang YY, Kao YH. Validation of snoring detection using a smartphone app. *Sleep Breath*. 2022;26(1):81–87. doi:10.1007/s11325-021-02359-3
23. Chiang JK, Lin YC, Lu CM, Kao YH. Correlation between snoring sounds and obstructive sleep apnea in adults: a meta-regression analysis. *Sleep Sci*. 2022;15(4):463–470. doi:10.5935/1984-0063.20220068
24. Kim SG, Cho SW, Rhee CS, Kim JW. How to objectively measure snoring: a systematic review. *Sleep Breath*. 2024;28(1):1–9. doi:10.1007/s11325-023-02865-6
25. O'Connor GT, Lind BK, Lee ET, et al. Variation in symptoms of sleep-disordered breathing with race and ethnicity: the sleep heart health study. *Sleep*. 2003;26(1):74–79. doi:10.1093/sleep/26.1.74

Nature and Science of Sleep

Publish your work in this journal

Nature and Science of Sleep is an international, peer-reviewed, open access journal covering all aspects of sleep science and sleep medicine, including the neurophysiology and functions of sleep, the genetics of sleep, sleep and society, biological rhythms, dreaming, sleep disorders and therapy, and strategies to optimize healthy sleep. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <https://www.dovepress.com/nature-and-science-of-sleep-journal>

Dovepress
Taylor & Francis Group